On Clause and Intonational Phrase in Japanese:
The Syntactic Grounding of Prosodic Constituent Structure

Elisabeth Selkirk

University of Massachusetts Amherst

Abstract: This paper reviews evidence from recent research on Japanese sentence prosody, in particular Kawahara and Shinya (2008) on coordinated clauses and Kubo (1989) et seq, Deguchi and Kitagawa (2002) et seq, Ishihara (2002) et seq and Hirotani (2003) et seq on matrix and embedded *wh*-questions, which suggests that syntactic clauses correspond to a domain for certain of the phonological and phonetic phenomena that define the intonational patterns of Japanese sentences. The finding that there is a clause-grounded intonational phrase domain of phonological representation in Japanese is predicted by a universal theory of the prosodic hierarchy as grounded in a universal theory of the syntax-phonology interface (Selkirk 2005). This paper lays out a new universal Match theory of the syntax-prosodic constituency interface, according to which designated syntactic constituent types are called on to match up with corresponding prosodic constituent types. Match theory may be construed as a component of the theory of Spell-Out in minimalist phase theory (Chomsky 2001). The data also shows that recursive intonational phrase structure is produced when the universal Match Clause constraint is satisfied on nested clausal domains. This is expected if indeed the availability of recursive prosodic structure derives from the organization of syntactic structure, through the agency of constraints on the syntax-phonology interface. This theory that prosodic constituent structure above the foot is syntactically grounded thus meshes well with the Itô and Mester (2007) claim that the prosodic hierarchy repertoire is universal and highly restricted and that recursivity in prosodic structure is systematic.

Key words: syntax-phonology interface, universal prosodic hierarchy, prosodic recursion, intonational phrase, Japanese

1. Beyond the standard theory of prosodic constituent structure

A prosodic structure consisting of both prosodic constituency and prosodic prominence is widely assumed to form the basic infrastructure of the phonological representation with respect to which phenomena of sentence phonology and phonetics are defined in any language. For example, both prosodic phrasing and prosodic prominence contribute to determining the phonological representation of tone in the sentence[1] as well as the tonal pitch scaling of the intonational contours

---

[1] In general, phonological markedness constraints call for tones to be located on prosodic heads/prominences or at prosodic constituent edges (Yip 2002); this may lead to the epen-

that are produced in the phonetic interpretation of that phonological representation[2]. In what one might refer to as the 'standard theory', prosodic constituent representation is defined as a well-formed labeled tree or bracketing, but one which has two fundamental properties that distinguish it from syntactic constituent structure representations—the prosodic hierarchy and strict layering (Selkirk 1978/1981a, 1981b, 1986: Nespor and Vogel 1986; Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988; and others). The *prosodic hierarchy* is the name for an ordered set of prosodic category types. These prosodic category types

---

thesis of tones into these positions. In the case of lexical or morphemic tones, satisfaction of these markedness constraints may lead to displacement into these positions, or, if the lexical or morphemic tone remains in situ, there may be effects on the prosodic structure of the sentence. Languages may differ in the level of prosodic prominence or the level of prosodic constituency that is relevant is establishing such tone-prosodic structure relations (see Hyman 2006, for example).

A classic example of the role for prosodic constituency in determining the distribution of tone in the sentence is provided by the L% and H- tones that produce the 'initial lowering' effect in Tokyo Japanese. These tones are neither morphemic or lexical, but rather appear by default at the edge of the prosodic constituent that has been referred to the minor phrase or accentual phrase in work on Japanese (Haraguchi 1977, Poser 1984, Kubozono 1988, 1993, Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, inter alia). They do not form part of the morphosyntactic representation of the sentence. Japanese lexical pitch accents, on the other hand, form part of the input to the phonology, where their presence in the representation may have an effect on the organization of the sentence into minor phrases: the restriction of one pitch accent per minor phrase in Japanese is plausibly a reflection of a requirement that a pitch accent be associated to an abstract prosodic prominence that is defined within each minor phrase. (See footnote 3 on the notion 'minor phrase.')

Bengali, by contrast, shows both a predictable phonological phrase-edge H tone and a predictable L* pitch accent that appears on the prosodic head prominence of a phonological phrase (Hayes and Lahiri 1991, Selkirk 2007). In Cairene Arabic, for example, it is the prosodic prominence/main stress of every prosodic word that is necessarily associated with a predictable pitch accent (Hellmuth 2007).

[2] The role for prosodic constituent structure in the phonetic interpretation of sentences of Japanese has been richly documented (Poser 1984; Kubozono 1988, 1989, 1993; Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988; Sugahara 2002, 2003/2005a, 2005b; Ishihara 2008, Kawahara and Shinya 2008), as it has in other languages (see e.g. Jun 2005). As for a role for prosodic structure prominence in phonetic interpretation, there is little crosslinguistic research on this matter at levels above prosodic word. It seems likely that the notion of prosodic prominence—by which I mean the abstract notion of prosodic constituent headedness (see Selkirk 1980 on the headedness of prosodic word and feet and Selkirk 2007 on prosodic head prominence above prosodic word)—will emerge as a significant factor in cross-linguistic phonetic research once it is recognized that such prosodic prominence may be differently interpreted by the phonetics in different languages. Japanese, for example, shows no durational reflex of prosodic headedness at any level, but arguably shows a reflex of this head prominence in the F0 domain (see e.g. Sugahara 2003/2005a), while English shows both durational and pitch-related reflexes of prosodic structure (see, e.g. Katz and Selkirk 2009).

constitute possible node labels for prosodic structures and in the standard view are stipulated by phonological theory.

(1)   Prosodic category types of a commonly posited prosodic hierarchy

Intonational Phrase (ι)
Phonological Phrase (φ)
Prosodic Word (ω)
Foot
Syllable

In the standard theory of prosodic constituent structure no inherent relation is assumed to exist between the prosodic category types found in phonological representations and the category types of syntactic representation. Categories of the prosodic hierarchy are simply assumed to be primitives of phonological theory.

There is debate about just how many prosodic categories above the foot may play a role in language, and about whether they are universally attested. Some proposals posit a higher Utterance-level constituent that may group together intonational phrases; some proposals posit two different levels of phonological phrase, a higher level termed Major Phrase or Intermediate Phrase and a lower level termed Minor Phrase or Accentual Phrase. In recent work, Itô and Mester (2007, 2009b) have proposed that just one category of phonological phrase (φ) should be posited; they argue that the major/minor phrase distinction is a syntagmatic notion, a function of the status of an instance of φ as maximal or minimal in the overall prosodic structure of the sentence[3]. Below, it will be argued that the prosodic category Utterance plays no role in Japanese, or universally. So it does seem quite probable that phonological and phonetic evidence motivates just the three levels above the foot that are displayed in (1): ι ('intonational phrase'), φ ('phonological phrase'), and ω ('prosodic word'). Yet the standard theory, which simply stipulates the prosodic categories that are primitives of phonological representation, does not provide an explanation for this limitation in the repertoire of supra-foot category types. This is a shortcoming that is overcome by the hypothesis put forward in the present paper that the prosodic hierarchy of sentence phonology—the repertoire of prosodic categories that group words and sequences of words—is syntactically grounded, with its source in the theory of the interface between syntactic and prosodic structure.

It should be noted that at this point in time it is not possible to point to any phonological or phonetic properties which hold universally of any of the particular supra-foot prosodic category types that have been posited in the context of the standard theory of prosodic constituency. Indeed, to find out if there are in fact such properties, it is necessary to have an independent basis for identifying constituents of these different types. A syntactically grounded theory of the

---

[3] For Itô and Mester (2007), a φ which is *maximal* is dominated by no other φ, and a φ which is *minimal* dominates no other φ. The major phrase of earlier accounts of Japanese is re-analyzed as a maximal φ. The minor phrase would correspond to a minimal φ.

prosodic hierarchy repertoire like that presented in section 2 provides just such a basis; a theory like this makes it possible to undertake crosslinguistic investigation of the phonological and phonetic properties of intonational phrases, phonological phrases and prosodic words with a common ground for identifying constituents of the different prosodic category types.

In the standard theory of prosodic constituency the nature of domination relations within a prosodic constituent structure is also determined by phonological theory. The *strict layer hypothesis* (SLH) is the name given to the idea that a prosodic structure representation is arranged according to the ordered set of categories in the prosodic hierarchy and shows no recursivity at all. According to the SLH, a constituent of category-level *n* in the prosodic hierarchy immediately dominates only constituents at category-level *n-1* in the hierarchy. The prosodic representation in (2a) satisfies the SLH while the representation in (2b) does not.

(2)   a.

b.
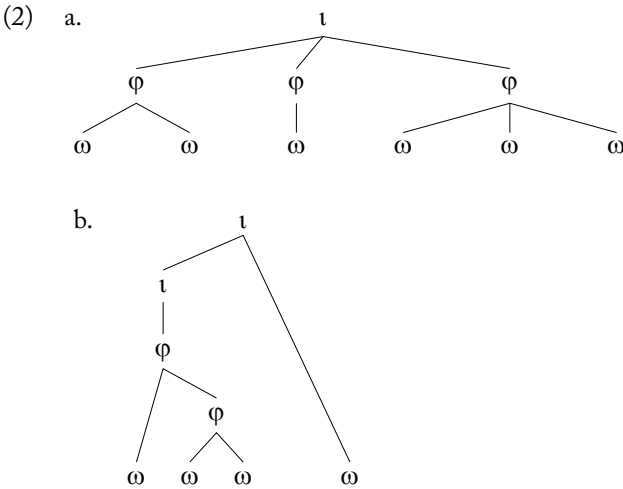
The strict layer hypothesis constitutes a purely phonological theory of the formal relations holding between constituents of the different prosodic category types in a prosodic structure. If representations like (2a) that obey the strict layer hypothesis were the rule in phonology, phonological representations would indeed differ fundamentally from syntactic representations. In (2b) are instances of a configuration in which a constituent of a particular prosodic category type dominates another of the same category type (ι/ι and φ/φ), namely instances of recursivity; there is also an instance of a configuration in which a constituent of category level *n* in the prosodic hierarchy immediately dominates a constituent of category level n-2 (ι/ω). These configurations both represent violations of the strict layer hypothesis. Yet they are arguably attested in language. Evidence has emerged over the years that prosodic structure representations that fail to obey the SLH are not at all uncommon[4]. Selkirk (1996) proposes that the SLH be factored into a set of universal

---

[4]   On recursivity in phonological domain structure, see e.g. Ladd (1986) et seq, Truckenbrodt

optimality theoretic prosodic structure markedness constraints in which the constraints Nonrecursivity and Exhaustivity, which are violated by the configurations described in (2b), are understood to be violable. Yet this revision to the phonological theory of prosodic constituent representation says nothing about a further generalization which has been emerging, namely that the domain structure for sentence phonology and phonetics tends to be much closer to syntactic constituent structure than the strict layer hypothesis would predict[5]. The hypothesis explored in the current paper is that the prosodic constituent structure relations observed in phonological representations are largely the consequence of constraints on the interface between syntax and prosodic constituency: violations of the strict layer hypothesis come about in the first place because of the influence of the syntactic constituent structure of the sentence on phonological representation.

So it seems, as Itô and Mester (2007) have emphasized, that two fundamental aspects of the general nature of prosodic constituent structure are at variance with the standard theory of prosodic structure: the systematic tendency towards recursivity and nonexhaustivity in prosodic representation (contra the strict layer hypothesis) and the repertoire of category types in the prosodic hierarchy above the foot as consisting only of ɩ ('intonational phrase'), φ ('phonological phrase'), and ω ('prosodic word'). The present work shares with that of Itô and Mester the goal of providing a restrictive, universal, theory of the prosodic hierarchy repertoire and of giving recursivity the place it is due in a theory of prosodic constituent representation. The contention of this paper is that these two general properties of prosodic structure representation above the foot can be derived from the theory of the syntax-prosodic structure interface and are not independent stipulations of phonological theory. The paper builds on an earlier proposal by Selkirk (2005) that the prosodic hierarchy repertoire is grounded in the interface and on the Itô and Mester (2006) proposal that recursivity in the prosodic structure of compounds in Japanese derives, in part, from their recursive morphosyntactic structure.

Section 2 of the paper presents a theory of the syntax-phonology interface which has the property that it explains both the prosodic hierarchy repertoire above the foot, and the systematic availability of recursivity in prosodic structure.

(1999, 2005), Frota (2000), Dobashi (2003), Féry and Truckenbrodt (2005), Wagner (2005, to appear), Itô and Mester (2006, 2007), Pak (2008). Exhaustivity violations are common within prosodic word (see eg. Hayes 1995 on syllables not dominated by feet in word stress systems, and Itô and Mester 1992/2003 on the allowability of syllable sisters to feet in Japanese prosodic morphology). They are the stock in trade of cases of prosodic clisis (Inkelas 1990, Selkirk 1996, Peperkamp 1997, Itô and Mester 2007, 2009a, Werle 2009) where they accompany instances of ω recursivity.

[5] Like the Itô and Mester (2006) analysis of compound prosody in Japanese, the Selkirk (2009) analysis of tonal phenomena in Tsonga shows that recursivity in prosodic representations like that in (2b) mirrors syntactic structure. In theories where phonology has direct access to syntactic structure Ishihara (2004), Wagner (2005), Pak (2008), Tokizaki (2008), it has never been assumed that the domain structure for phonology would *not* display the properties of syntactic structure.

Section 3.1 is devoted to a review of data from Japanese coordinate structures (Kawahara and Shinya 2008) that motivates the presence of the prosodic category intonational phrase as a reflex of syntactic clause structure. Section 3.2 proposes that the notion 'clause' be identified as a complement of one of the syntactic functional heads of the Rizzi (1997) 'complementizer' layer, which includes Force^0, as well as lower instances of Comp^0. In section 3.3 issue is taken with Kawahara and Shinya's contention that there is evidence from Japanese for postulating a prosodic category Utterance which would stand above intonational phrase in the ordered set of universal prosodic categories. In section 3.4 evidence from the prosody of embedded *wh*-questions in the Fukuoka and Tokyo dialects of Japanese is used to support the notion that embedded clauses correspond to intonational phrases in Japanese, creating recursive intonational phrase structures. And section 3.5 briefly reviews arguments against alternatives to the Match theory of the interface that is presented in this paper.

