

## 9 Recursive prosodic phrasing in Japanese\*

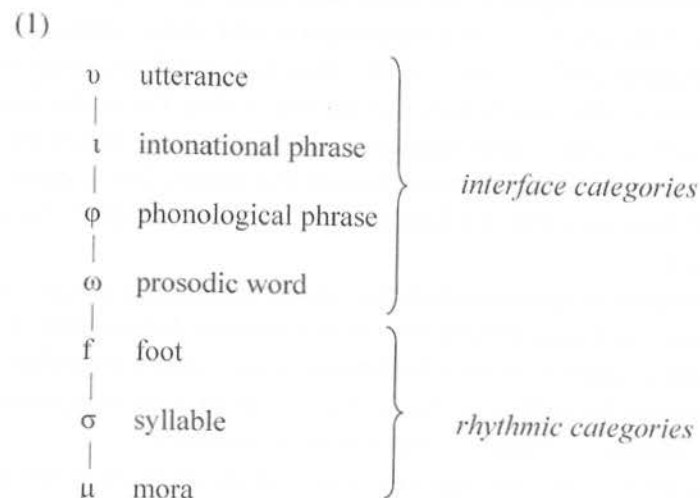
Junko Ito and Armin Mester<sup>a</sup>

### 9.1 Background and identification of the problem

A key result of studies in phrasal phonology since the 1970s has been the finding that, cross-linguistically, phrase-level phonological processes do not make use of the vast set of potential domains that are in principle made available by grammatical (i.e., syntactic and morphological) structure. Rather, they are localized in a small set of domains that are phonological in nature, even though defined in reference to grammatical structure, and that turn out to play a decisive role in language after language. The model that developed in response to this central finding is *prosodic hierarchy theory* (Selkirk, 1978; Nespor and Vogel, 1983; Inkelas, 1989: 4, etc.), building on key insights in earlier work (such as Halliday, 1960 and Pike, 1967): Speech is organized into a set of prosodic domains that form a hierarchy of containment, with each non-terminal constituent made up of a sequence of constituents at the next level down (the *Strict Layer Hypothesis*, see Selkirk, 1984 and Nespor and Vogel, 1986, among others).

The hierarchy comprises two groups of categories, as shown in (1). The word-internal units (syllable, foot, and perhaps mora) are intrinsically defined in terms of sonority-related phonetic factors and speech rhythm, whereas the parsing of higher-level units (prosodic word, phonological phrase, intonational phrase, etc.) is regulated by constraints, alignment-based and other, on the correspondence between syntactic/morphological and phonological

constituents. We will sometimes refer to the word-internal prosodic units as *rhythmic* categories, and the larger prosodic units as *interface* categories.



While the general form of the rhythmic categories is relatively uncontroversial, apart from specific issues (such as the status of the mora as a genuine prosodic constituent vs. a property of syllables, see Ito and Mester, 2003 [1992]), the picture is far less clear for the interface categories. The overall research program has been vastly successful in advancing our understanding of the relation between syntactic structure and phonological form, but many questions of both detail and principle have remained open, even as to the exact number and/or content of the units of the hierarchy. To make matters worse, categories are sometimes literally defined in terms of the processes associated with their instantiations in specific languages, resulting in labels like ‘accentual phrase’, ‘tone group’, etc. While this is mnemonically useful for the description of a single language, the lack of truly cross-linguistically valid and constant properties associated with these units creates additional obstacles in identifying categories between languages and grammars (see Selkirk, 2005 and Truckenbrodt, 2006 for discussion). The underlying research program has valued the postulation of new descriptive categories, designed to serve as domains for various processes in various languages, over restrictiveness.

One of the main points of dispute is, perhaps unsurprisingly, the mid-range of the prosodic hierarchy, where at least two distinct phrasal categories have

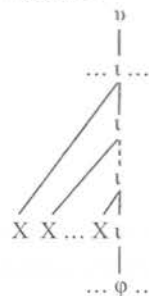
<sup>a</sup> Junko Ito and Armin Mester: University of California, Santa Cruz, CA, USA. Email: ito@ucsc.edu, mester@ucsc.edu

been proposed, the minor phrase and the major phrase (alternatively named 'accentual phrase' and 'intermediate phrase'). The distinction seems to have grown out of research on Japanese, one of the best-studied prosodic systems. The two kinds of phrases were first distinguished in McCawley (1968), followed by Haraguchi (1977), Poser (1984), Beckman and Pierrehumbert (1986), Kubozono (1988), and Selkirk and Tateishi (1988). The distinction was then adopted for many other languages, including Basque (Elordieta, 2007), English (Selkirk, 1996), German (Kratzer and Selkirk, 2007), Italian (Petroni and D'Imperio, 2008), and Korean (Jun, 1998). See Jun (2005) for a general overview.

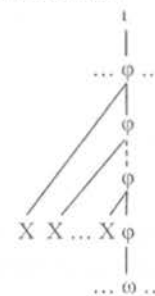
This paper reopens the question of whether two distinct phrasal categories are truly necessary. Is a model of prosodic parsing possible that accounts for all the facts, both in Japanese and in other languages, with a single phonological phrase category? We will argue that such a conception is not only possible, but in fact necessary: multiple categories create problems.

The starting point of the new approach is the simple observation that a single phrasal category does not mean a single layer of structure at a given level of prosody. Equating the two presupposes subscribing to the doctrine, long abandoned as part of orthodox strict layering, that only a new category can introduce a new level of structure. Even though prosodic structure shows nothing like the depth of embedding created by recursion in syntax,<sup>1</sup> it is a far cry from this uncontroversial observation to the strict layering conclusion that, given a prosodic hierarchy with  $n$  categories, each path (from root to terminal node) in a prosodic tree must have a length of exactly  $n$  layers. On the contrary, both level skipping and level repetition have been well motivated, resulting in paths with fewer and more levels, respectively. Level skipping occurs in weakly layered structures, as argued in Ito and Mester (2003 [1992]), and level repetition is found in recursive prosody, as was assumed in early prosodic theory (see, for example, Nespor and Vogel, 1983) and forcefully argued by Ladd in a number of publications (see Ladd 1996 for a summary). Following this growing body of work, we allow for additional layers to arise through recursion, in particular, through adjunction, as shown in (2).