2. Matching up syntactic and prosodic constituents

A new theory of universal constraints on the interface between syntax constituency and prosodic constituency is put forward in this paper (see also Selkirk 2009/ to appear). This theory defines the simplest possible correspondence between constituents of syntactic structure and constituents of prosodic structure, namely a match between the two:

(3)   A Match theory of the syntax-prosodic structure interface
    (i) Match Clause
        A clause in syntactic constituent structure must be matched by a constituent of a corresponding prosodic type in phonological representation, call it ɩ.
    (ii) Match Phrase
        A phrase in syntactic constituent structure must be matched by a constituent of a corresponding prosodic type, in phonological representation, call it φ.
    (iii) Match Word
        A word in syntactic constituent structure must be matched by a constituent of a corresponding prosodic type in phonological representation, call it ω.

The Match constraints in (3) pare syntactic constituent types to the minimum, exploiting the notions clause, phrase and word, which presumably play a role in any theory of morphosyntax. This set of Match constraints embodies the claim that, in the ideal case, the grammar allows the fundamental syntactic distinctions between clause, phrase and word to be reflected in, and retrieved from, the phonological representation of a sentence. In the current theory, then, the prosodic category types ɩ ('intonational phrase'), φ ('phonological phrase'), and ω ('prosodic word') are syntactically grounded, in the sense that they owe their existence to constraints on the syntax-phonology interface: they constitute the ideal phonological

correspondents of the syntactic category types that are singled out by the syntax-phonology interface constraints of the grammar. Rather than a stipulation of the phonological component of grammar, as in the standard theory, the repertoire of prosodic category types above the foot constitutes a phonological mirroring of the syntactic constituent types playing a role at the interface.

The Match theory being entertained here could be called a *constituent faithfulness theory* of the syntax-prosodic structure mapping. In the spirit of optimality theory approaches to faithfulness (e.g. McCarthy and Prince 1995), input-output Match constraints would call for constituency in the input syntactic representation to correspond to (prosodic) constituency in the output phonological representation, and output-input Match constraints would call for surface prosodic constituents to correspond to input syntactic constituents. But whether or not these Match Clause, Match Phrase and Match Word interface constraints are respected in surface phonological representations depends, optimality-theory-wise, on their interaction with phonological markedness constraints. For example, higher ranked prosodic markedness constraints may be responsible for the presence in phonological representation of a prosodic phrase that does not correspond to a syntactic constituent, or for the absence of a prosodic phrase corresponding to a syntactic constituent[6].

A general "match" theory of the syntax-prosodic structure interface that would encompass the three levels of prosodic constituency ι, φ and ω has not before been proposed, though Match theory has precursors in the proposals by Ladd (1986) and Nespor and Vogel (1986) that a clause in the syntax corresponds to an intonational phrase in prosodic structure. Something like Match theory has emerged in recent proposals within minimalist phase theory (Chomsky 2001) which hold that the Spell-Out domains of phases correspond to prosodic phrases and/or that certain phrase types identifiable in terms of phase theory are spelled out as corresponding prosodic constituents (Seidl 2001, Dobashi 2003, Revithiadou and Spyropoulos 2005, in press, Adger 2007, Selkirk 2006, Ishihara 2007, Kratzer and Selkirk 2007, among others).

The present theory makes the claim that in all languages the phonological representation of a sentence is organized into units at the intonational phrase, phonological phrase and prosodic word levels. This claim is entailed by three assumptions: (i) that the sentences of any language are organized into clauses, phrases and words in syntactic representation, (ii) that these three general types of syntactic

---

[6] A binarity constraint requiring that a prosodic phrase consists of at least two prosodic words might rule out single word instances of phonological phrases (Ghini 1993, Selkirk 2000) or intonational phrases (Dehé to appear), that would otherwise be called to correspond to a syntactic phrase or clause by Match Phrase or Match Clause, while a markedness constraint that a phonological phrase may contain at most one accent (cf. footnote 1) may force the analysis of an accented word as a phonological phrase, even when it lacks the status of a syntactic phrase. Such cases are attested in Japanese (cf. Kubozono 1988, 1993, Selkirk and Tateishi 1988).

constituent are required by a universal theory of the syntax-phonology interface to correspond to constituents of phonological representation, as in Match theory, and (iii) the additional, highly plausible assumption that there are no phonological markedness constraints which categorically rule out the presence of intonational phrase, or phonological phrase, or prosodic word in phonological representation. The claim that there is a universally realized prosodic hierarchy is controversial. Jun (2005) explicitly assumes that languages may vary in the repertoire of prosodic categories that appear in phonological representation. Venditti (1997, 2005) construes the major/intermediate phonological phrase as the highest prosodic constituent type in Japanese (and renames it intonational phrase); Venditti, Maekawa and Beckman (2008) concur in this. The working hypothesis here, by contrast, is that when putative cases of the absence of a prosodic constituent of a particular category type in a particular language are reexamined from the perspective of a theory that grounds prosodic constituent structure in syntactic representation, evidence of the presence of that prosodic category type will emerge. As we will see, Kawahara and Shinya (2008) have shown this is the case for the clause-grounded intonational phrase in Japanese. In what follows, this and other evidence from recent research on Japanese[7] will be assembled in support of the specific claim made by the Match theory of the syntax-prosodic structure interface that there is a systematic correspondence relation between clauses in syntactic representation and intonational phrases in phonological representation. With this case made, the prosodic structure of Japanese is entirely in line with the universal prosodic hierarchy hypothesis.

3. Clause and ι ('intonational phrase') in Japanese

Relatively little linguistic research has been undertaken in Japanese over the past decades that would reveal any generalizations about the phonological or phonetic properties of clauses. An investigation of the prosody of coordinated clauses has recently been made by Kawahara and Shinya (2008), though, and recent investigations comparing the prosody of wh-questions in matrix and embedded sentence contexts by a number of scholars including Kubo, Deguchi and Kitagawa, Ishihara, Hirotani and Smith arguably bring additional data to bear on an understanding of Japanese clause prosody. These investigations will be reviewed below. We will see (i) that there is indeed an effect of clause structure on sentence prosody, and (ii) that the data is consistent with the assumption made by interface Match theory that clauses in syntactic representation are mirrored in phonological representation by the presence of corresponding prosodic ι-domains. These clause-grounded prosodic constituents will be referred to as intonational phrases, as is common practice in the literature (Ladd 1986, Nespor and Vogel 1986 and many others)[8]. Data

---

[7] The literature on Japanese sentence prosody that is in Japanese is unfortunately unavailable to the author. It is hoped that this shortcoming will not damage the argument being made.

[8] The terms 'intonational phrase' and 'phonological phrase' as used in the context of the

from phonological constraints on intonational phrase structure in Japanese, such as minimal size requirements and the like, is not yet available, though, so we are not now in a position to use data from Japanese to clinch the argument that it is intonational phrases and not syntactic clauses themselves that constitute the domain over which apparently clause-sensitive phonological and phonetic phenomena are defined[9]. In any case, the evidence shows that, directly or indirectly, clause structure affects the domain structure relevant for the phonology and phonetics of the sentence in Japanese.

3.1. Evidence from conjoined clauses for intonational phrase in Japanese

Given the current state of knowledge of sentence prosody crosslinguistically, it is not possible to assume that intonational phrases, or clauses, necessarily show any phonological or phonetic properties that distinguish them—crosslinguistically—from other constituent types of phonology or syntax. A useful methodology for discovering distinctive prosodic properties of intonational phrases in a particular language would therefore consist in comparing data on the phonetics and phonology of the edges or spans of clauses with data on the phonetics and phonology of the edges or spans of lower order syntactic phrases, in contexts that are appropriately controlled. In their investigation of the intonation of Japanese sentences containing coordinated clauses, Kawahara and Shinya (2008) do just this. They explicitly set out to test the universalist claim that in every language the syntax-phonology interface calls for a clause to match up with a corresponding prosodic constituent—the intonational phrase—and they find five pieces of evidence motivating an appeal to intonational phrase (or to clause) in Japanese. The two that will be discussed here involve the phonetic scaling of pitch: clause-initial F_0 rises and F_0 lowering in clause-final position.

Japanese is well known for two competing trends in the phonetic scaling of the H and L tone entities in the phonological representation of sentence intonation. One is the phenomenon that Poser (1984), Beckman and Pierrehumbert (1986), and Pierrehumbert and Beckman (1988) (together, PPB) refer to as catathesis— the lowering and narrowing of the pitch range following a lexical H*+L pitch accent, understandable as an instance of what's generally termed downstep. The other is the upward reset or 'boost' of pitch that is found at the left edge of a phrasal constituent (see PPB, Kubozono 1993, Selkirk and Tateishi 1991, inter alia). Data from investigations of sentences that contain sequences of lexically

present theory are nicknames for 'prosodic constituent that corresponds to clause', 'prosodic constituent that corresponds to phrase', as in (3). The novel usage of the term 'intonational phrase' in Venditti (1997, 2005) for the phrase-grounded prosodic constituent that is alternatively referred to as major phrase or intermediate phrase in the earlier literature on Japanese sentence prosody is not adopted here, in the interests of terminological consistency and theoretical clarity.

[9] The corpus-based study of intonational phrasing in English by Dehé (to appear) shows there are minimal size requirements on intonational phrase, for example.

accented words has revealed a significant upward pitch reset at the left edge of clause-medial XP (= syntactic maximal projection), despite the presence of a potentially downstep-inducing lexical pitch accent in the preceding word (Selkirk and Tateishi 1991, Nagahara 1994, Sugahara 2003, Ishihara 2008, Kawahara and Shinya 2008). Further evidence for upward pitch reset at the left edge of XP comes from Selkirk, Shinya and Sugahara (2003), who show that in sentences consisting of sequences of unaccented words, there is also a significant pitch rise at the left edge of a clause-medial XP, a rise that far exceeds the rise seen at the beginning of a branching constituent that is not an XP. Selkirk and Tateishi (1991) propose that the upward pitch reset pattern observed at left edge of XP is a consequence of (i) the alignment of the left edge of XP with the left edge of a major phonologi-cal phrase produced by the syntax-phonology interface and (ii) the sensitivity of upward pitch register resetting (a phenomenon of phonetic interpretation) to the presence of that left major phrase boundary in phonological representation. The input syntactic representations of all-accented and all-unaccented sentence types like those in (4a) and (5a) would accordingly have the surface prosodic structure representations in (4b) and (5b), where the major phrases are notated with φ[10].

(4)   Sentence with lexically accented nouns
       a.   [$_{XPsubj}$[N´-no N´-ga] $_{XPobj}$[N´-o] verb]   syntactic representation
       b. $_\varphi$(N´-no N´-ga)$_\varphi$ $_\varphi$(N´-o)$_\varphi$ verb   prosodic phrasing
       c.   ↑N´⇓-no N´⇓-ga ↑N´⇓-o verb   phonetic interpretation
(5)   Sentence with lexically unaccented nouns
       a.   [$_{XPsubj}$[N-no N-ga] $_{XPobj}$[N-o] verb]   syntactic representation
       b. $_\varphi$(N-no N-ga)$_\varphi$ $_\varphi$(N-o)$_\varphi$ verb   prosodic phrasing
       c.   ↑N-no ↓N-ga ↑N-o ↓verb   phonetic interpretation

Phonetic interpretation of prosodically structured phonological representations like those in the (b) sentences would produce the intonational contour of the sentence as its output[11]. A portion of that contour is schematically represented

---

[10]The phonological phrasings in (4) and (5) are those produced by Match Phrase. The Selkirk and Tateishi (1991) left-edge alignment theory of Japanese major phrasing produces representations in which the verb is grouped in a φ with a preceding XP. The facts concern-ing left-edge upward pitch reset do not decide between these theories of the interface (see section 3.5).

[11] Following Itô and Mester (2007), it is assumed here that what's been referred to as major phrase is a maximal phonological phrase of the general category φ, namely one that is not dominated by any other φ. Principles of phonetic interpretation would single out the left edge of maximal φ for substantial reset. Ishihara (2008) reports that the upward pitch reset seen at the left edge of the maximal φ that coincides with the left edge of XP is slightly less when the preceding word is accented than when the preceding word is unaccented. This finding is consistent with the findings of Kubozono (2007), who documents a downstep-ping effect of the accented status of a preceding phrase-final word on the realization of a subsequent, upstepped, *wh*-word. These findings show that the major/maximal phrase, can-not be the domain of downstep. Instead these results show that the effects of downstep and

here in (4c) and (5c), where in lieu of pitch tracks of the intonational contours, the relevant phonetic scaling is shown with arrows. The upward bold arrow marks the upward pitch reset found at left edge of maximal φ and the down arrows ⇓ and ↓ represent, respectively, the substantial downstep found after an accent and the less pronounced downtrend seen in unaccented sequences. Equipped with this understanding of the effects of phonological phrasing on pitch scaling in Japanese sentence prosody, Kawahara and Shinya (2008) are able to show the distinctive contribution of clause structure to the intonational contours of sentences.

The materials for the Kawahara and Shinya (2008) experiment include sentences conjoining together three S(IO)OV clauses containing the same verb in all three clauses (the coordinate condition) and sentences in which the first two of the three conjoined S(IO)OV clauses show gapping of the verb, i.e. deletion of the verb under identity with the verb in the last conjoined clause (the gapped condition). Sentences of the coordinate condition were identical to those of the gapped condition, except for the presence/absence of verbs in the first two clauses, as in (6). The words of the sentences were all lexically accented.

(6)   i. Coordinate condition: $[\,_{\text{clause}}[\text{S (IO) O V}_i]\,_{\text{clause}}[\text{S (IO) O V}_i]\,_{\text{clause}}[\text{S (IO) O V}_i]\,]$
       ii. Gapped condition: $[\,_{\text{clause}}[\text{S (IO) O }t_i]\,_{\text{clause}}[\text{S (IO) O }t_i]\,_{\text{clause}}[\text{S (IO) O V}_i]\,]$

The use of sentences with coordinated clauses in which the component clauses contain sequences of XP allows for a comparison of the pitch scaling found at the edge of clause-medial XP with the pitch scaling found at the edge of clause-initial XP. An important finding of their study is that the pitch rise found at the left edge of a clause-initial XP is significantly greater than the pitch rise found at the left edge of a clause-medial XP, as shown schematically in (7c) for example, where the size of the arrow reflects the relative size of the rise.

(7)   a. $[\,_{\text{clause}}[\text{S O }t_i]\,_{\text{clause}}[\text{S O }t_i]\,_{\text{clause}}[\text{S O V}_i]\,]$
       b. $_\iota(_\varphi(\text{S})_\varphi\,_\varphi(\text{O})_\varphi\,)_\iota\,_\iota(_\varphi(\text{S})_\varphi\,_\varphi(\text{O})_\varphi\,)_\iota\,_\iota(_\varphi(\text{S})_\varphi\,_\varphi(\text{O})_\varphi\text{V})_\iota$
       c. [Nino'miya-wa] [⬆Nira'saki-no ore'nji-o],   ⬆[Yone'kura-wa] [⬆Naga'saki-no
           Ninomiya-TOP   Nirasaki-GEN orange-ACC Yonekura-TOP   Nagasaki-GEN
           ama'guri-o],   ⬆[Ima'moto-wa] [⬆Ono'michi-no ae'mono-o]   era'nda.
           chestnut-ACC, Imamoto-TOP   Onomichi-GEN mixed salad-ACC chose.
           "Ninomiya chose an orange from Nirasaki, Yonekura a chestnut from Nagasaki,
           and Imamoto a mixed salad from Onomichi."

In other words, the degree of upward pitch reset seen at the left edges of XP correlates with the position of that XP as clause-initial or clause-medial. Kawahara and Shinya take this finding to show that the phonological representation of

---

of upward φ-edge pitch reset are not mutually exclusive, but rather that they are 'additive'. The finding that the effects of downstep are not confined to the maximal φ (= major phrase) goes contrary to the assumption of Pierrehumbert and Beckman (1988) and Selkirk and Tateishi (1991). If downstep has a domain at all in Japanese, it might be the intonational phrase.

multi-clause sentences cannot consist merely of a sequence of major phonological phrases, but must include in addition a phonological representation of constituents at the intonational phrase level which would correspond to the clause structure of syntactic representation, as in (7b). As they point out, this analysis is consistent with a generalization that has emerged from previous studies in articulatory phonetics, namely that domain-initial articulatory strengthening is greater at higher prosodic constituent boundaries (see for example, Cho and Keating (2001); Fougeron and Keating (1997); Keating et al (2003); Onaka (2003)), and it is consistent as well with findings concerning upward pitch reset at left clause edges in the intonational patterns of other languages (see for example Cooper and Sorenson (1981); Ladd (1988); Truckenbrodt (2002)).

The novel use of gapped sentences in the experimental materials permits a comparison of the realization of clause-final phrases in the gapped condition with the same phrases when they lie before a verb in the coordinate condition. So for example, the pitch scaling of an underlined single word direct object XP in sentences like (8a), where it is followed by a verb, can be compared to the scaling of the corresponding direct object in gapped sentences like (8b), where it is clause-final.

(8)   a.   Coordinate condition
Mura'sugi-wa <u>nama'uni-o</u>   ↓mori'tsuke,   Muna'kata-wa
Murasugi-TOP raw sea urchin-ACC   put on dish   Munakata-TOP
<u>mame'mochi-o</u> ↓mori'tsuke,   Mori'mura-wa   <u>ae'mono-o</u>   mori'tsuketa.
bean rice-cake-ACC   put on a dish   Morimura-TOP   mixed salad-ACC   put on dish
"Murasugi put raw sea urchin on a dish, Munakata put a bean rice-cake on a dish, and Morimura put a mixed salad on a dish."

   b.   Gapped condition
Mura'sugi-wa   ↓<u>nama'uni-o,</u>   Muna'kata-wa   ↓<u>mame'mochi-o,</u>
Murasugi-TOP   raw sea urchin-ACC   Munakata-TOP   bean rice-cake-ACC
Mori'mura-wa   <u>ae'mono-o</u>   mori'tsuketa.
Morimura-TOP   mixed salad-ACC   put on dish
"Murasugi put raw sea urchin on a dish, Munakata a bean rice-cake, and Morimura a mixed salad."

The generalization that emerges from comparisons of the pitch contours of pairs like (8ab) is that the H* of the direct object in gapped sentences like (8b) is systematically lower in pitch than the H* in the identical direct object in preverbal (non-clause-final) position in sentences like (8a). (In all the materials, the pitch accent H* is located on the second mora of the word.) More generally, F_0 lowering affects the scaling of any words that are clause-final/intonational phrase-final, whether they are phrases, as with the direct objects here, or not, as with the final verbs in (8a)[12]. This final lowering effect is claimed not to be found at the left edge of φ.

---

[12]When the final word of the clause is an XP/φ, as in (8b), the final lowering effect exhibited on the pitch accented syllable subtracts from the significant upward pitch reset which is

Summarizing, Kawahara and Shinya (2008) have discovered pitch scaling phenomena in Japanese sentences with conjoined clauses which reveal an influence of clause structure/intonational phrasing. These phenomena distinguish phrases located in clause-initial and clause-final position from phrases that are clause-medial. The conclusion they draw on the basis of this and other data is (i) that clauses of Japanese sentences correspond to intonational phrases in phonological representation, as in (9b),

(9)   a.   [$_{clause}$[.....] $_{clause}$[.....] $_{clause}$[.....]   ]
     b. $_\iota$(.....)$_\iota$ $_\iota$(.....)$_\iota$ $_\iota$(.....)$_\iota$

and (ii) that the position of a phrase within this intonational phrase structure provides the distinctive context for the pitch-related aspects of phonetic interpretation they report. Note that an alternative account based on the assumption that phonetic interpretation has direct access to syntactic representation would hold that this phonetic scaling was a direct function of the syntactic clause structure of the utterance. But direct access theories like this are untenable in the face of evidence that, in general, phonological markedness constraints may contribute to determining the domain structure for phonetic interpretation (see footnote 6, and Selkirk 2009 for a review). The role for such constraints indicates that the domain structure is phonological in kind. With Kawahara and Shinya it is therefore assumed here that the clause-grounded domains for phonetic phenomena are indeed defined in terms of prosodic structure constituents in phonological representation. Their findings can be taken as validating the predictions of Match theory, according to which the repertoire of prosodic categories for any language will include a distinctive ι-level (intonational phrase) constituent that corresponds to the clause.

3.2. What counts as a clause for the syntax-prosodic structure interface?

We need now to say something more about the notion of 'clause' that has been used pretheoretically in the preceding discussion to denote the linguistic unit consisting of a subject and a predicate that corresponds to an intonational phrase. Given current assumptions about the syntactic phrase structure of sentences as consisting of a spine of functional projections above the VP, there are a variety of possible projections that could be identified as the clause, and a variety of possible locations for the subject in syntactic representation. Suppose we were to assume that 'clause' refers to the phrasal constituent that is the sister of the complementizer. This complement of C is often assumed to be the phrasal projection of the Tense[0] head, TP, and consequently the subject could occupy the specifier position of TP, or could appear in lower positions. In minimalist phase theory (Chomsky 2001), CP is a phase of the syntactic derivation and C is a phase head; the complement of C constitutes the Spell-Out domain (SOD) of the CP phase. At Spell-Out, the syntactic structure comprising the current phasal Spell-Out domain is given phonological form. It is theoretically appealing, then, to construe the clause

normally found at the left edge of XP/φ.

as the complement of C and to construe Match Clause as a provision of the theory of Spell-Out which states that the Spell-Out domain of the CP phase must have a corresponding prosodic constituent in phonological representation (namely ι, the intonational phrase).

(10)   Match Clause as a constraint calling for Spell-Out of the complement of C as ι

$$_{CP}[\text{ Spec } _{C'}[\text{ C } _{TP}[\text{ .... }]]]$$
$$\Downarrow \text{ SPELL-OUT on the CP phase}$$
$$_\iota(\text{ .... })_\iota$$

But this account of the syntactic constituent corresponding to intonational phrase in grammatical theory requires further elaboration, since it has not been generally recognized that just any sort of clause corresponds to a prosodic unit like the intonational phrase. For example, there is the well-known proposal that in English only *root* clauses correspond to intonational phrases (Downing 1970, 1973, Nespor and Vogel 1986). According to Emonds (1970, 1976), the root clause is a clause that is not dominated by any other clause; it displays a wide variety of grammatical properties that are not found in embedded clauses. Emonds proposed that parenthetical expressions, nonrestrictive relative clauses and other appositives are root clauses, in not being embedded and not being syntactically (or semantically) integrated into the surrounding matrix sentence. A nonrestrictive relative clause of English would therefore be parsed as an intonational phrase, but a restrictive relative clause, which is embedded within the noun phrase and the matrix clause of the sentence, would not be.

New theories of the root clause have emerged in the last decade or so. Rizzi (1997) et seq proposes an elaboration of the theory of the 'complementizer layer' of phrase structure which multiplies the types of 'complementizer phrase'. Most relevant to our concerns is the proposal that at the topmost complementizer layer is a functional head Force[0] representing the illocutionary force of the sentence. Given the assumption that only the topmost sentence can carry illocutionary force, the ForceP that is projected from the functional head Force[0] is not embeddable. Clauses that are embeddable would be headed by lower order complementizer elements. So the constituent that is the complement of the Force[0] complementizer could be identified with the Emonds root clause. Let's call this the Force[0] clause[13].

---

[13] Potts (2005) presents arguments showing that the Emonds characterization of parentheticals, nonrestrictive relatives and other supplement clauses as root clauses—namely as unembedded clauses that are outside the structure of the surrounding sentence—is incorrect. These clauses are indeed syntactically embedded within the larger sentential structure, he shows, and he proposes that they are identified by a [+comma] feature which instructs the semantics to treat them as if they were root clauses, and instructs the phonology to parse them with 'comma intonation'. The multi-tiered semantics developed in Potts (2005) allows for the semantic interpretation of these embedded Comma Phrases as what are in effect distinct speech acts from the speech act made by the surrounding sentence. Both the Potts term Comma Phrase and the Rizzi term Force Phrase, then, refer to a constituent that carries illocutionary forece and constitutes a sentence's speech act. In what follows, the Rizzi

In distinguishing different types of complementizer phrase and different types of clause, then, the Rizzi theory makes available the possibility that the theory of the syntax-phonology interface might distinguish between these clause types. Conceivably, Match Clause could be limited to the case of Match $Force^0$ Clause. Or, alternatively, the rendition of Match Clause given in (10) could be expanded so as to include two interface constraints, the specific $Force^0$ version and a general version that applies to the complement of any functional head of the 'complementizer layer'.

(11) Match Clause: Spelling Out the complements of complementizer heads as ι
  a. Match $Force^0$ Clause

$$_{\text{ForceP}}[\ \text{Spec} \ _{\text{Force}'}[\ Force^0 \ _{\text{CP}}[\ \dots\dots\dots\dots\ ]\ ]\ ]$$
$$\Downarrow \ \text{SPELL-OUT on the ForceP phase}$$
$$_{\iota}(\ \dots\dots\dots\dots\ )_{\iota}$$

  b. Match $Comp^0$ Clause

$$_{\text{CP}}[\ \text{Spec} \ _{\text{C}'}[\ Comp^0 \ _{\text{FncP}}[\ \dots\dots\dots\dots\ ]\ ]\ ]$$
$$\Downarrow \ \text{SPELL-OUT on any complementizer–level phase}$$
$$_{\iota}(\ \dots\dots\dots\dots\ )_{\iota}$$

  (where $Comp^0$ designates any functional head of the 'complementizer layer')

The complement-of-$Comp^0$ theory of the clause and the Match theory of the clause-intonational phrase interface together make predictions that have yet to be systematically tested crosslinguistically. The data available from Japanese begins making a contribution to the evaluation of these combined theories. One prediction is that in a language where the general version of the Match Clause interface constraint (11b) is higher ranked than any prosodic markedness constraint, the complement of every $Comp^0$, including every $Force^0$ of the sentence should correspond to an intonational phrase. In other words, a recursive intonational phrase structure would be predicted:

(12)

$$_{\text{ForceP}}[\dots[\ Force^0\ [\ \dots\dots[\dots[\ Comp^0\ [\ \dots\dots\dots\dots\ ]\ ]\ ]\ \dots\dots]\ ]$$

Successive SPELL-OUTS: $\Downarrow_{n+m}$ $\Downarrow_n$
$$\dots \ _{\iota}(\ \dots\dots\dots\dots \quad _{\iota}(\ \dots\dots\dots\dots\ )_{\iota}\ \dots\dots)_{\iota}$$

---

term Force Phrase will be used for this semantically 'topmost' constituent of a sentence, but it will be assumed, following Potts (2005), that a Force Phrase is syntactically embedded in the case of parentheticals, nonrestrictive relatives and so on, and that it triggers the multi-tier semantic interpretation Potts proposes which gives rise to multiple speech acts. In an earlier paper on intonational phrasing in English (Selkirk 2005), it was assumed that the Potts (2005) Comma Phrase correlated with the presence of intonational phrase. But evidence reported in Truckenbrodt (2005), Pak (2007, 2008), Selkirk (2009) suggesting that clauses which are not Comma Phrases may also serve as distinctive domains for phonological or phonetic phenomena makes it desirable to adopt the Rizzi theory of the complementizer layer, which allows for clauses of different types to be defined, and for all of them to be matched up with intonational phrase, as in (11b). The Potts proposal does not generalize in this way.

In the case of clauses that are syntactically coordinated, it would be predicted that the sequence of coordinated clauses should correspond to a sequence of intonational phrases. In the next section we will examine the cases of clausal coordination discussed above from this perspective. In section 3.4 we will review relevant data from the studies of the prosody of embedded questions in Japanese[14].