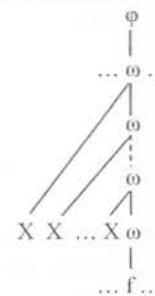
## (2) Prosodic recursion for interface categories:

a.  $t$ -recursion

e.g., Ladd (1986, 1988) on nested coordinate structures in English

b.  $\varphi$ -recursion

e.g., Gussenhoven (1991, 2005) on the rhythm rule in Dutch and English

c.  $\omega$ -recursion

e.g., Ito and Mester (2007) on Japanese compounds and Ito and Mester (2009a, 2009b) on English and German function word complexes

Constituent does not equal category once recursive structures are admitted. The crucial point is that recursion imposes further levels of structure on the string being parsed without claiming that each time a different category is involved. We are assuming an optimality-theoretic analysis (Selkirk, 1996; Truckenbrodt, 1999), where a NonRecursivity constraint prevents unlimited recursion and vacuous recursion. In fact, one layer for each category is optimal given appropriate parsing constraints, and recursive structures arise only if they serve to fulfill a higher ranking constraint (e.g., a syntax-prosody alignment constraint, a binarity restriction, etc.).

## 9.2 Interface categories in Japanese

Most previous work on the phrase-level prosody of Japanese (e.g., McCawley, 1968; Haraguchi, 1977; Poser, 1984; Beckman and Pierrehumbert, 1986; Kubozono, 1988; Selkirk and Tateishi, 1988) distinguishes the two phrasal interface categories in (3).

## (3)

- |     |  |
|-----|--|
| MaP | major ( $\approx$ 'intermediate') phrase |
|     |  |
| MiP | minor ( $\approx$ 'accentual') phrase    |

The rationale for this supposedly irreducible distinction is that MaP and MiP are domains for different processes. The three main generalizations are summarized in (4).

- (4)
- |         |                                |
|---------|--------------------------------|
| a. MiP: | Domain of accent culminativity |
| b. MiP: | Domain of initial rise         |
| c. MaP: | Domain of downstep             |

How can a model with a single and undifferentiated category  $\phi$  ('phonological phrase') make the necessary distinctions? The key lies in a better understanding of what the facts really imply about 'domains', and of the ways in which domains relate to categories.

The first generalization (4a) is more of a definition than an argument for a domain, and simply states that there can be at most one accent (H\*L) in a MiP (hence the alternative name 'accentual phrase'). The generalization in (4b) is schematically illustrated in (5) (after Selkirk and Tateishi, 1991), where several unaccented prosodic words are joined into a single MiP. The central observation is that MiP, the domain of accent culminativity, is also the locus of the initial rise (a %L boundary tone followed by a phrasal H- tone).

(5) Initial rise within MiP

- 
- a. [<sub>MaP</sub>(<sub>MiP</sub> Oomiya-no) (<sub>MiP</sub> Inayama-no yuujin-ga inai)]  
friend isn't there  
'Mr. Inayama's friend from Oomiya isn't there.'
- b. [<sub>MaP</sub>(<sub>MiP</sub> Oomiya-no Inayama-ga)(<sub>MiP</sub> yuujin-o yonda)]  
friend called  
'Mr. Inayama from Oomiya called his friend.'

Downstep (or catathesis) refers to the lowering of the pitch register following an accented syllable (Poser, 1984). This is illustrated in (6) (after Selkirk and Tateishi, 1991), where each prosodic word is accented and therefore

projects a MiP of its own. The domain of downstep is MaP in the sense that post-accent lowering takes place throughout MaP, and the pitch register is reset only at the beginning of the next MaP, not at the beginning of each MiP.

(6) Downstep within MaP

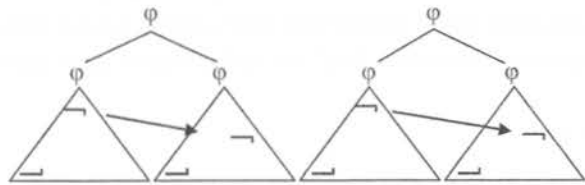
- 
- a. [<sub>MaP</sub>(<sub>MiP</sub> Aoyama-no)] [<sub>MaP</sub>(<sub>MiP</sub> Yamaguchi-no)(<sub>MiP</sub> anyome-ga inai)]  
sister-in-law isn't there  
'Mr. Yamaguchi's sister-in-law from Aoyama isn't there.'
- b. [<sub>MaP</sub>(<sub>MiP</sub> Aoyama-no) (<sub>MiP</sub> Yamaguchi-ga)] [<sub>MaP</sub>(<sub>MiP</sub> anyome-o yonda)]  
sister-in-law called  
'Mr. Yamaguchi from Aoyama called his sister-in-law.'

A schematized diagram (with two MaPs, each with two MiPs) illustrating the initial rise (at the beginning of each MiP) and downstep (indicated by arrows within MaP) is given in (7).

- (7)
- 
- ┘ = %L,    ㄣ = H\*L (lexical pitch accent)

How solid are these domain arguments? Are they sufficient grounds to motivate distinct categories? What goes wrong if both MaP and MiP are simply recursive undifferentiated  $\phi$ s, as in (8)?

(8)



The surprising result is that nothing goes wrong: As far as the initial rise and downstep are concerned, there is no reason to distinguish between different kinds of phrases, MaP and MiP (or intermediate phrase and accentual phrase).

First, by accent culminativity a MiP contains maximally one accent; since downstep requires two accents, the first one downstepping the second, it cannot have any effect within a single MiP. Since it already follows from the structural description of downstep that it cannot apply within MiP, there is no need to specifically exclude it from this domain. Only one kind of phonological phrase  $\phi$  is needed, then, as in (8), where the lowest  $\phi$ s (not containing other  $\phi$ s) are automatically excluded as ‘domains of downstep’.

In structures like (8), minimal  $\phi$ s alone serve as domains of accent culminativity. Ito and Mester (2007: 100) argue that accent is a *head feature*: Each accent must be the head of some prosodic phrase  $\phi$ . This ensures that every minimal  $\phi$  has at most one accent (any additional accent would be a non-head), but non-minimal  $\phi$ s are free to contain more than one accent. The accent-as-head idea explains why it is the minimal  $\phi$  that is the domain of accent culminativity.

In a similar way, there is no need to ban an initial rise from the left edge of MaP. As long as every MaP begins with a MiP, it is already predicted in any case: [<sub>MaP</sub> [<sub>MiP</sub> %LH ...]...]. Moreover, the observation that the degree of initial rise is more extreme at MaP edges than at MiP edges (Selkirk *et al.*, 2003) casts doubt on the idea that all instances of initial rise can be reduced to MiP-rises. For this reason alone, it seems more adequate to locate a %LH sequence at the left edge of every phrase  $\phi$  (we will return to this issue below in Section 9.4.3).