3.3. Fleshing out the syntactic and prosodic structure of coordinate clauses in Japanese

The Kawahara and Shinya (2008) analysis of the prosody of the coordinated structures in (6)–(8) assumes that the coordinated clause structure schematically represented in (9a) is responsible for the presence of the sequence of intonational phrases in (9b). We need to contend with two issues regarding this analysis. The first concerns the actual syntactic analysis of the conjoined clauses, and the second the identity of the prosodic constituent which groups the intonational phrases together. If the generalized complement-of-Comp^0 theory of the syntactic entity corresponding to intonational phrase in Japanese is correct, then a variety of coordination structures—whether at the ForceP level or lower—could supplant the pretheoretic syntactic representation in (9a). (13) provides one possible analysis; it represents a coordination of TP, the presumed complement to the lowest possible Comp^0 head from the 'complementizer level' of the sentence.

(13) Japanese sentence coordination with a complement-of-Comp^0 theory of the clause

$$_{\text{ForceP}}[\ _{\text{CP}}[\ _{\text{TP}}[\ _{\text{TP}}[\ S\ O\ V\ ]\ _{\text{TP}}[\ S\ O\ V\ ]\ _{\text{TP}}[\ S\ O\ V]\ ]_{\text{TP}}\ \text{Comp}^0]_{\text{CP}}\ \text{Force}^0]_{\text{ForceP}}$$

But this analysis of Japanese coordinate structures is not obviously consistent with morphosyntactic details of Japanese sentences with coordination. Of central importance is the fact that only the final verb of coordinate sentences bears the past tense marker –*ta*, as illustrated in (8a); the verbs in the nonfinal conjuncts appear in bare, tenseless, form. This restriction on the distribution of the tense morpheme in coordinate structures has led to syntactic accounts in which the coordinated clauses are analyzed as syntactic constituents that are embedded under TP and c-commanded by the single, final, Tns morpheme (Takano 2004, Hirata 2006). (14) gives the analysis of coordination offered by Hirata (2006):

(14) Japanese sentence coordination: the 'clause' as lower than complement of Comp^0 (Hirata 2006)

$$_{\text{CP}}[\ _{\text{TP}}[\ _{\text{AgrP}}[\ _{\text{AgrP}}[\ S\ O\ V]\ _{\text{AgrP}}[\ S\ O\ V]\ _{\text{AgrP}}[\ S\ O\ V]\ ]_{\text{AgrP}\ \text{Tns}}[\text{-}ta]\ ]_{\text{TP}}\ \text{Comp}^0\ ]_{\text{CP}}$$

If Takano and Hirata are right about the structure of Japanese coordinated clauses, then the syntactic character of the unit that counts as the 'clause' for the purposes of the interface constraint Match Clause would need to be redefined, since it

---

[14] Space does not permit discussion of unpublished results of recent experimental work by Kawahara (to appear) on a class of embedded parentheticals which have the syntax of noun phrases, not clauses, but share some prosodic properties of clauses.

could not be grounded in a constituent type that is a complement-of-Comp^0. But because of the appeal of the general proposal that complement-of-Comp^0 (whether at the Force^0 level or lower) corresponds to intonational phrase, it may be worth entertaining alternatives to syntactic analyses like (14) for sentence coordination. Might it be possible, for example, that in Japanese Tns is realized as part of the complementizer? If there were across-the-board raising of Tns to Comp^0, for example, then the TP that is complement-to-Comp^0 could count as the 'clause' and correspond to intonational phrase, consistent with the hypothesis that was programmatically laid out above. This question cannot be resolved here. The purpose of this discussion has been to open up relevant issues, and help define questions for future research. Clearly, consideration of the theory of the syntax-phonology interface potentially brings new evidence to bear on the evaluation of syntactic analyses, just as theories of syntax bring issues to bear in considering the nature of the syntax-phonology interface.

Let's turn now to the question whether there is evidence for a yet higher constituent of prosodic structure representation that groups together the intonational phrases that correspond to the coordinated clauses structures in (13), (14) or the pretheoretic (9a). If there is, does that evidence allow the prosodic category type of that superordinate prosodic constituent to be identified? Assuming a Rizzian theory of Comp^0 and the syntactic-interface grounded theory of prosodic constituency developed in this paper, the analysis in (15b) would be predicted, in which there is a recursive intonational phrase structure. The intonational phrases corresponding to the coordinated clauses would be grouped under another instance of intonational phrase corresponding to the topmost Force^0 clause. If, however, there were to exist a universal prosodic structure markedness constraint which ruled against such recursivity in ι-structure and this constraint were moreover higher ranked than Match Clause in the grammar of Japanese, then a simple sequence of intonational phrases, as in (15a), would conceivably be produced.

(15) Conceivable analyses of the grouping of the coordinate clause ι sequence
  a. simple sequence of ι's: $_\iota(\ \ldots)_\iota\ {}_\iota(\ \ldots)_\iota\ {}_\iota(\ \ldots)_\iota$
  b. ι-grouped sequence of ι's: $_\iota(\ {}_\iota(\ \ldots)_\iota\ {}_\iota(\ \ldots)_\iota\ {}_\iota(\ \ldots)_\iota\ )_\iota$
  c. unlabelled grouping of ι's: $(\ {}_\iota(\ \ldots)_\iota\ {}_\iota(\ \ldots)_\iota\ {}_\iota(\ \ldots)_\iota\ )$
  d. Utt-grouped sequence of ι's: $_{Utt}(\ {}_\iota(\ \ldots)_\iota\ {}_\iota(\ \ldots)_\iota\ {}_\iota(\ \ldots)_\iota\ )_{Utt}$

The two other options in prosodic structure that need to be considered would not be produced by a sentence-based theory of grammar such as this, but would conceivably be produced by a complementary discourse grammar. If that discourse grammar were to organize sentences into unlabelled units of a higher order discourse organization, then it would group the sequence of intonational phrases corresponding to coordinated clauses as in (15c). A variant of such a discourse-based prosodic constituent grouping would be (15d), where that discourse constituent is labeled Utt(erance). The Utterance has appeared as the topmost category in the repertoire of possible prosodic constituent types assumed by a number of authors (Nespor and Vogel 1986, Pierrehumbert and Beckman 1988, Gussenhoven 2004).

But the category type Utterance has no standing in the syntactic interface-grounded theory of prosodic category types put forward here, in which the highest category type is the correspondent of the clause, whether at the Force^0 level or below. If Utterance has any theoretical basis at all, it is as a discourse entity, a name to give to the speaking event constituted by the laboratory production of a sentence, for example. So (15d) is simply a notational variant of the representation in (15c).

With this in mind, let us review the claim by Kawahara and Shinya (2008) that there is evidence to support the representation in (15d), in which the intonational phrases corresponding to the coordinated clauses in their materials are grouped together within a prosodic category of type Utterance. The potentially most decisive evidence concerns the appearance of a H tone within the rightmost intonational phrase, a tonal event that is not found in any of the preceding intonational phrases. For two speakers in their experiment, this tone appears on the case-marker of the noun phrase preceding the clause-final verb, and for the other two, they propose that this same H tone docks onto the pitch accented syllable of that pre-verbal noun phrase, where it is responsible for the higher than expected F0 of the H*. (16a) illustrates the tonal representation that Kawahara and Shinya posit for the final intonational phrase of the first two speakers:

(16)

a. $\qquad$ $\overset{H^{*}+L}{\qquad}$ $\overset{H^{*}+L}{\qquad}$ $\overset{H}{\qquad}$ $\overset{H^{*}+L}{\qquad}$ $\overset{L\%}{\qquad}$
   ….(Mori'mura-wa ae'mono-o moritsu'ke-ta)$_\iota$)$_{Utt}$

b. $_{ForceP}[$ $_{FncP}[$ ….[Mori'mura-wa ae'mono-o moritsu'ke] –ta] $[\varnothing]_{Fnc0}$]$_{FncP}$
      $[\varnothing]$]$_{ForceP}$

If this extra H tone were merely a phonological entity, as Kawahara and Shinya assume, then their claim that the presence of that H tone toward the right edge of the coordinated sentence structure reveals the presence of a distinctive prosodic Utterance constituent as in (15d) would be on potentially solid footing. Kawahara and Shinya see this H tone as a peripheral tone that automatically appears as a reflex of the prosodic Utterance status of the coordinated sentence structure, on a par with the peripheral H- and L% tones that Pierrehumbert and Beckman (1988) and others have proposed automatically appear in Japanese at the left and right edges of the minor/accentual phonological phrase (see footnote 1). Because this sentence-final H tone does not appear in all the intonational phrases of the sentence, the Kawahara and Shinya argument goes, it must be a yet higher type of prosodic constituent that the phonology picks out for this H tone epenthesis, namely the Utterance. As for the surface appearance of this H tone on the penultimate minor phrase of the Utterance, they analyze this non-final positioning in a fashion parallel to the nonfinal instances of phrase accent in standard Greek analyzed by Grice et al (2000). But there is an alternative source in the grammar for this H tone. It might be a morphological entity, namely a tonal morpheme [H]$_{Fnc}$ that is present in the morphosyntactic representation and syntactically located along with other discourse particles, whether tonal or segmental, at the right edge

of the matrix sentence that groups together all the coordinated clauses; it could be the realization of one of the functional projections of the Rizzian complementizer layer, for example, as in (16b). Such a morphosyntactic analysis would equally well explain the restriction of this H tone to the rightmost intonational phrase of the coordinate sentences, and it does not require the postulation of a prosodic category of level Utterance. Given its status as a functional head high in the syntactic tree, this $[H]_{Fnc}$ morpheme would morphosyntactically linearized at the right periphery of the coordinate clause structure. And as a part of its lexical specification, this tonal functional head morpheme could require association to a specific position within the prosodic structure of the sentence, namely the right edge of penult minor phrase (in the case of the first pair of speakers) or the prominent syllable of the penult minor phrase (in the case of the second pair of speakers)[15]. So until more is known about the potential meaning-related status of this H tone, a case can't be made that its presence in the phonological representation of this set of coordinate sentences reveals the presence of an Utterance-level prosodic constituent.

The other, phonetic, evidence that Kawahara and Shinya bring to bear on the choice between the analyses in (15) does suggest that the coordinated intonational phrases of their study are together organized into a higher prosodic unit, but says nothing about whether that unit is an Utterance, another intonational phrase ι, or

---

[15]It is extremely common in Japanese that discourse-related particles that are either segmental, or tonal, or constituted of both segments and tone appear at the right periphery of the sentence (see works cited in Venditti et al. 2008, for example). The proposal that entities such as these are functional heads in the Rizzian complementizer layer of the sentence has been made for Chinese by Law (2002), among others. Of course, any such tonal morphemes that appear in situ in that clause-edge position will end up appearing at the right edge of intonational phrase in phonological representation and as such will have the status of 'boundary tones', without there needing to be any statement to this effect in the grammar.

It does seem, however, that tonal and even segmental morphemes may impose idiosyncratic, lexically specified, requirements for positioning within the prosodic structure of the sentence. Inkelas (1988) proposes that the focus particle *fa* in Hausa must be realized as the edge of a (local) phonological phrase; Aissen (1992) proposes that in various Mayan languages, topic and focus particles must appear at the edge of intonational phrase. The H phrase accent of Greek whose distribution is reported in Grice et al. (2000) is plausibly a clause-edge functional morpheme which is specified for association to the prosodic prominence of the final word of the clause/intonational phrase. Selkirk (2007) argues that the [H] focus morpheme in Bengali remains in situ at the right of a focused phrase in the syntax, and that a lexical requirement that it appear at the right edge of a phonological phrase in phonological representation forces a prosodic phrase edge to coincide with the [H] focus morpheme.

So there is definitely precedent for the proposal that the H tone that Kawahara and Shinya observe in the rightmost conjoined clause in their materials is in fact a tonal morpheme $[H]_{Fnc}$ which furthermore lexically requires a phonological realization at some remove from the normal position for the clause-final functional heads that serve as discourse particles of one kind or another.

simply unlabelled. One type of evidence concerns the relative heights of the H* of the intial pitch accents in the three intonational phrases in sequence:

(17)

$$(\ _\iota(\ H^*.....\ )_\iota \ _\iota(\ H^*.....)_\iota \ _\iota(\ H^*.....)_\iota \ )$$

There is an overall pattern of downtrend, with the first higher than the second, and the second higher than the third; the difference between the second and third, though significant, is not of the same order of magnitude as the difference between the first and the second. Kawahara and Shinya take the presence of the overall downtrend that is seen across the span of the three intonational phrases to indicate their sisterhood within a higher order prosodic constituent, and take the greater height of the first H* as a reflex of the higher initial pitch rise that would be expected if the left edge of the first intonational phrase were to coincide with the left edge of some yet higher prosodic constituent. This sort of reasoning is common in the study of intonational phrase groupings in English (Ladd 1988) and German (Truckenbrodt 2002). In addition, they find a significant effect of position at the right edge of the rightmost intonational phrase. The pitch rise up to the H* of the verb in this very rightmost position is significantly lower than even its position at the right edge of intonational phrase would imply, suggesting an increased final lowering effect due to a yet higher constituent of prosodic structure. So the degree of the phonetic scaling effects at the right and left edges of the entire sequence of intonational phrases as well as the existence of a downtrend across the span of that sequence all converge on the conclusion that there is a higher order prosodic unit that groups together the component intonational phrases. But since these effects are matters of quantitative degree, rather than involving qualitatively different phonetic or phonological properties which are not found with intonational phrases themselves, it is not possible to establish which among the representations (15b)–(15d) actually forms part of the phonological representation. These phenomena do not testify to the presence of a prosody constituent of type Utterance.