Downstep and initial rise, then, work without problems in the one- $\phi$  model. Our immediate conclusion is that the initial rise applies to all  $\phi$ -phrases (not just to MiP), and likewise downstep applies to all  $\phi$ -phrases (not just to MaP). The more interesting conclusion is that we can contemplate abolishing the entire MaP/MiP distinction: Let there be only one phonological phrase  $\phi$ !

### 9.3 Recursion-based subcategories

What the facts and generalizations seem to demand is not enough *categories*, as the standard view of prosodic form has it, but rather enough *levels of structure*. But structure can be provided in a number of ways, the ‘level = category’ approach being just one of a number of possibilities, and arguably not the optimal one. In this context, it is significant that work over the last twenty years has firmly established recursivity as an indispensable attribute of prosodic form in a number of languages. The existence of recursive phrasing has been demonstrated at the level of the intonational phrase by Ladd (1986, 1988) and Ladd and Campbell (1991), and at the level of the phonological phrase by Gussenhoven (2005); prosodic word recursion has been shown to hold for compounds and function word complexes by Booij (1996) and Ito and Mester (2007, 2009a, 2009b). Further development, with additional evidence, is found in Kubozono (1988, 1993, 2005), van den Berg *et al.* (1992), Truckenbrodt (2002), and Féry and Truckenbrodt (2005), among others. Extensive study and motivation of recursive structures in prosody is provided in two recent dissertations (Wagner, 2005 (MIT); Schreuder, 2006 (Groningen); see also Wagner, 2010).

Building on this line of work, we will here outline a model that we refer to as *recursion-based subcategories*. Each prosodic category defines its own network of projections, where the usual tree-structural notions apply, such as minimal and maximal projection and head vs. non-head. Phonological and phonetic processes are part of the realization of this structure, and signal important boundaries by selecting different subconstituents as their domains. Using standard tree-structural terminology, the largest projection of a prosodic category  $\kappa$  is the ‘maximal  $\kappa$ ’, and its smallest projection is the ‘minimal  $\kappa$ ’, as defined in (9).

(9)

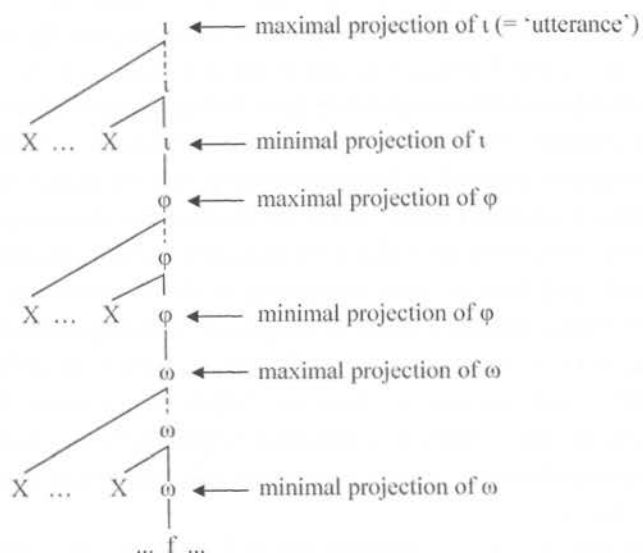
$$\begin{aligned} \kappa_{\max} &= \kappa \text{ not dominated by } \kappa \\ \kappa_{\min} &= \kappa \text{ not dominating } \kappa \end{aligned}$$

The schematic structure in (10) shows how these definitions apply to the interface categories *intonational phrase* ( $\iota$ ), *phonological phrase* ( $\phi$ ), and *prosodic word* ( $\omega$ ). Taking up a suggestion by Shigeto Kawahara, we propose that *utterance* ( $\upsilon$ ), usually posited as the highest category in the prosodic hierarchy, is not a separate category, but rather the maximal projection of the



*intonational phrase* ( $\iota$ ). The empirical prediction is that *utterance* cannot be recursive; its only role is to gather up the smaller chunks of prosodic structure.

(10) Prosodic adjunction:



A reviewer raises the very pertinent question why the intonational phrase and the phonological phrase need to be different categories. Could the intonational phrase, as well as the utterance, simply be projections of the phonological phrase? This is obviously an empirical question that cannot be decided on a priori grounds. A relevant finding is that there seem to be substantial and 'categorical' (i.e., not merely gradient) differences such as pause and other phonetic correlates identified in Kawahara and Shinya (2008) between these levels of prosody. A similar question can be asked about the prosodic word and the phonological phrase: Might the phrase just be a higher projection of the word? Again this can only be decided on empirical grounds: For example, Kahnemuyipour's (2003) finding that stress in Persian involves opposite headedness at the word and at the phrase level argues that we are dealing with different units of prosody.

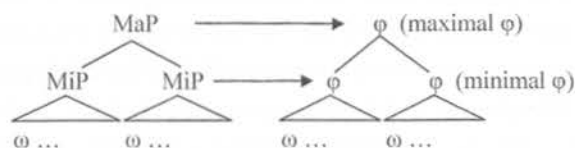
Rather than pursuing a bottom-up phonetic approach, as in Jun (2005), Venditti *et al.* (2008), and other works, where higher prosodic categories are defined solely in terms of their phonetic manifestation, we continue to assume

the syntax-prosody mapping hypothesis in the tradition of prosodic theory, as inaugurated in the work of Selkirk (1978): all interface categories, i.e., prosodic categories above the rhythmically defined foot, are defined in relation to grammatical/syntactic structure and information structure. A reasonable guiding idea is to map *Lex*, the syntactic word (the lexical categories N, V, and A) to the prosodic interface unit  $\omega$ , the syntactic phrase *LexP* (the maximal projection of *Lex*) to  $\phi$  (roughly corresponding to the Major Phrase), and the syntactic clause (CP) to  $\iota$ . This idea is one of the cornerstones of the Match Theory recently developed by Selkirk (2009, to appear).