Summarizing, the Kawahara and Shinya (2008) study of the prosody of coordinate clauses in Japanese provides evidence that each coordinated clause corresponds to a unit of prosodic structure, the intonational phrase. The intonational phrase serves as a unit with respect to which various pitch scaling phenomena of phonetic interpretation are defined, in addition to other phonetic phenomena (including presence of creakiness in segments at the right edge and of a following pause) which have not been discussed here. Evidence also supports the grouping of these coordinate intonational phrases into a larger prosodic unit, though this unit is one whose identity can't be established on the basis of the data at hand. This said, there is no basis for assuming there is a category Utterance which would 'label' this constituent grouping. As argued above, the evidence is consistent with the predictions of the Match theory of the syntax-prosodic constituency interface in (3) and with the universal prosodic hierarchy that it entails. It is not consistent

with a theory that gives no place to clause structure and its corresponding intonational phrase structure in the theory of Japanese sentence prosody.

3.4. Embedded clauses as intonational phrases: evidence from Japanese indirect questions

The general version in (11b) of the syntactic clause-intonational phrase interface specified by Match Clause predicts that *any* sort of clause—whether the topmost clause of the sentence or one that is deeply embedded within the sentence—would require the presence of a corresponding intonational phrase in phonological representation. Recursive embedding of intonational phrases should be produced in satisfaction of this general Match Clause constraint, in some languages at least:

(18) $_\iota($ ... $_\iota($ ..... $)_\iota$ ... $)_\iota$

An embedded intonational phrase structure like this is just what Ladd (1986) proposes for sentences containing medial parenthetical phrases in English. Yet English seemingly does not display such embedding of intonational phrases with 'normal' embedded clauses, namely ones that are *not* marked by the Pottsian [+comma] feature (which we are understanding here to be a property of the Rizzian Force^0 head). Since cross-linguistic investigation of clause-grounded intonational phrasing is in its infancy, it is not possible to do more than speculate about the (alleged) absence of embedded intonational phrasing in English that would correspond to these embedded complement types in English. The current hypothesis is that this would be due to phonological constraints on ι-structure. What is needed at this point, in order to show some support for a general complement-to-Comp^0 provision of Match Clause (11b) independent of the Match Force^0 Clause provision in (11a) is crosslinguistic evidence of the existence of embedded intonational phrase domains that correspond to embedded clauses of the complement-to-lexical head variety, for example. Selkirk (2009) suggests that the evidence from Luganda and Huave that Pak (2008) takes to show that the Spell-Out domain of the CP phase serves as a domain for phonological phenomena can be seen as supporting an intonational phrasing that corresponds to an embedded instance of complement-to-Comp^0. Below it is suggested that data which is already available from studies of the prosody of embedded and matrix *wh-* questions in Japanese provide support for the proposal that embedded clauses in Japanese correspond to intonational phrases.

3.4.1. The prosody of *wh-* questions in the Fukuoka dialect

Investigation of the tonal patterns of wh-questions in the Fukuoka dialect of Japanese by by Hayata (1985) and Kubo (1989) et seq indicate a very close correlation between tonal pattern and syntactic configuration which reveals, it can be argued, the presence of intonational phrases. The data reviewed here are from Kubo (2005)[16]. In *wh*-questions in Fukuoka, a H tone plateau extends right-

---

[16]Jennifer Smith informed me some years ago of the work of Hayata and Kubo on the *wh-*

ward from a *wh*-word up to the edge of the clause which corresponds to the *wh*-question. Following Smith (2005) this stretch between the *wh*-word and the right edge of its clausal scope will be called the *wh*-domain. The problem is how to account for the H tone plateau of the *wh*-domain. Part of the puzzle concerns the fact that, except if they are located in the *wh*-domain, all words in the Fukuoka dialect bear a H*+L pitch accent. (In nouns the position of the accent in a word is lexically determined, while in verbs and adjectives accent always appears on the penult syllable, by default.) A declarative sentence like (19) illustrates the normal appearance of H*+L on every word, noted with ⌉, while the *wh*-questions in (20ab) show the absence of pitch accents in the *wh*-domain and the presence of the H tone plateau[17]:

(19) kyo⌉o   bi⌉iru   non⌉da
      today   beer   drank
      'I drank beer today.'

(20) a. kyo⌉o da⌐re-ga   biiru   nonda
          today WHO-nom   beer   drank ∅COMP
          'Who drank beer today?'

      b. da⌐re-ga   kyoo   biiru   nonda
          WHO-nom   today beer   drank ∅COMP
          'Who drank beer today?

When the *wh*-word forms part of an embedded *wh*-question, however, the H tone plateau extends only to the end of the embedded complement-to-Comp[0].

(21) [[[da⌐re-ga kyoo biiru nonda]⌐ka] sitto⌉o]

          WHO-nom today beer drank COMP know ∅COMP
          'Do you know who drank beer today?'

A L tone appears on the complementizer *ka* of the embedded clause; the matrix verb *sittoo* 'know', which introduces the embedded *wh*-question, retains its word-level H*+L pitch accent.

In Kubo's account of these facts *wh*-questions display a special prosodic constituency, one in which the span from the *wh*-word to the complementizer of the *wh*-question constitutes a minor phonological phrase. (22) gives the proposed prosodic constituency of (20a), and (23) that for (21):

(22) MiP(kyo⌉o)MiP MiP(dare-ga biiru non⌉da)MiP   [cf. (20a)]
                    WHO-nom   ∅COMP

---

patterns of the Fukuoka dialect, which except for Kubo (2005), is in Japanese. Smith (2005) presents the results of the Kubo (1989, 1990a, 1990b, 1992, 2001, 2005) investigations in the context of an attempt to account for these results in the context of the Multiple Spell-Out theory of the syntax-phonology interface laid out by Ishihara (2004).

[17]Note the slight rise at the end of the H tone plateau. It is conceivably derived from the upstepping of a question-final H% boundary tone after the H of the preceding plateau.

(23) $_{\text{MiP}}$(dare-ga kyoo biiru nonda ⌐ka)$_{\text{MiP}}$ $_{\text{MiP}}$(sitto⌐o)$_{\text{MiP}}$ [cf. (21)]
　　　WHO-nom　　　　　　　COMP　　 ∅COMP

(Minor phrases in Fukuoka are otherwise assumed by Kubo to correspond to a single word, each bearing a pitch accent.) A default pitch accent would be assigned to the penult syllable of the minor phrase of the *wh*-domain, explaining the fall to the L tone on the embedded *wh*-complementizer *ka* in (21). (The absence of this L on the final syllable of the *wh*-domain minor phrase in the matrix version in (22) could explained by the appearance of the H% boundary tone.) In addition to the pitch accent, the minor phrase in Fukuoka shows an initial H- peripheral tone that appears on the second mora of the MiP when neither the first or second mora is pitch accented, as in the Tokyo dialect. In the minor phrase of the *wh*-domain, this H- tone would appear on the second mora of the *wh*-word, from which it would spread rightward to the pitch accent, just as in word-sized minor phrases. This analysis thus accounts for the shape of the tonal pattern observed. But the loss of the lexical pitch accents on nouns within the *wh*-domain remains to be explained: why wouldn't one of those lexical pitch accents survive as the pitch accent for the *wh*-minor phrase? And a deeper problem is the stipulative character of *wh*- minor phrase formation. What general principle(s) could explain this departure from syntactic constituent-grounded prosodic constituency, and from the normal correspondence of minor phrases to much smaller morphosyntactic domains?

　　I believe that an alternative account based on intonational phrasing can shed new, better, light on this phenomenon. Suppose we make the assumption that in the Fukuoka dialect any clause whether embedded or not, i.e. any complement-to-Comp^0, corresponds to an intonational phrase, in accordance with the general syntax-phonology interface constraint Match Clause, (11b). This would predict the intonational phrase structures in (24) and (25) for (20a) and (21):

(24) $_{\iota}$(kyoo dare-ga biiru nonda)$_{\iota}$ [cf. (20a)]
　　　　WHO-nom　　　∅COMP

(25) $_{\iota}$($_{\iota}$(dare-ga kyoo biiru nonda)$_{\iota}$ka sittoo)$_{\iota}$ [cf. (21)]
　　　　WHO-nom　　　　　COMP　　∅COMP

The right edge of the intonational phrase containing the *wh*-word provides a characterization of the right edge limit of the high tone plateau that extends rightward from the *wh*-word. This intonational phrasing provides a crucial piece of a very different, but typologically plausible, solution to the puzzle of the rightward spreading of the H tone from the *wh*-word, the loss of intervening lexical pitch accents in conjunction with the H tone spread, and the presence of the H tone on the *wh*-word in the first place. The idea is that the loss of the intervening pitch accents is a consequence of (i) the presence of a high tone *wh*-morpheme $[\text{H}]_{+wh}$ that is affixed to *wh*-words like *dare* and (ii) a requirement that this $[\text{H}]_{+wh}$ morphemic tone also be associated (autosegmentally) to the right edge of the intonational phrase that immediately dominates it in the phonological representation. Given the linear nature of tonal representations, specifically the necessary absence of crossing tonal association lines, satisfaction of a requirement for the multiple linking/spreading

of the $[H]_{+wh}$ morpheme would entail the deletion of any intervening lexical tones. (26) would be the phonological representation corresponding to (21).

(26) $[H]_{+wh}$ H*+L   H*+L

     ₁(₁(dare-ga kyoo biiru nonda)₁ ka sittoo)₁
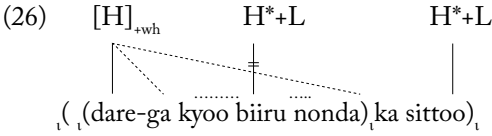
(Dashed association lines indicate the tonal associations called for by H tone spread. '=' marks the elimination of a previously established tonal association, in this case a lexical pitch accent.)

   As for the proposal that a $[H]_{+wh}$ morpheme marks *wh*-words in the first place, one could see this morpheme as providing a morphological mark of the distinction between the *wh*- and non-*wh*- forms of indeterminate pronouns in Japanese (Shimoyama 2001, Kratzer and Shimoyama 2002). It's possible that the $[H]_{+wh}$ is present in the Tokyo dialect as well, where it provides an explanation for the rather high pitch that is found on *wh*-words in Japanese (see e.g. Kubozono 2007). (This sort of tonal prominence for *wh*-words is not typically found cross-linguistically.) And what of the proposal that the grammar of the Fukuoka dialect requires that this $[H]_{+wh}$ be linked as well with the right edge of IP, giving rise to the observed H tone plateau? Study of the phonology of tonal morphemes crosslinguistically reveals that certain tonal morphemes may be morphosyntactically located in one position in the sentence but also apparently require simultaneous association to some distant position that is definable in terms of prosodic structure (cf. footnote 14). In the Bantu language Tsonga, for example, the presence of a H-toned negative morpheme that is morphosyntactically positioned in the post-subject verbal complex correlates with the appearance of a H level plateau that extends from that negative morpheme to the right edge of intonational phrase; at the same time there is a loss of the lexical tones of the words that lie between Neg and the right edge of intonational phrase (Kisseberth 1995). In Tsonga it is clear that these two related phenomena must be explained in terms of the tonal representation, in a manner analogous to what is proposed for Fukuoka *wh*-questions in (26), since the prosodic phrasing of a sentence in Tsonga does not correlate with the presence or absence lexical tone, but only with the manner in which it spreads (see Kisseberth 1995, Selkirk 2009). Given the commonality in the formal properties of the Negative H spreading in Tsonga and *wh*- H spreading in Fukuoka, I would like suggest that they both involve instances of tonal morphemes that are located in a designated position within the morphosyntactic structure of the sentence and also lexically specified for association to an intonational phrase edge within the prosodic structure of the sentence. What happens in these specific cases, is that both the tonal association specified by the morphosyntactic linearization and the further tonal association to the edge of the intonational phrase specified by the morpheme itself are realized at Spell-Out, and as a consequence a multiply linked tonal structure is created that results in the deletion of any intervening lexical or

morphemic tones.

A central assumption of phase theory is that Spell-Out applies on successively higher phases. This assumption permits an explanation for data from multiple *wh*-questions in Fukuoka. Kubo (2005) reports that in cases where an embedded *wh*-question lies between a matrix *wh*-word and the end of the matrix clause, a H tone plateau extends from the matrix *wh*-word to the end of the matrix clause, as in (27).

(27) [[da︱re-ga [[dare-ga   biiru nonda] ka]   sittoo]
      WHO-nom WHO-nom beer   drank COMP know   ∅COMP
      'Who knows who drank beer ?'

Match Clause predicts that (27) has the center-embedded intonational phrase structure in (28).

(28) $[H]_{+wh}$ $[H]_{+wh}$
             | |
      $_\iota$(dare-ga $_\iota$(dare-ga biiru nonda)$_\iota$ka sittoo)$_\iota$

This structure does not, however, explain why the H tone spread from the matrix *wh*-word takes precedence, in effect obliterating any sign of the predicted H tone spread from the embedded *wh*-word to the intonational phrase corresponding to the of the embedded clause. But assuming Multiple Spell-Out does. There would first be H tone spread from $[H]_{+wh}$ on the lower phase corresponding to the embedded question. Then, on the higher phase corresponding to the matrix question, Spell-Out would call for H tone spread from the $[H]_{+wh}$ attached to the *wh*-word in the matrix Spell-Out domain. That H tone spread takes precedence over any other tonal representation that would conflict with it, resulting in the obliteration of the *wh*-H tone plateau established on the lower phase, just as it would result in the obliteration of the tonal material of lexical pitch accents:

(29) Spell-Out on the phase of the matrix *wh*-question: spreading of the matrix
            $[H]_{+wh}$ $[H]_{+wh}$ H*+L

      $_\iota$(dare-ga $_\iota$(dare-ga biiru nonda)$_\iota$ka sittoo)$_\iota$
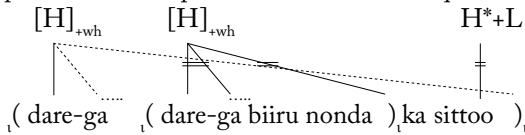
(Dashed association lines indicate the H tone spread called for on the matrix phase.) In the proposed derivation, the constraint requiring the establishment of the *wh*-plateau on the later, matrix, phase outranks any faithfulness constraint requiring realization of the tonal input on that phase, which is the output of Spell-Out on the earlier phase. It is not obvious how a non-derivational, non-Multiple Spell-Out account would account for the precedence of the matrix *wh*-plateau realization.