In connection with defining minor phrasing in Japanese, Selkirk and Tateishi (1988) point out that the syntax-prosody mapping hypothesis may be too strong since it is not the case that there is a syntactic constituent type whose edge must coincide with the edge of a Minor Phrase, nor that a Minor Phrase edge must coincide with the edge of a syntactic constituent of a particular type. Selkirk and Tateishi's observation actually reveals higher ranking constraints at work, such as Accent Culminativity, which dominate the mapping constraints and result in  $\phi$ s that are smaller than XPs. A syntactic phrase  $x_p[a a]$ , where *a* is an accented word, would be expected to map onto a phonological  $\phi$  as  $\phi(a a)$ . But because of Accent Culminativity, the two *a*s cannot both be heads of  $\phi$ , and are optimally mapped as recursive  $\phi(\phi(a)\phi(a))$ . On the other hand, an unaccented combination  $x_p[u u]$  simply maps onto a single phonological  $\phi$  as  $\phi(u u)$ . Thus with recursive prosodic subcategories and violable constraints, there is no need to weaken the syntax-prosody mapping hypothesis – or better, the weakening involved here reveals itself simply as an intrinsically optimality-theoretic effect: The domination of an interface constraint by prosodic constraints.<sup>2</sup>

#### 9.4 Recursive $\phi$ vs. MiP/MaP

How do recursion-based subcategories and MiP/MaP theory match up? In specific instantiations, MiP simply corresponds to the minimal  $\phi$ , MaP to the maximal  $\phi$ , as depicted in (11), leaving us wondering whether we have just recreated MiP and MaP under different names.

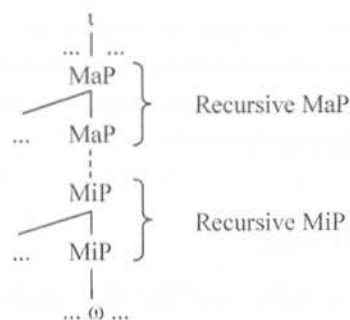
(11) MiP/MaP vs. minimal- $\phi$ /maximal- $\phi$ 

Closer inspection reveals, however, that the two theories are far from being notational variants. There are significant differences between the two, and the evidence favors the single  $\phi$ -category approach. MiP/MaP theory faces a dilemma in that it gives rise to two diametrically opposed problems at the same time: It provides too much structure in some respects, and too little structure in others.

## 9.4.1. Too much structure in MiP/MaP

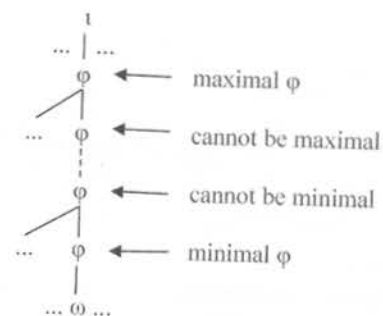
Recursivity is an established attribute of prosodic form – in OT-terms, the anti-recursivity constraint is violable (Selkirk, 1996). But whenever this constraint is low ranking in the grammar, MiP/MaP theory in principle allows both phrases to appear recursively, as in (12).

(12)



Such nested MaP/MiP structures have no equivalent in a theory with recursion-based subcategories. While MaP and MiP are actual categories and can be recursive, 'maximal  $\phi$ ' and 'minimal  $\phi$ ' are relational terms, not separate categories, and it makes no sense for something to be 'recursively maximal' or 'recursively minimal' in a single projection: Only one instance of a category is maximal, and only one is minimal (see (9)).

(13)



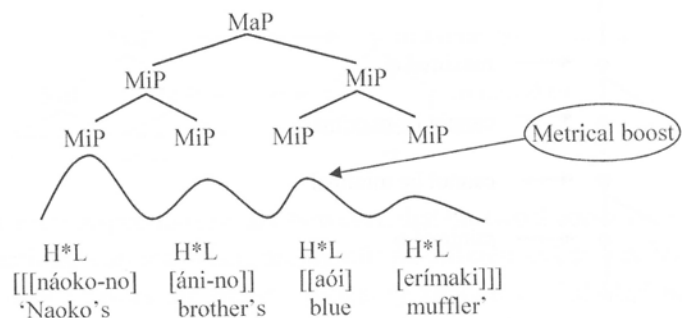
Given the independent need for recursive structures, the single- $\phi$  approach is thus inherently more restrictive than the MiP/MaP approach in (12).

Whether this difference is of consequence remains an empirical question, and will require careful investigation. Interestingly, Selkirk (2000: 25) argues explicitly that in English recursive MaP structures need to be ruled out by specifically assuming high ranking NONRECURSIVITY-MAP. If 'MaP'  $\approx$  maximal  $\phi$ , this follows automatically without invoking other constraints or ranking.<sup>3</sup>

9.4.2 *Too little structure in MiP/MaP*

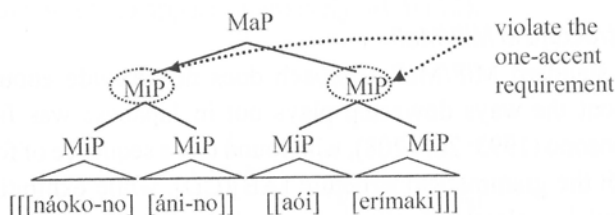
Evidence that the standard MiP/MaP approach does not provide enough structure to represent the ways downstep plays out in Japanese was first pointed out by Kubozono (1993: 205–208), who found that a sequence of four accented MiPs with the grammatical structure [ÁB][ĆD], while exhibiting downstep throughout, i.e., clearly constituting a single MaP, has a systematically higher pitch on Ć than what the flat prosodic structure  $_{\text{MaP}}[\text{MiP MiP MiP MiP}]$  predicts. Kubozono (1989: 58–59) argues that one way of understanding this metrical boost is as a phonetic reflex of a binary, recursively restructured MiP-MiP sequence as in (14).

(14) Recursive MiPs



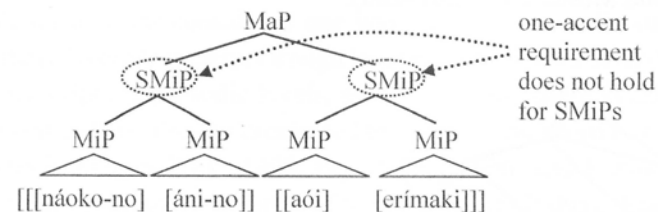
This is a subtle finding with important consequences, and the proposed recursive MiP structure clearly makes sense of the metrical boost, which remains baffling under the standard view.<sup>4</sup> But now a different and unexpected problem arises: Each of the higher MiPs contains two accents, inherited from the two subordinate MiPs, and therefore violates accent culminativity, the defining property of MiP.