This phase-based account of the tonal patterns in embedded and matrix *wh*-questions in the Fukuoka dialect has been inspired by the proposal of Ishihara (2004) et seq to understand the pitch patterns of embedded and matrix *wh*-

questions in the Tokyo dialect in terms of Multiple Spell-Out, as well as by the demonstration by Smith (2005) that a phase-based Multiple Spell-Out account of the Fukuoka data which retains the essential features of the Kubo analysis is untenable. The analysis given here, which combines a new proposal for the morphophonology of Fukuoka *wh-* elements and the Match Clause theory that the Spell-Out domain of a $Comp^0$-headed corresponds to intonational phrase, when couched within the phasal theory of Multiple Spell-Out, has allowed for an insightful account of these patterns in the Fukuoka dialect. It remains to be seen, of course, whether the predictions of Multiple Spell-Out, rather than a non-derivational theory, are more generally borne out when it comes to accounting for patterns of sentence phonology.

3.4.2. The prosody of *wh-*questions in the Tokyo dialect

Studies of the intonation of matrix and embedded *wh-*questions in Tokyo Japanese provide a second source of data from Japanese that bears on the hypothesis being investigated here that embedded clauses correspond to intonational phrases (unless such a configuration is ruled out by phonological constraint). Deguchi and Kitagawa (2002) and Ishihara (2002) put forward the factual claim that the intonational pattern of a *wh-*question correlates with the embedded vs. matrix scope of the question, a claim that has since been the subject of experimental investigation, and debate (see in particular Hirotani (2003, 2005); Ishihara (2003, 2004, 2007); Kitagawa (2005); Kitagawa and Fodor (2003, 2006)). The question for the present paper is whether these findings constitute evidence for, or against, the hypothesis that embedded clauses in Japanese correspond to intonational phrases. The specific claim put forward by the authors cited above is that in a *wh-* question, the *wh-*word shows both (i) greater upward pitch excursion and (ii) significant pitch compression following it which extends up to the end of the clause which constitutes the scope of the *wh-*word. Thus the matrix question in (30a) displays a different pattern from the sentence with an embedded question in (30b):

(30) a. Matrix scope question
       [[Naoya-ga   ↑ $nani_i$-o↓ <u>nomiya-de nonda</u>] no_i]?
         Naoya-NOM what-ACC   bar-LOC   drink   Q
       "What did Naoya drink at the bar?
    b. Embedded scope question
       [Naoya-wa [[Mari-ga   ↑ $nani_i$-o↓ <u>nomiya-de nonda</u>] ka_i]
         Naoya-TOP Mari-ACC what-ACC bar-LOC   drink   Q
       ↑Yumi-ni   morasita]
         Yumi-DAT   divulge
       "Naoya divulged to Yumi what Mari drank at the bar."

The thick down arrow indicates the initial locus of downward pitch compression after the wh-word; the thin up arrow indicates upward pitch reset after the stretch of pitch compression, which is indicated by underlining. In (30a) the post *wh-* pitch compression extends to the end of the matrix sentence, which coincides

with the matrix *wh*- scope. In (30b), the pitch compression extends only to the end of the embedded clause, which is an embedded, indirect, *wh*- question. At the left edge of the matrix constituent *Yumi-ni* that follows the right edge of the embedded clause, there is upward pitch reset. (It should be mentioned that lexical pitch accents are retained under post-*wh* pitch compression, though they are quite reduced (Hirotani 2005, Ishihara 2004), just as they are retained by reduced under post-Focus pitch compression (Sugahara 2003/2005a).) The facts in (30) would be consistent with an alternative generalization, namely that the pitch compression extends up to the edge of the intonational phrase that corresponds to the clause that contains the *wh*- word, if phonological representations like those given in (24) and (25) for Fukuoka were assumed. But a further type of fact is seemingly not consistent with this intonational phrase analysis of the domain of post-*wh* pitch compression. These authors also observe that when a *wh*-word which is located in the embedded clause has matrix clause scope, the post-*wh* pitch compression extends from the *wh*-word through the embedded clause up to the edge of the matrix clause as indicated by the underlining in (31):

(31)   Matrix scope question with embedded *wh*-word
　　　Naoya-wa [[Mari-ga ↑nani$_i$-o↓   nomiya-de nonda] to]
　　　Naoya-TOP Mari-NOM what-ACC bar-LOC   drink COMP
　　　Yumi-ni   morasita no$_i$
　　　Yumi-DAT   divulge   Q
　　　"What did Naoya divulge to Yumi that Mari drank at the bar?"

Unlike (30b), no upward pitch reset is reported at the edge of a phrase like *Yumi-ni* belonging to the matrix clause in a sentence of the type in (31). An intonational phrase-based account of the domain of post-*wh* pitch compression would predict that even in cases like (31), pitch compression would not extend beyond the embedded clause.

　　There is debate, however, about the status of the intonational contrast reported between (31) and (30b). Hirotani (2005) reports on a series of experiments that were designed to test the claim that there is such a contrast in the production of these question types. The results show that there is no systematic intonational difference between the two, except in circumstances where the speakers have explicitly been made aware of the meaning contrasts. Hirotani's experiments employed sentences which were identical in the content of the embedded clause and complementizer but unambiguous in the scope of the *wh*-question; the matrix complementizer indicates the difference in scope:

(32)　a.　Matrix scope question with embedded *wh*-word
　　　　　[[Yama'zaki-wa   [[Nino'miya-ga   da're$_i$-o   XP* maneita] ka]   XP*
　　　　　　Yama'zaki-TOP   Nino'miya-NOM   who-ACC   invited   Q
　　　　　mora'sita]-ndai$_i$]?
　　　　　revealed   Q
　　　　　"Who did Yamazaki reveal whether Ninomiya invited?"
　　　b.　Embedded scope question with embedded *wh*-word

    [[Yama'zaki-wa   [[Nino'miya-ga   da're$_i$-o   XP*   maneita]   ka$_i$]   XP*
      Yama'zaki-TOP   Nino'miya-NOM   who-ACC   invited   Q
    mora'sita]-nokai]?
    revealed   Q
    "Did Yamazaki reveal who Ninomiya invited?

The experiments varied the length of the questions; this is indicated by XP*. Hirotani found that, in circumstances where speakers were not attempting to make a contrast between the two scopes, the appearance of upward pitch reset at the left edge of a matrix XP following the embedded clause did not distinguish between the two cases. Both matrix scope and embedded scope questions showed that upward pitch reset most of the time, as in (30b). The facts are more consistent with the phonological, intonational phrase-based account being entertained here. First of all, the absence of contrast between these matrix and embedded scope cases is what a theory like this that gives phonetic interpretation no access to syntactic representation or scope relations would predict for the prosody of questions with embedded *wh*-words. And if embedded clauses are indeed parsed as intonational phrases in Tokyo Japanese, as in (18), regardless of their status as *wh*-scope, then it is predicted that upward pitch reset would be systematically possible after the embedded clause, as the Hirotani data indicates. The data does apparently show some variation: the absence of upward reset is also possible, in both embedded and matrix scope cases. Conceivably this might be a result of optionality in the parsing of embedded clauses as intonational phrases. Finally, there is the fact that when speakers were aware of the meaning contrast between the two sentence types, they did reliably produce the intonational contrast reported in earlier work ((30b) vs. (31)). Conceivably these distinct productions for the two scope cases under circumstances of speakers' metalinguistic awareness are a reflection of the prosody-meaning correlation found in the comprehension experiments conducted by Hirotani (2005). The Hirotani comprehension experiments employed embedded *wh*-word questions that were ambiguous as to embedded or matrix scope, as in (33).

(33)   [[John-wa   [[Mary-ga   nani-o   katta]   ka]   zimusitu-de
        John-TOP   Mary-NOM   what-ACC   bought Q   office-LOC
    kikimasita] ka]
    asked   Q
    "Did John ask in the office what Mary bought?" (embedded scope)
OR:
    "What did John ask in the office whether Mary bought?" (matrix scope)

The stimuli varied in the presence/absence of upward pitch reset following the embedded clause. The results showed that in response to stimuli containing that upward pitch reset, speakers showed a strong bias for the embedded scope interpretation, but that when no pitch reset was present, both the embedded and matrix interpretations were available. Hirotani suggests that these comprehension results reflect a psycholinguistic processing principle—Scope-Prosody Correspondence—

which privileges scope or binding relations between elements of the sentence that are contained within the same prosodic constituent[18]. The production results from her experiments suggests that Scope-Prosody Correspondence does not influence pronunciation, expect perhaps when speakers are attempting to convey the meaning difference.

Taking stock of the empirical situation, clearly there are differences in findings that need to be sorted out and understood concerning the prosody of *wh*-questions with embedded *wh*-words in Japanese. This paper is not the place to attempt a detailed review of the full set of findings from experimental research on the prosody of *wh*-questions that has or is currently being conducted (see e.g. footnote 19). The purpose of this discussion has been to open up the possibility that investigations of the prosody of *wh*-questions in Tokyo Japanese might provide evidence for the hypothesis being explored here that embedded clauses correspond to intonational phrases in Japanese. The ongoing debate around the various types of explanation for the prosodic patterning of *wh*-questions is likely to continue to produce highly relevant data. And it will hopefully inform research on the more general question of the correspondence between intonational phrasing and embedded clauses of all types in Japanese, and other languages.

An important property of the phonological representations that Match Clause, (11b), derives for sentences with embedded questions in Japanese, and with any other type of embedded clause, is the recursive center-embedded character of the corresponding intonational phrasing, schematically represented in (18). The embedded intonational phrase is flanked at the left and right, not by other intonational phrases, but by the phonological phrases or words which correspond to the surrounding material in the matrix sentence. Do these embedded intonational phrases display the same range of phonetic behavior that is exhibited by intonational phrases that correspond to coordinate clauses? Does one find clause-initial upward pitch reset at the left edge of the embedded intonational phrase that exceeds in degree what would be found at the left edge of a merely clause-medial phrase? And does one also find the clause-final lowering that is attested in Kawahara and Shinya's study? These are the sorts of questions that await further research. Answering these will help decide whether there is further supporting evidence for the embedded intonational phrasing that the Match Clause hypothesis predicts should in principle exist.

3.5. Against alternative accounts of the syntax constituency-phonology interface

Theories of the syntactic constituency-phonology interface in grammar fall into two broad classes, those that presuppose the existence of a prosodic constituent structure that mediates between syntactic structure and domain-sensitive phono-

---

[18]Hirotani (2005) proposes that Scope-Prosody Correspondence disprefers cases where scope-taking and scoped elements like *wh*-complementizers and *wh*-words are in separate major (maximal) phonological phrases. Given the discussion above, an alternative characterization of the relevant prosodic constituent might be intonational phrase.
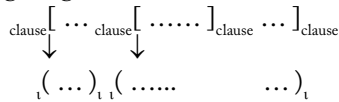
logical and phonetic phenomena, and those which do not assume any prosodic structure above the foot and instead see domain-sensitive phonology and phonetics as being defined directly in terms of syntactic constituency or the syntactic derivation. The Match theory of the syntax-prosodic structure interface proposed in the present paper, which calls for syntactic constituents of the type clause, phrase, and word to each correspond to a distinct prosodic constituent type, belongs to the first class. So does the earlier edge-based syntactic-prosodic structure alignment theory that is proposed in Selkirk (1986, 1996), for example. The phase-based theory of domains put forth in Ishihara (2004) as well as in Seidl (2001), Pak (2008), and Tokizaki (2008) belongs to the second, 'direct-access-to-syntax', class, as does the syntactic constituency-based theory of Wagner (2005), to appear. The contention of this paper has been that the theory of the syntactic constituency-phonology interface must be a theory of the first class, and specifically that it must be a Match theory of the interface. In this section, a brief review will be made of the reasons which suggest that neither the edge-based theory in that first class, nor any of the theories of the second class are able to provide the insight into syntax-influenced sentence prosody that the Match theory of the interface can provide.

Comparing the prosodic structure-based Match theory with 'direct access' theories of the syntactic constituency-phonology/phonetics interface, the most striking difference is the implicit claim made by the latter that the domains for sentence phonology and phonetics are defined only in terms of syntactic constituency or in terms of the syntactic derivation (as corresponding to the Spell-Out domain of the phase, for example). Match theory, on the other hand, assumes that the grammar defines a prosodic structure representation independent of the syntactic representation. It is consistent, therefore, with evidence that might show the existence of intonational phrases that do not correspond to clauses, and with evidence that particular clauses do not correspond to intonational phrases. The hypothesis is that such divergences between clause structure and intonational phrase strucure would have their source in phonological markedness constraints on prosodic structure (see, e.g. Selkirk 2009). In the material reviewed in this paper, we have not seen any evidence of this sort of divergence from the syntax at the intonational phrase level. But there is already compelling evidence at the phonological phrase level in Japanese that purely phonological factors may influence the phonological domain structure of the sentence. These included the accented vs. unaccented status of words, as well as binarity constraints on the size of phrases (Selkirk and Tateishi 1988; Kubozono 1993; Sugahara 2002; Dobashi 2003; Selkirk, Shinya, and Kawahara 2004; Shinya, Selkirk and Kawahara 2004). There is also ample evidence from other languages that prosodic markedness constraints contribute to determining the parsing of the sentence into phonological phrases (see Selkirk 2009 for a review). There is evidence of size constraints on intonational phrasing in other languages as well (e.g. Dehé to appear on English). For these reasons, it is necessary to posit an independent prosodic representation of phonological domain structure as part of the theory of grammar, and it is therefore necessary to construe the theory of the syntax-phonology interface as consisting, in part, of a theory of
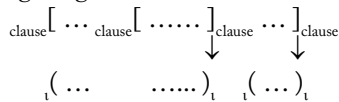
the relation between syntactic constituency and prosodic constituency.