(15)



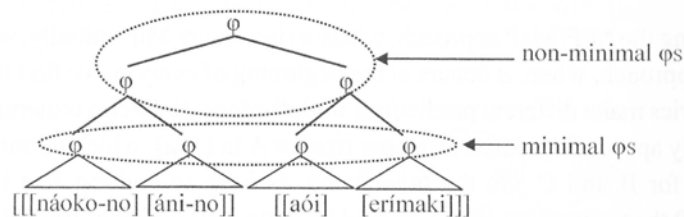
The contradictory domain desiderata – nonrecursive MiPs to observe accent culminativity vs. recursive MiPs to account for the metrical boost – did not go unnoticed (Shinya *et al.*, 2004), and the response was the standard one of introducing yet another category, ‘SMiP’ (‘Superordinate Minor Phrase’), between MiP and MaP. SMiP, and not MiP, takes care of the metrical boost, and the one-accent requirement holds of MiP, not of SMiP.

(16)



In the  $\phi$ -only model, the problematic contradiction does not arise, because the one-accent requirement holds of minimal  $\phi$ , and the branching  $\phi$ s are necessarily non-minimal. The metrical boost, on the other hand, is associated, as in Kubozono’s conception, with  $\phi$  in a right-branching recursive configuration, which is necessarily non-minimal.

(17)

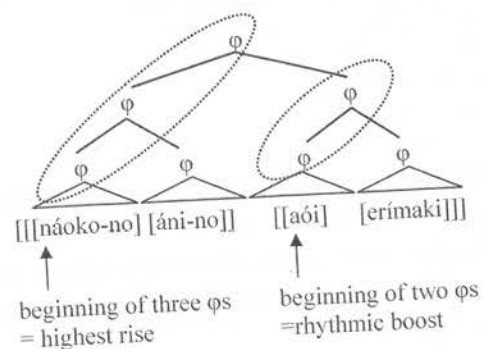


No new intermediate category (such as SMiP) is necessary in the  $\phi$ -only approach. The larger lesson to be learned here is that separate labels become a liability where recursive structures are called for: They require further elaboration of the labeled hierarchy, dimming the prospects for a cross-linguistically valid hierarchy.<sup>5</sup>

Given the prosodic parses assigned by the  $\phi$ -only approach, which are accessed by phonetics, an important question is whether phonetics can ‘count’ the nodes – for example, determine how many instances of  $\phi$  are initiated by a given word. The answer is that such information must in principle be accessible and is of obvious relevance since it is a manifestation of what is often informally referred to as ‘boundary strength’. Thus the first word *náoko* in the example below stands at the beginning of three  $\phi$ s, *aói* at the beginning of two  $\phi$ s, and *áni* and *erímaki* are at the beginning of only one  $\phi$ . Since the initial rise is computed on the basis of beginnings of  $\phi$ s, these distinctions are crucial

(the same basic idea is expressed by the OT-mechanism of constraint conjunction in Ito and Mester 2003: 201–206).

(18)



### 9.4.3 Cumulative rise in recursive φ

Comparing the MiP/MaP approach, where a rise occurs MiP-initially, with the φ-only approach, where it occurs at the beginning of every φ, we find that the two theories make different predictions for left edges. If the rise is cumulative, the φ-only approach predicts a steeper rise for A in (19a), which initiates two φs, than for B and C. On the other hand, as long as nothing else is said, MiP/MaP theory predicts that A should be in no way different from B and C in (19b).

(19)

- a.  $[\phi \ [ \phi \ A ] \ [ \phi \ B ] \ [ \phi \ C ] ]$   
 b.  $[\text{MaP}[\text{MiP} A] \ [ \text{MiP} B ] \ [ \text{MiP} C ] ]$

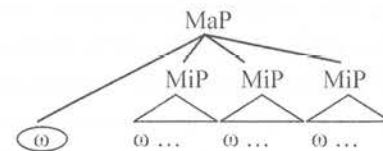
The facts here favor the φ-only approach: Selkirk *et al.* (2003) and Kawahara (2010, this volume) found that the degree of initial rise is more extreme at left edges – a puzzling result for the view that the initial rise is a MiP-exclusive property. Such upwards-inheritance of strengthening effects is a general feature of the prosodic hierarchy, according to Fougeron and Keating (1997) and Flack (2007). What accounts for the different degrees of initial rise? In the φ-only approach, a rise occurs at the beginning of all φs, and if it is cumulative,<sup>6</sup> more rise is immediately predicted at the left edge of structures like (19a) without special pleading. The MiP/MaP approach, on the other hand, needs a separate stipulation that MaP edges have a more extreme rise.<sup>7</sup>

### 9.4.4 Initial rise in weakly layered structures

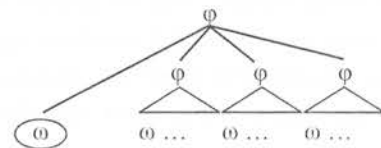
We have so far considered one way in which prosodic structures are not strictly layered, namely, through level repetition (recursivity). A second way is by skipping prosodic levels, as in situations where syllables remain unfooted and are directly dominated by the prosodic word (see Ito and Mester, 2003 [1992]; Selkirk, 1996; Peperkamp, 1996; and Kabak and Schiering, 2006 for examples). In terms of Ito and Mester (2009a), these involve violations of the constraint PARSE-INTO-X, where X is some level of the prosodic hierarchy.<sup>8</sup> It turns out that the MiP/MaP approach and the φ-only approach make different predictions in cases where level-skipping is involved, i.e., when MaP does not begin with MiP and directly dominates φ (skipping the MiP level), as in (20a), to be compared with the structurally equivalent φ-only structure (20b). The circled prosodic word is not MiP-initial in (20a), which would mean no initial rise. On the other hand, it is φ-initial in (20b), predicting a rise. In order to force an initial rise MaP-initially, the MiP/MaP approach must stipulate that every MaP begins with a MiP:  $[\text{MaP}[\text{MiP}\dots]\dots]$ .

(20)

a. MiP/MaP theory:



b. φ-only theory:



No examples are known to us that would directly bear on this issue, but plausibility is certainly on the side of the φ-only theory, especially in light of the fact that the initial rise effects in general increase with the level of phrasing (see the previous section).

### 9.5 Summary and further consequences

In prosodic hierarchy theory, a large number of different interface categories have been proposed in order to provide enough separate domains for different



processes, including *utterance*, *intonational phrase*, *phonological phrase*, *major phrase*, *intermediate phrase*, *minor phrase*, *accentual phrase*, *tone group*, *clitic group*, *prosodic word*, and *minor word*. The totality of these categories has never been instantiated in a single language, however, and their cross-linguistic identification (Does the X-Phrase of Language A really correspond to the X-Phrase of Language B?) has remained a largely unsolved problem. Even within a single language, the insistence on strictly layered representations has led to a considerable multiplication of categories. Whenever a process is found to operate in a slightly different domain than some other process, the model requires setting up two separate categories. Once repetition of levels through recursion becomes an option, however, 'domain' no longer equals 'category', raising the suspicion that perhaps some of the categories proposed in the earlier prosodic literature are in reality only larger recursive structures built on a single basic interface category. Loosening the doctrine of strict layering allows us to strengthen the theory on the category side, and limit the interface categories to a small and universally well-defined set, much like the broadly agreed-on set for rhythmic categories (foot, syllable, and mora).

In this paper, we reviewed the evidence for the distinction between the central interface categories proposed for Japanese, major phrase and minor phrase. While everyone agrees that constituents of different sizes are involved, we have argued here that there is no need to postulate two separate interface categories, and that the evidence favors a model with a single category  $\phi$  'phonological phrase', with the option of recursion. Since the  $\phi$ -only theory is more restrictive, given that recursion has been shown to be necessary on independent grounds, it is up to the proponents of theories with a larger number of interface categories to show that additional categories are in fact necessary.

In a more general vein, we hypothesized that there are only three universal interface categories: *intonational phrase* ( $t$ ), *phonological phrase* ( $\phi$ ), and *prosodic word* ( $\omega$ ). Additional structure is imposed on the string through recursion.

Investigations along these lines, where relational notions such as maximal and minimal projections of categories play a natural role, have resulted in some interesting and promising results in a variety of languages. Woodbury (2002) shows in detail, using both prosodic and segmental phenomena as evidence, that Cup'ik prosody distinguishes two 'word' constituents; The gram-

matical word minus all enclitics, and the grammatical word plus all enclitics. In our terms, the two domains ('PW-' vs. 'PW' in Woodbury's notation) correspond exactly to the minimal and the maximal prosodic word ( $\omega$ ), respectively. The analysis of the word-level stress domain and the *Ezafe* construction in Persian (Kahnemuyipour, 2003) can be interpreted along similar lines, with an appeal to minimal and maximal projections of  $\omega$ .

In work on the prosody of Irish, Dowd (2009) has identified the maximal  $\omega$  as the domain of synthetic agreement; and Bennett and McCloskey (2008) show with several diagnostics (including pause, downstep, and lengthening) that phonological phrases ( $\phi$ ) are right-aligned to XP, and that the syntactically baffling distribution of weak pronouns has a prosodic rationale (see also Elfner, 2008): They appear at the right edge of the maximal  $\phi$ . Henderson (2008) shows that in K'ichee' *h*-final allomorphs occur as phonological phrase markers. In conjoined phrases, they are optionally found on the first conjunct, which can be understood as recursive  $\phi$ -phrasing, with *h*-final allomorphs appearing either on the minimal or the maximal  $\phi$ .

In our own work, we have shown that a theory with recursion-based subcategories can be fruitfully applied to illuminate the interplay between *rendaku*, accent, downstep, and initial rise in Japanese complex compounds (Ito and Mester, 2007), the prosodic conditioning of intrusive-*r* in non-rhotic English (Ito and Mester, 2009b), and the phonology of English and German function word complexes, which involves various segmental fusion processes (Ito and Mester, 2009a).

While these preliminary results inspire some confidence in the general outlook on prosody presented here, it remains to be seen in future work whether a theory with such a minimal set of interface categories and recursion-based subcategories can be upheld cross-linguistically.

## Notes

- \* An earlier version of this paper appeared in the conference proceedings of Japanese/Korean Linguistics 18 (CSLI, Stanford). The present version takes into consideration the comments and critiques of the reviewers and editors of the *Festschrift* volume. For helpful comments and suggestions on this work, we would like to thank Mary Beckman, Ryan Bennett, Kenneth de Jong, Andrew Dowd, Caroline Féry, Steven Franks, Shosuke Haraguchi, Shin Ishihara, Dan Kaufman, Shigeto Kawahara, Yoshi Kitagawa, Haruo Kubozono, Jim McCloskey, Bruce

Morén, Marina Nespó, Jeremy O'Brien, Marc van Oostendorp, Dave Odden, Jaye Padgett, Anthi Revithiadou, Tomas Riad, Curt Rice, Shin-Ichi Tanaka, and Satoshi Tomioka.

- As shown by time-honored examples like the following (after Chomsky and Halle, 1968), where syntactic structure and  $\iota$ -phrasing are radically different:

$\iota$ ( This is the cat ) $\iota$   $\iota$ ( that chased the rat ) $\iota$   $\iota$ ( that ate the malt ) $\iota$   
 CP[ DP[ CP[ DP[ CP[ DP[ ]DP]CP]DP]CP]DP]CP

All embedding is removed in the prosodic representation, resulting in a flat sequence of three  $\iota$ -phrases whose left edges coincide with those of clauses (CPs), and not of the syntactically superordinate DPs.

- The basic idea here goes back to the earliest work on minor phrasing in McCawley (1968: 177–180), who inserts minor phrase boundaries by a strictly syntax-based rule and then adjusts the phrasing with rules sensitive to (un)accentedness. Just as minor phrases can be understood as part of recursive  $\phi$ s, Clitic Groups have been argued to be recursive  $\omega$ s in Ito and Mester (2009a), taking up an earlier proposal by Inkelas (1989), among others.
- Kratzer and Selkirk (2007) use recursive MaPs in explaining the accentual variation found with main clause word order in German. MaP recursivity is not essential to the proposal, however, which can be straightforwardly recast with extended word projections ( $\omega$ -recursion) for MiP and  $\phi$ -phrase recursion for MaP (see Ito and Mester, 2008).
- Among other things, it involves the recognition that the notion downstep has a *paradigmatic* component, besides the obvious *syntagmatic* one: Kubozono argues that  $\acute{C}$  in [ÁB][ĈD] in (14) is downstepped relative to  $\acute{B}$ , even though at roughly the same pitch as  $\acute{B}$ , because it is significantly lower in pitch than  $\acute{C}$  in [ÁB][ĈD], where it follows unaccented B.
- See also Wagner (2005, 2010) for a more radical departure from the standard labeled hierarchy, with arguments for a 'label-free' purely metrical model of prosodic structure.
- For concreteness, assume that each  $\phi$  begins with %L, a lowering operator, and that stacked lowering operators are interpreted in a cumulative way.
- A reviewer reminds us that Kawahara and Shinya (2008) found even more extreme rises at  $\iota$ -edges than at  $\phi$ -edges. This does not mean, however, that  $\phi$ -only theory, like MiP/MaP theory, still has to have recourse to separate stipulations of the same property for separate categories. The slogan of MiP/MaP theory – and, more generally, of theories insisting on strictly layered structures – is that each level is