Turning to a comparison within the class of prosodic-structure based theories, the end-based theory proposed in Chen (1987) and Selkirk (1986), exploited in the Selkirk and Tateishi (1991) account of phonological phrasing in Japanese, and re-articulated in Selkirk (1996) in terms of generalized alignment theory (McCarthy and Prince 1993), differs from the current Match theory in a number of important ways. Because it calls for only the right or left edge of a designated syntactic constituent (e.g. the clause) to correspond to a prosodic constituent (e.g. the intonational phrase), the Align theory systematically allows the creation in prosodic structure of constituents that do not correspond to syntactic constituents. Moreover it provides no pressure for the creation of nested, recursive, prosodic structures. Together these could be called the 'prosodic grouping consequences' of the alignment theory. But the evidence from the phonology of wh-questions in Fukuoka Japanese sentences, for example, shows that it is the recursive ι-domain structure in (33c) that is predicted by Match Clause, and not the left- or right-aligned intonational phrasing in (33ab), which provides the correct domain structure.
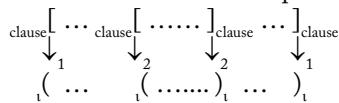
(33) a. Left-edge alignment of clause and intonational phrase

$$_{\text{clause}}[\; \cdots \; _{\text{clause}}[\; \cdots\cdots \;]_{\text{clause}} \; \cdots \;]_{\text{clause}}$$
$$\downarrow \qquad\qquad \downarrow$$
$$_{\iota}(\; \cdots \;)_{\iota} \; _{\iota}(\; \cdots\cdots \qquad\qquad \cdots \;)_{\iota}$$

   b. Right-edge alignment of clause and intonational phrase

$$_{\text{clause}}[\; \cdots \; _{\text{clause}}[\; \cdots\cdots \;]_{\text{clause}} \; \cdots \;]_{\text{clause}}$$
$$\downarrow \qquad\qquad\qquad \downarrow$$
$$_{\iota}(\; \cdots \qquad\qquad \cdots\cdots \;)_{\iota} \; _{\iota}(\; \cdots \;)_{\iota}$$

   c. Match of clause and intonational phrase

$$_{\text{clause}}[\; \cdots \; _{\text{clause}}[\; \cdots\cdots \;]_{\text{clause}} \; \cdots \;]_{\text{clause}}$$
$$\downarrow_1 \qquad \downarrow_2 \qquad \downarrow_2 \qquad \downarrow_1$$
$$_{\iota}(\; \cdots \qquad _{\iota}(\; \cdots\cdots \;)_{\iota} \; \cdots \quad)_{\iota}$$

More evidence from other languages of similar problematic predictions made by the alignment theory is given in Selkirk (2009). Thus the Alignment theory makes the wrong predictions about prosodic grouping. Its earlier apparent success, as in the Selkirk and Tateishi (1991) account of major phrasing in Japanese, was due to in large part to the accidental fact that the domain-sensitive phonological or phonetic phenomena that were examined in individual languages diagnosed only one edge of the relevant prosodic constituent. In the case of Japanese the phenomenon at issue was left-edge-sensitive upward pitch reset. When a language displays phenomena that diagnose both edges of a domain, as in Tsonga (Kisseberth 1995, Selkirk 2009), or which show a phenomenon applying across smaller and larger instances of recursively nested domains, it's Match theory that makes the right predictions.

4. Conclusion and prospectus

This paper has reviewed evidence suggesting that the clause structure of a sentence in Japanese corresponds to a domain for certain of the phonological and phonetic phenomena that contribute to defining the intonational patterns of Japanese sentences. This domain has been referred to as the ι-domain, or intonational phrase. The finding that there is a clause-grounded intonational phrase domain of phonological representation in Japanese—one that is distinct from the (major/intermediate) phonological phrase—is predicted by a universal theory of the prosodic hierarchy, and more specifically, by the hypothesis that the universal repertoire of prosodic category types is grounded in a universal theory of the syntax-prosodic structure interface like Match theory. Match theory, which may be construed as a component of the theory of Spell-Out in minimalist phase theory (Chomsky 2001), specifies that a set of syntactic constituent types match up with corresponding prosodic constituents in phonological representation. The specific proposal made here is that the syntactic constituent type referred to by the interface constraint Match Clause is the complement of $Comp^0$, which denotes the general class of functional heads of the 'complementizer layer' (Rizzi 1997). Assuming that all maximal phrasal projections of the general type $Comp^0$ are phases, in the sense of Chomky (2001), then the 'clause' as understood here can be identified with the Spell-Out domain of any CompP phase.

Evidence from Japanese also suggests that both embedded clauses and matrix clauses correspond to intonational phrases. The nested, recursive, intonational phrase structure that is motivated by the prosody of matrix and embedded *wh*-questions in Japanese is at variance with the strict layer hypothesis. In other words, the attested recursivity does not have its source in the phonology. Rather, recursive intonational phrase structure is produced when the Match Clause constraint is satisfied on nested clausal domains (and when that recursive structure is not ruled out by phonological constraints which might call for that recursive structure to be avoided). This is what is expected given the hypothesis being pursued here that prosodic constituent structure organization above the foot derives from the organization of syntactic structure, through the agency of the interface Match constraints.

The data from Japanese on the relation between clause and intonational phrase structure that has been considered so far, then, is precisely what is predicted by a syntactically grounded theory of prosodic constituent structure—both the repertoire of prosodic categories and the manner in which the prosodic constituent structure may fail to show strict layering. It opens up the prospect that crosslinguistic study, as well as further research on Japanese, will provide further support for the Match theory of the syntax-prosodic constituency interface from which these fundamental properties are derived. This universal theory of the interface does not, it should be noted, predict lack of variation in prosodic constituent organization, either within a language or between languages. It has been hypothesized that the Match constraints may interact with prosodic markedness constraints in a language-particular ranking. This predicts the possibility of crosslinguistic varia-

tion in constraint ranking and hence the possibility of variation in prosodic structure organization. Moreover, another source of variation is in a language's syntactic structure itself, e.g. the position of the subject above or below the complementizer. Relevant syntactic structure variation will be mirrored in the phonological representation via the interface Match constraints. Further research is required, needless to say, to determine if these hypotheses concerning the range of variation in possible phonological domain structures in language are borne out in fact.

References

Adger, David (2007) Stress and phasal syntax. *Linguistic Analysis* 33: 238–266.

Aissen, Judith (1992) Topic and focus in Mayan. *Language* 68: 518–536.

Beckman, Mary and Pierrehumbert, Janet (1986) Intonational structure in English and Japanese. *Phonology Yearbook* 3: 255–310.

Chen, Matthew (1987) The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4: 109–150.

Cho, T. and Keating, Pat (2001) Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics* 29: 155–190.

Chomsky, Noam (2001) Derivation by phase. In: Michael Kenstowicz (ed.) *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Cooper, William E. and Sorenson, J. (1981) *Fundamental frequency in sentence production*. New York: Springer.

Deguchi, Masanori and Kitagawa, Yoshihisa (2002) Prosody and wh-questions. In: Masako Hirotani (ed.) *Proceedings of the 32nd annual meeting of the Northeastern Linguistics Society*, 73–92. Amherst MA: GLSA, UMass Amherst.

Dehé, Nicole (To appear) Clausal parentheticals, intonational phrasing and prosodic theory. *Journal of Linguistics* 45.

Dobashi, Yoshihito (2003) Phonological phrasing and syntactic derivation. Unpublished doctoral dissertation, Cornell University.

Downing, Bruce (1970) Syntactic structure and phonological phrasing in English. Unpublished doctoral dissertation, University of Texas at Austin.

Downing, Bruce (1973) Parenthetical rules and obligatory phrasing. *Papers in Linguistics* 6: 108–128.

Emonds, Joseph (1970) Root and structure-preserving transformations, Linguistics, MIT.

Emonds, Joseph (1976) *A transformational approach to English syntax*. New York: Academic Press.

Fougeron, Cécile and Keating, Pat (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 106: 3728–3740.

Frota, Sonia (2000) *Prosody and focus in European Portuguese: Phonological phrasing and intonation*. Outstanding dissertations in linguistics. New York, London: Garland Publishing Co.

Féry, Caroline and Truckenbrodt, Hubert (2005) Sisterhood and tonal scaling. *Studia Linguistica. Special issue "Boundaries in intonational phonology"* 59: 223–243.

Ghini, Mirco (1993) φ-formation in Italian: A new proposal. *Toronto Working Papers in Linguistics* 12: 41–78.

Grice, Martine, Ladd, D. Robert, and Arvaniti, Amalia (2000) On the place of phrase accents in intonational phonology. *Phonology* 17: 143–185.

Gussenhoven, Carlos (2004) *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Haraguchi, Shosuke (1977) *The tone pattern of Japanese: An autosegmental theory of tonology.* Tokyo: Kaitakusha.

Hayata, Teruhiro (1985) *Hakata hougen no akusento / keitairon [The accent and morphology of the Hakata dialect].* Fukuoka: Kyushu University Press.

Hayes, Bruce and Lahiri, Aditi (1991) Bengali intonational phonology. *Natural Language and Linguistic Theory* 9: 47–96.

Hayes, Bruce (1995) *Metrical stress theory: Principles and case studies.* Chicago: Chicago University Press.

Hellmuth, Sam (2007) The relation between prosodic structure and pitch accent distribution: Evidence from Egyptian Arabic. *The Linguistic Review* 24: 291–316.

Hirata, Ichiro (2006) Coordination, subject raising and AgrP in Japanese. *Linguistic Inquiry* 37: 318–329.

Hirotani, Masako (2003) Prosodic effects on the interpretation of Japanese wh-questions. In: Luis Alonso-Ovalle (ed.) *University of Massachusetts Occasional Papers in Linguistics 27: On semantic processing*, 117–137. Amherst MA: Graduate Linguistics Student Association.

Hirotani, Masako (2005) *Prosody and LF interpretation.* Unpublished doctoral dissertation, University of Massachusetts Amherst.

Hyman, Larry (2006) Word-prosodic typology. *Phonology* 23: 225–258.

Inkelas, Sharon (1988) Prosodic constraints on syntax: Hausa 'fa'. In: Hagit Borer (ed.) *Proceedings of the Seventh West Coast Conference on Formal Linguistics*, 375–389. Palo Alto CA: Stanford Linguistics Association.

Inkelas, Sharon (1990) *Prosodic constituency in the lexicon.* Outstanding dissertations in Linguistics. New York: Garland Publishing Co.

Ishihara, Shin (2002) Invisible but audible wh-scope marking. In: Line Mikkelson and Christopher Potts (eds.) *Proceedings of the 21st West Coast Conference on Formal Linguistics*, 180–193. Somerville, MA: Cascadilla Press.

Ishihara, Shin (2003) Intonation and interface condition. Unpublished doctoral dissertation, MIT.

Ishihara, Shin (2004) Prosody by phase: Evidence from focus intonation-WH scope correspondence in Japanese. In: Shin Ishihara, M. Schmitz and A. Schwarz (eds.) *Interdisciplinary Studies on Information Structure 1: Working Papers of SFB632*, 77–119. Potsdam: University of Potsdam.

Ishihara, Shin (2007) Major Phrase, Focus Intonation, Multiple Spell-Out. *The Linguistic Review* 24: 137–167.

Ishihara, Shin (2008) Japanese downstep revisited. Manuscript, University of Potsdam. Submitted for publication.

Itô, Junko and Mester, Armin (2003) Weak layering and word binarity. In: Takeru Honma, Masao Okasaki, Toshiyuki Tabata, and Shin-ichi Tanaka (eds.), *A New Century of Phonology and Phonological theory: A Festschrift for Professor Shosuke Haraguchi on the Occasion of his 60ᵗʰ Birthday*, 26–65. Tokyo: Kaitakushi. A slightly revised version of an unpublished 1992 UC Santa Cruz paper of the same title.

Itô, Junko and Mester, Armin (2006) Prosodic Adjunction in Japanese Compounds. Paper presented at *Formal Approaches to Japanese Linguistics 4*, Osaka.

Itô, Junko and Mester, Armin (2007) Categories and projections in prosodic structure. handout of talk delivered at *OCP4 [Old World Conference in Phonology]*. Rhodes, Greece.

Itô, Junko and Mester, Armin (2009a) The onset of the prosodic word. In: Steve Parker (ed.) *Phonological Argumentation: Essays on Evidence and Motivation* (forthcoming). London:

Equinox.

Itô, Junko and Mester, Armin (2009b) Trimming the prosodic hierarchy. In: Toni Borowsky, Shigeto Kawahara, Takahito Shinya and Mariko Sugahara (eds.) *Prosody matters: Essays in honor of Elisabeth Selkirk* (forthcoming). London: Equinox Publishers.

Jun, Sun-Ah (2005) Prosodic typology. In: Sun-Ah Jun (ed.) *Prosodic typology: The phonology of intonation and phrasing*, 430–458. Oxford, New York: Oxford University Press.

Katz, Jonah and Selkirk, Elisabeth (2009) Contrastive focus vs. discourse-new: Evidence from prosodic prominence in English. Manuscript, MIT and University of Massachusetts Amherst. Submitted for publication.

Kawahara, Shigeto and Shinya, Takahito (2008) The intonation of gapping and coordination in Japanese; Evidence for Intonational Phrase and Utterance. *Phonetica* 65: 62–105.

Kawahara, Shigeto (To appear) The intonation of parentheticals in Japanese. In: Toni Borowsky, Shigeto Kawahara, Takahito Shinya and Mariko Sugahara (eds.) *Prosody matters: Essays in honor of Elisabeth Selkirk*, London: Equinox.

Keating, Pat, Cho, T., Fougeron, Cécile, and Hsu, C.-S. (2003) Domain-initial strengthening in four languages. *Papers in laboratory phonology 6: Phonetic interpretation*, 145–163. Cambridge: Cambridge University Press.

Kisseberth, Charles (1995) On domains. In: Jennifer Cole and Charles Kisseberth (eds.) *Perspectives in Phonology*, 133–166. Stanford, CA: CSLI Publications.