identified with a particular phonetic property, constituting its 'domain': MiP is the domain of the initial rise, MaP the domain of downstep, etc. In such a theory, it is odd to find more pronounced rises at the left edge of MaP, which is not a rise domain. On the other hand, since our theory does not correlate 'domain' and 'phonetic property' in this way, we do not expect a one-to-one correspondence between them. In particular, a specific phonetic property can characterize several domains. Intonational phrases and phonological phrases are both phrases, and nothing is more natural than certain properties, such as the presence of %L, being associated with all phrasal categories.

- Or in the terminology of Selkirk (1996), violations of EXHAUSTIVITY<sub>Y</sub>, which demands that the higher category Y dominate only X.

## References

- Beckman, Mary and Pierrehumbert, Janet. (1986) Intonational structure in English and Japanese. *Phonology* 3: 255–309.
- Bennett, Ryan and McCloskey, James. (2008) Donegal Irish and the Syntax/Prosody Interface. UC Santa Cruz. Presented at PRIG (Prosody Interest Group), UC Santa Cruz.
- Booij, Geert. (1996) Cliticization as prosodic integration: The case of Dutch. *The Linguistic Review* 13: 219–242.
- Chomsky, Noam and Halle, Morris. (1968) *The Sound Pattern of English*. New York: Harper & Row.
- Dowd, Andrew. (2009) Aspects of the Prosody of Irish Synthetic and Analytic Forms. UC Santa Cruz. Presented at PRIG (Prosody Interest Group), UC Santa Cruz.
- Elfner, Emily. (2008) The Interaction of Linearization and Prosody: Evidence from Pronoun Postposing in Irish and Scottish Gaelic. Amherst, MA: University of Massachusetts.
- Elordieta, Gorka. (2007) Minimum size constraints on intermediate phrases. <http://www.icphs2007.de/conference/Papers/1682/1682.pdf>.
- Féry, Caroline and Truckenbrodt, Hubert. (2005) Sisterhood and tonal scaling. *Studia Linguistica* 59: 223–243.
- Flack, Kathryn Gilbert. (2007) *The Sources of Phonological Markedness*. University of Massachusetts: Doctoral dissertation.
- Fougeron, Cécile and Keating, Patricia. (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101: 3728–3740.
- Gussenhoven, Carlos. (1991) The English rhythm rule as an accent deletion rule. *Phonology* 8: 1–35.
- Gussenhoven, Carlos. (2005) Procliticized phonological phrases in English: Evidence from rhythm. *Studia Linguistica* 59: 174–193.

- Halliday, M. A. K. (1960) Categories of the theory of grammar. *Word* 17 (3): 241–292.
- Haraguchi, Shosuke. (1977) *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology*. Tokyo: Kaitakusha.
- Henderson, Robert. (2008) Allomorphy and K'ichee' Prosodic Structure. UC Santa Cruz. Presented at PRIG (Prosody Interest Group), UC Santa Cruz.
- Inkelas, Sharon. (1989) *Prosodic Constituency in the Lexicon*, Stanford University: Doctoral dissertation. Published 1990, Outstanding Dissertations in Linguistics Series. New York: Garland Press.
- Ito, Junko and Mester, Armin. (2003) *Japanese Morphophonemics: Markedness and Word Structure*. Linguistic Inquiry Monographs, 41. Cambridge, MA, and London, England: MIT Press.
- Ito, Junko, and Mester, Armin. (2003 [1992]) Weak layering and word binarity. In Takeru Honma, Masao Okazaki, 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 Sixtieth Birthday* 26–65. Slightly revised version of 1992 UC Santa Cruz Linguistics Research Center working paper.
- Ito, Junko and Mester, Armin. (2007) Prosodic adjunction in Japanese compounds. In *Formal Approaches to Japanese Linguistics* 4 97–111. Cambridge, MA: MIT Department of Linguistics and Philosophy.
- Ito, Junko and Mester, Armin. (2008) Rhythmic and Interface Categories in Prosody. UC Santa Cruz. Presented at PRIG (Prosody Interest Group), UC Santa Cruz.
- Ito, Junko and Mester, Armin. (2009a) The extended prosodic word. In Janet Grijzenhout and Barış Kabak (eds) *Phonological Domains: Universals and Deviations*, Berlin: Mouton de Gruyter.
- Ito, Junko and Mester, Armin. (2009b) The onset of the prosodic word. In Steve Parker (ed.) *Phonological Argumentation: Essays on Evidence and Motivation* 227–260. London: Equinox.
- Jun, Sun-Ah. (1998) The Accentual Phrase in the Korean prosodic hierarchy. *Phonology*. 15: 189–226.
- Jun, Sun-Ah. (2005) Prosodic typology. In Sun-Ah Jun (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing* 430–458. Oxford: Oxford University Press.
- Kabak, Barış and Schiering, René. (2006) The phonology and morphology of function word contractions in German. *Journal of Comparative Germanic Linguistics* 9: 53–99.
- Kahnemuyipour, Arsalan. (2003) Syntactic categories and Persian stress. *Natural Language and Linguistic Theory* 21 (2): 333–379.
- Kawahara, Shigeto. (2011) The intonation of nominal parentheticals in Japanese. In Toni Borowsky, Shigeto Kawahara, Takahito Shinya and Mariko Sugahara (eds) *Prosody Matters* 320–356. London: Equinox.
- Kawahara, Shigeto and Shinya Takahito. (2008) The intonation of gapping and coordination in Japanese: Evidence for Intonational Phrase and Utterance. *Phonetica* 65: 62–105.
- Kratzer, Angelika, and Selkirk, Elisabeth. (2007) Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24: 93–135.
- Kubozono, Haruo. (1989) Syntactic and rhythmic effects on downstep in Japanese. *Phonology* 6: 39–67.
- Kubozono, Haruo. (1988) *The Organization of Japanese Prosody*, University of Edinburgh: Doctoral dissertation.
- Kubozono, Haruo. (1993) *The Organization of Japanese Prosody*. vol. 2: Studies in Japanese Linguistics. Tokyo: Kurocio Publishers. Revision of 1988 University of Edinburgh thesis.
- Kubozono, Haruo. (2005) Focus and intonation in Japanese: Does focus trigger pitch reset?. In S. Ishihara, M. Schmitz and A. Schwarz (eds) *Interdisciplinary Studies on Information Structure (Working Papers of the Sonderforschungsbereich 632)*. Potsdam: Universitätsverlag Potsdam.
- Ladd, D. Robert. (1986) Intonational phrasing: The case for recursive prosodic structure. *Phonology* 3: 311–340.
- Ladd, D. Robert. (1988) Declination 'reset' and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84 (2): 530–544.
- Ladd, D. Robert. (1996) *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladd, D. Robert, and Campbell, Nick. (1991) Theories of prosodic structure: Evidence from syllable duration. In *Proceedings of the 12th International Congress of the Phonetic Sciences* 290–293. Aix-En-Provence, France.
- McCawley, James D. (1968) *The Phonological Component of a Grammar of Japanese*. The Hague, The Netherlands: Mouton.
- Nespor, Marina and Vogel, Irene. (1983) Prosodic structure above the word. In A. Cutler and D. R. Ladd (eds) *Prosody: Models and Measurements* 123–140. Berlin, Heidelberg, New York, Tokyo: Springer.
- Nespor, Marina, and Vogel, Irene. (1986) *Prosodic Phonology*. Dordrecht: Foris.
- Peperkamp, Sharon. (1996) On the prosodic representation of clitics. In Ursula Kleinhenz (ed.) *Interfaces in Phonology* 102–127. Berlin: Akademie Verlag.
- Petrone, Caterina, and D'Imperio, Mariapaola. (2008) Tonal structure and constituency in Neapolitan Italian: Evidence for the Accentual Phrase in statements and questions. ISCA Archive, <http://www.isca-speech.org/archive>.
- Pike, Kenneth L. (1967) *Language in Relation to a Unified Theory of the Structure of Human Behavior*. Janua Linguarum series major 24. The Hague: Mouton.
- Poser, William J. (1984) *The Phonetics and Phonology of Tone and Intonation in Japanese*. MIT: Doctoral dissertation.
- Schreuder, Maartje. (2006) *Prosodic Processes in Language and Music*. Rijksuniversiteit Groningen, The Netherlands: Doctoral dissertation.