Kitagawa, Yoshihisa and Fodor, Janet (2003) Default prosody explains neglected syntactic analysis in Japanese. In: W. McClure (ed.) *Japanese/Korean Linguistics* 12, 267–279. Stanford CA: CSLI Publications.

Kitagawa, Yoshihisa (2005) Prosody, syntax and pragmatics of Wh-questions in Japanese. *English Linguistics* 22: 302–346.

Kitagawa, Yoshihisa and Fodor, Janet (2006) Prosodic influence on syntactic judgments. In: Gisebert Fanselow, Caroline Fery, Ralf Vogel and Matthias Schlesewsky (eds.) *Gradience in grammar: Generative Perspectives*, Ch. 17. Oxford: Oxford University Press.

Kratzer, Angelika, and Shimoyama, Junko (2002) Indeterminate pronouns: The view from Japanese. In: Y. Otsu (ed.) *The proceedings of the third Tokyo conference on psycholinguistics*, 1–25. Tokyo: Kaitakusho.

Kratzer, Angelika and Selkirk, Elisabeth (2007) Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24: 93–135.

Kubo, Tomoyuki (1989) Hukuoka-si hougen no, dare/nani tou no gimonsi o hukume bun no pittipaan [The pitch patterns of sentences containing wh-words in the Fukuoka-city dialect]. *Kokugogaku* 156: 71–82.

Kubo, Tomoyuki (1990a) Hukuoka-si hougen no toi-gaesi gimonsi gimonbun (wh-echo) no pittipataan [The pitch patterns of wh-echo questions in Fukuoka-city dialects]. *Bungaku kenkyuu* 87: 153–179.

Kubo, Tomoyuki (1990b) Hukuoka-si hougen no gimonsi hyougen no akusento kisoku [Accent rules for question expressions in Fukuoka-city dialect]. *Kyuudai gengogaku kenkyuusitu houkoku* 11: 103–118.

Kubo, Tomoyuki (1992) Hukuoka-si hougen ni okeru akusento syouka ni tuite [On accent deletion in Fukuoka-city dialects]. In: Tensuya Kunihiro (ed.) *Nihongo intoneesyon no zittai to bunseki*. Tokyo: Monbusho, "Nihongo no onsei" research project.

Kubo, Tomoyuki (2001) Hukuoka hougen ni okeru tougoron to on'inron no kyoukai ryouiki [Syntax-phonology interface in the Fukuoka dialect]. *Journal of the Phonetic Society of Japan* 5: 27–32.

Kubo, Tomoyuki (2005) Phonology-syntax interfaces in Busan Korean and Fukuoka

Japanese. In: Shigeki Kaji (ed.) *Cross-linguistic studies of tonal phenomena: Historical development, tone-syntax interface, and descriptive studies*, 195–210. Tokyo: Research Institute for languages and cultures of Asia and Africa, Tokyo University of Foreign Studies.

Kubozono, Haruo (1988) The organization of Japanese prosody. Doctoral dissertation, University of Edinburgh. [Kurosio Publishers, 1993]

Kubozono, Haruo (1989) Syntactic and rhythmic effects on downstep in Japanese. *Phonology* 6: 39–67.

Kubozono, Haruo (1993) *The organization of Japanese prosody*. Tokyo: Kurosio.

Kubozono, Haruo (2007) Focus and intonation in Japanese. In: Shin Ishihara (ed.) *Interdisciplinary Studies on Information Structure 9: Proceedings of the 2nd Workshop on Prosody, Syntax and Information Structure*, 1–27. Potsdam: University of Potsdam.

Ladd, D. Robert (1986) Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook* 3: 311–340.

Ladd, D. Robert (1988) Declination 'Reset' and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84: 538–544.

Law, Ann (2002) Cantonese sentence-final particles and the CP domain. *UCL Working Papers in Linguistics* 14: 375–398.

McCarthy, John and Prince, Alan (1993) Prosodic morphology: Constraint interaction and satisfaction, 48–1201: Rutgers Optimality Archive.

McCarthy, John and Prince, Alan (1995) Faithfulness and reduplicative identity. In: Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*: *University of Massachusetts Working Papers in Linguistics* (UMOP) 18, 249–384.

Nagahara, Hiroyuki (1994) Phonological phrasing in Japanese. Unpublished doctoral dissertation, UCLA.

Nespor, Marina and Vogel, Irene (1986) *Prosodic phonology*. Dordrecht: Foris Publications.

Onaka, A. (2003) Domain-initial strengthening in Japanese: an acoustic and articulatory study. *Proceedings of the 15th International Congress of Phonetic Sciences*: 2091–2094. Barcelona.

Pak, Marjorie (2007) Phrasal tone domains in San Mateo Huave. *Proceedings of the 33rd annual meeting of the Berkeley Linguistics Society* 33.

Pak, Marjorie (2008) The postsyntactic derivation and its phonological reflexes. Unpublished doctoral dissertation, University of Pennsylvania.

Peperkamp, Sharon (1997) *Prosodic Words*. HIL dissertation series 34. The Hague: Holland Academic Graphics.

Pierrehumbert, Janet and Beckman, Mary (1988) *Japanese Tone Structure*. Linguistic Inquiry Monograph Series. Cambridge MA: The MIT Press.

Poser, William (1984) The phonetics and phonology of tone and intonation in Japanese. Unpublished doctoral dissertation, MIT.

Potts, Christopher (2005) *The logic of conventional implicatures*. Oxford, New York: Oxford University Press.

Revithiadou, Anthi and Spyropoulos, Vassili (2005) The Multiple Spell-Out Hypothesis and the phonological component: Evidence from Greek. Paper presented at *NELS 35*, University of Connecticut.

Revithiadou, Anthi and Spyropoulos, Vassili (In press) A dynamic approach to the syntax-phonology interface: A case study from Greek. In: Kleanthes Grohmann (ed.) *InterPhases: Phase-theoretic investigations of linguistic interfaces*, Oxford: Oxford University Press.

Rizzi, Luigi (1997) The fine structure of the left periphery. In: Liliane Haegeman (ed.) *Elements of grammar: Handbook in generative syntax*, 75–116. Dordrecht: Kluwer Publishing.

Seidl, Amanda (2001) *Minimal indirect reference: A theory of the syntax-phonology interface*. Outstanding dissertations in Linguistics. London, New York: Routledge.

Selkirk, Elisabeth (1978) On prosodic structure and its relation to syntactic structure. Ms. University of Massachusetts Amherst.

Selkirk, Elisabeth (1980) The role of prosodic categories in English word stress. *Linguistic Inquiry* 11: 563–605.

Selkirk, Elisabeth (1981a) On prosodic structure and its relation to syntactic structure. In: Thorstein Fretheim (ed.) *Nordic Prosody II: Papers from a Symposium*, 111–140. Trondheim: TAPIR.

Selkirk, Elisabeth (1981b) The nature of phonological representation. In: John Anderson, John Laver and Terry Myers (eds.) *The cognitive representation of speech*, 379–388. Dordrecht: North Holland Publishing Company.

Selkirk, Elisabeth (1986) On derived domains in sentence phonology. *Phonology Yearbook* 3: 371–405.

Selkirk, Elisabeth (1996) The prosodic structure of function words. In: James L. Morgan and Katherine Demuth (eds.) *Signal to syntax: Prosodic bootstrapping from speech to grammar in early acquisition*, 187–214. Mahwah, NJ: Lawrence Erlbaum.

Selkirk, Elisabeth (2000) The interaction of constraints on prosodic phrasing. In: Merle Horne (ed.) *Prosody: Theory and experiments*, 231–262. Dordrecht: Kluwer.

Selkirk, Elisabeth (2005) Comments on intonational phrasing. In: Sonia Frota, Marina Vigario and Maria Joao Freitas (eds.) *Prosodies*, 11–58. Berlin: Mouton de Gruyter.

Selkirk, Elisabeth (2006) Strong minimalist Spell-Out of prosodic phrases. Paper delivered at *GLOW*, Barcelona.

Selkirk, Elisabeth (2007) Focus prominence drives focus phrasing: Bengali intonation revisited. In: Chung-min Lee, Matthew Gordon and Daniel Büring (eds.) *Topic and focus: A cross-linguistic perspective*, 217–246. Dordrecht: Kluwer.

Selkirk, Elisabeth (2009) The syntax-phonology interface. In: John Goldsmith, Jason Riggle and Alan Yu (eds.) *The Handbook of Phonological Theory* (forthcoming), Oxford: Blackwell.

Selkirk, Elisabeth, Shinya, Takahito, and Kawahara, Shigeto (2004) Phonological and phonetic effects of phrase length on F0 in Japanese. In: Bernard Bel and Isabelle Marlien (eds.) *Proceedings of the second international conference on speech prosody 2004*, 183–186. Nara, Japan.

Selkirk, Elisabeth, Shinya, Takahito, and Sugahara, Mariko (2003) Degree of initial lowering in Japanese as a reflex of prosodic structure organization. *Proceedings of the 15th International Congress of Phonetic Sciences*, 491–494. Barcelona.

Selkirk, Elisabeth and Tateishi, Koichi (1988) Constraints on minor phrase formation in Japanese. In: M.G. Larson and Diane Brentari (eds.) *Proceedings of the 24th annual meeting of the Chicago Lingusitics Society*, 316–336. Chicago IL: Chicago Linguistics Society.

Selkirk, Elisabeth and Tateishi, Koichi (1991) Syntax and downstep in Japanese. In: Carol Georgopoulos and Roberta Ishihara (eds.) *Interdisciplinary approaches to langauge: Essays in honor of S.-Y. Kuroda*, 519–543. Dordrecht: Kluwer.

Shimoyama, Junko (2001) *Wh-constructions in Japanese*. Unpublished doctoral dissertation, University of Massachusetts Amherst.

Shinya, Takahito, Selkirk, Elisabeth, and Kawahara, Shigeto (2004) Rhythmic boost and recursive minor phrase in Japanese. In: Bernard Bel and Isabelle Marlien (eds.) *Proceedings of the second international conference on speech prosody*, 183–186. Nara, Japan.

Smith, Jennifer (2005) On the WH-question intonational domain in Fukuoka Japanese. In: Shigeto Kawahara (ed.) *Papers in Prosody, University of Massachusetts Occasional Papers in Linguistics*, 219–237: Amherst MA: GLSA.

Sugahara, Mariko (2002) Conditions on post-Focus dephrasing in Tokyo Japanese. In: Bernard Bel and Isabelle Marlien (eds.) *First International Conference on Speech Prosody*, 655–659. Aix-en-Provence, France.

Sugahara, Mariko (2003) Downtrends and Post-FOCUS intonation in Japanese. Unpublished doctoral dissertation, University of Massachusetts Amherst.

Sugahara, Mariko (2005a) *Downtrends and post-Focus intonation in Japanese*. Amherst MA: GLSA, University of Massachusetts Amherst.

Sugahara, Mariko (2005b) Post-FOCUS prosodic boundaries in Tokyo Japanese: Asymmetric behavior of an F0 cue and domain-final lengthening. *Studia Linguistica* 59: 144–173.

Takano, Yuji (2004) Coordination of verbs and two types of verbal inflection. *Linguistic Inquiry* 35: 168–178.

Tokizaki, Hisao (2008) *Syntactic structure and silence: A minimalist theory of syntax–phonology interface*. Tokyo: Hituzi Syobo Publishing.

Truckenbrodt, Hubert (1999) On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30: 219–256.

Truckenbrodt, Hubert (2002) Upstep and embedded register levels. *Phonology* 19: 77–120.

Truckenbrodt, Hubert (2005) A short report on intonation phrase boundaries in German. *Linguistische Berichte* 203: 273–296.

Venditti, Jennifer (1997) Japanese ToBI labelling guidelines. *Ohio State Working Papers in Linguistics* 50: 127–162.

Venditti, Jennifer (2005) The J_ToBI model of Japanese intonation. In: Sun-Ah Jun (ed.) *Prosodic typology: The phonology of intonation and phrasing*, 172–200. Oxford, New York: Oxford University Press.

Venditti, Jennifer, Maekawa, Kikuo, and Beckman, Mary (2008) Prominence Marking in the Japanese intonation system. In: Shigeru Miyagawa and Mamoru Saito (eds.) *The Oxford handbook of Japanese linguistics*, 456–512. Oxford, New York: Oxford University Press.

Wagner, Michael (2005) Prosody and recursion. Unpublished doctoral dissertation, MIT.

Wagner, Michael (To appear) Prosody and recursion in coordinate structures and beyond. *Natural Language and Linguistic Theory*.

Werle, Adam (2009) Word, phrase and clitic prosody in Bosnian, Serbian and Croatian. Unpublished doctoral dissertation, University of Massachusetts Amherst.

Yip, Moira (2002) *Tone*. Cambridge: Cambridge University Press.

Author's contact information:
Department of Linguistics
South College
150 Hicks Way
University of Massachusetts Amherst
Amherst MA 01003, USA

【要　旨】

日本語の節とイントネーション句について
―韻律構造の統語論的基盤―

Elisabeth Selkirk

（エリザベス・セルカーク）

マサチューセッツ大学

　日本語の文プロソディーの研究では最近 Kawahara and Shinya（2008）が日本語の並列節のプロソディーを，Deguchi and Kitagawa（2002），Ishihara（2002），Hirotani（2003）が wh 疑問文のプロソディーをそれぞれ分析し，統語的な節（clause）が一部の音韻・音声現象の適用範囲と一致することを示唆するデータを示した。本稿はこれらの研究成果を概説し，日本語が示す統語構造とプロソディー構造の対応関係が統語論と音韻論のインターフェースに関する普遍的な理論から導かれることを示す。本稿ではまた，統語構造と音韻構造の一般的な対応関係を構築する新しいインターフェース理論（Match 理論）の概要を紹介し，この理論が Chomsky（2001）が提唱する Spelll-Out 理論の一部門となりうることを示す。さらに，この理論が再帰的（recursive）なイントネーション構造が一定の条件下で起こることを予測し，Itô and Mester（2007）が提唱する新しい韻律構造分析とも相容れることを論じる。