- Selkirk, Elisabeth. (1978) On prosodic structure and its relation to syntactic structure. In T. Fretheim (ed.) *Nordic Prosody* 111–140. Trondheim: TAPIR.
- Selkirk, Elisabeth. (1984) *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, Elisabeth. (1996) The prosodic structure of function words. In James L. Morgan and Katherine Demuth (eds) *Signal to Syntax* 187–213. Mahwah, NJ: Lawrence Erlbaum.
- Selkirk, Elisabeth. (2005) Comments on intonational phrasing in English. In Sónia Frota, Marina Vigário and Maria João Freitas (eds) *Prosodies: Selected Papers from the Phonetics and Phonology in Iberia Conference, 2003 (Phonetics and Phonology Series)*. Berlin and New York: Mouton de Gruyter.
- Selkirk, Elisabeth. (2009) On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyū* 136: 35–73.
- Selkirk, Elisabeth. (to appear) The syntax-phonology interface. In John Goldsmith, Jason Riggle and Alan Yu (eds) *The Handbook of Phonological Theory. 2nd edition*. Oxford: Blackwell Publishing.
- Selkirk, Elisabeth, Shinya, Takahito and Sugahara, Mariko. (2003) Degree of initial lowering in Japanese as a reflex of prosodic structure organization. In *Proc. 15th ICPhS, Barcelona, Spain*. 491–494.
- Selkirk, Elisabeth and Tateishi, Koichi. (1988) Constraints on minor phrase formation in Japanese. In *Papers from the 24th Annual Regional Meeting of the Chicago Linguistic Society* 316–336. Chicago, IL: Chicago Linguistic Society.
- Selkirk, Elisabeth, and Tateishi, Koichi. (1991) Syntax and downstep in Japanese. In C. Georgopoulos and R. Ishihara (eds) *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda* 519–544. Dordrecht: Kluwer.
- Shinya, Takahito, Selkirk, Elisabeth and Kawahara, Shigeto. (2004) Rhythmic boost and recursive minor phrase in Japanese. In *Proceedings of the Second International Conference on Speech Prosody* 345–348.
- Truckenbrodt, Hubert. (1999) On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30: 219–255.
- Truckenbrodt, Hubert. (2002) Upstep and embedded register levels. *Phonology* 19: 77–120.
- Truckenbrodt, Hubert. (2006) The syntax-phonology interface. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology* 435–456. Cambridge: Cambridge University Press.
- van den Berg, Rob, Gussenhoven, Carlos, and Rietveld, Toni. (1992) Downstep in Dutch: Implications for a model. In Gerard J. Dochtery and D. Robert Ladd (eds) *Papers in Laboratory Phonology, Vol. II: Gesture, Segment, Prosody* 335–358. Cambridge: Cambridge University Press.
- Venditti, Jennifer J., Maekawa, Kikuo and Beckman, Mary E. (2008) Prominence marking in the Japanese intonational system. In Shigeru Miyagawa and Mamoru

- Saito (eds) *The Oxford Handbook of Japanese Linguistics* 456–512. Oxford: Oxford University Press.
- Wagner, Michael. (2005) *Prosody and Recursion*, Department of Linguistics and Philosophy, MIT: Doctoral dissertation.
- Wagner, Michael. (2010) Prosody and recursion in coordinate structures and beyond. *Natural Language and Linguistic Theory* 28 (1): 183–237.
- Woodbury, Anthony. (2002) The word in Cup'ik. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds) *Word: A Cross-linguistic Typology* 79–99. Cambridge: Cambridge University Press.