Available online at www.sciencedirect.com

SciVerse ScienceDirect

Lingua 124 (2013) 20–40

Lingua

www.elsevier.com/locate/lingua

Prosodic subcategories in Japanese

Junko Ito [a,b,*], Armin Mester [a,c]

[a] University of California, Santa Cruz, USA
[b] International Christian University, Mitaka, Japan
[c] National Institute for Japanese Language and Linguistics, Tachikawa, Japan

Received 10 July 2011; received in revised form 26 August 2012; accepted 29 August 2012
Available online 8 October 2012

Abstract

Research on Japanese prosody, especially on the pitch accent system of the language, has for a long time found that a single domain ''phonological phrase'' is not sufficient. Rather, two domains need to be distinguished, which go by various names (Minor vs. Major Phrase, Accentual vs. Intermediate Phrase). While empirically well-founded, these developments, together with similar findings in other languages, have resulted in a dissolution of the originally tightly organized universal prosodic hierarchy into a collection of many prosodic types, each instantiated here and there in different languages, but never simultaneously realized within a single language. Two strands of recent work, that of Selkirk (2009:205–219, 2011a), and of Ito and Mester (2007, 2009a, 2009b) converge on a common theme from different directions. On the one hand, Selkirk has developed a vastly simplified approach to the syntax–prosody mapping which distinguishes only three levels (word, phrase, and clause) where syntactic constituents are systematically made to correspond to phonological domains (''Match Theory''). On the other hand, Ito and Mester have argued that the empirically necessary subcategories (such as Minor vs. Major Phrase) need to be understood not as additional categories existing in their own right, but rather as instances of recursively deployed basic categories. This paper carries forward this line of prosodic hierarchy research and shows that the recursion-based conception implemented within Match Theory allows for a conceptually and empirically cleaner understanding of the phonological facts and generalizations in Japanese as well as for an understanding of the respective roles of syntax and phonology in determining prosodic constituent structure organization, and the limitation in types of distinctions in prosodic category that are made in phonological representation. Finally, a formal constraint-based OT analysis is developed that provides an account of the varying tonal and accentual structures in syntactic collocations of varying sizes and structures.

Keywords: Prosodic hierarchy; Syntax–prosody mapping; Recursion in prosody; Match Theory; Minimal/maximal projections; Optimality Theory

1. Introduction: The syntax–prosody mapping hypothesis (SPMH)

In the tradition of Prosodic Hierarchy Theory inaugurated by Selkirk and others (Selkirk, 1978, 1980; Nespor and Vogel, 1983, 1986, etc.), the *syntax–prosody mapping hypothesis* (SPMH) holds that each constituent in prosodic structure (above the rhythmically defined foot) has a corresponding designated constituent in syntactic structure, and is defined in relation to it. Thus, the phonological word ω is defined in relation to the syntactic word (the lexical categories N, V, and A), the phonological phrase φ in relation to the syntactic phrase (maximal projection of such lexical categories), etc. Things are more complex at the intonational phrase level (ι), where differences in information structure, speed, and level of formality result in a large range of alternative parses (Bierwisch, 1966; Selkirk, 1984), but it seems clear that the syntactic clause serves as a basic anchoring point (Hirst, 1993; Selkirk, 2011a).

* Corresponding author at: Department of Linguistics, Stevenson Academic Services, UC Santa Cruz, Santa Cruz, CA 95064, USA. Tel.: +1 831 459 3340; fax: +1 831 459 3334.
E-mail address: ito@ucsc.edu (J. Ito).

0024-3841/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.lingua.2012.08.016

On the other hand, research on individual languages, in particular work on the Japanese pitch accent system, has for a long time found that a single prosodic constituent ''phonological phrase'' is not sufficient. Rather, two domains need to be distinguished which go by various names (Minor vs. Major Phrase, Accentual vs. Intermediate Phrase). This was already established in the pioneering contributions of Martin (1952) and McCawley (1968), with ample further motivation added in later work (Poser, 1984; Kubozono, 1988; Pierrehumbert and Beckman, 1988, among others). While empirically well-founded, this bifurcation at the phonological phrase level actually poses a serious challenge to the SPMH because, different from the Major Phrase, no syntactic category can be given as corresponding to the Minor Phrase, as shown in (1).

(1)
|
Major Phrase: corresponding syntactic category: *XP*
|
Minor Phrase: corresponding syntactic category: ????
|
Phonological Word: corresponding syntactic category: *X* (*N*, *V*, *A*)
|

The problem was already diagnosed by Selkirk and Tateishi (1988:332): ''[. . .] if the Minor Phrase is indeed a prosodic constituent, it is one unlike Phonological Word and Major Phrase, which have a strict correspondence to syntactic structure. 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. [. . .] The Minor Phrase is defined independently of syntactic structure.''

Selkirk and Tateishi (1988) go on to provide several constraints that define the Minor Phrase, such as the Accent Condition, the Ternary Branching Condition, and the Peripherality Constraint, whose definitions are purely phonological and not syntactic. If such phonological conditions are the defining characteristics of the Minor Phrase, then a serious question arises as to the viability of the SPMH. One may indeed reasonably interpret the situation as a vindication of a bottom-up phonetic approach, as advocated in Jun (2005), Venditti et al. (2008), and related work, where higher prosodic categories are set up, on a case-by-case basis, solely in terms of their phonetic manifestation, and where no attempt is made to define them in relation to syntactic structure.

It is somewhat surprising that most subsequent research has not taken up this issue as a problem and has tacitly gone on to adopt both the Minor Phrase and some necessarily weaker version of the SPMH, without questioning the hierarchy itself that contains the Minor Phrase sandwiched between the Major Phrase and the Phonological Word. In this paper, we argue that the strong version of the SPMH, with straightforward correspondences between syntactic and prosodic categories, as in (2), can be upheld, provided we recognize that the Minor Phrase/Major Phrase distinction does not involve two distinct categories, but rather two layers of a recursively deployed single category ''Phonological Phrase''.

(2)
¦
Phonological Phrase (φ): corresponding syntactic category: *XP*
|
Phonological Word (ω): corresponding syntactic category: *X* (*N*, *V*, *A*)
¦

The approach requires a better understanding not only of the way prosodic categories are recursively assigned to parse a syntactic string, matching major structural features of syntactic constituency in a well-defined way, but also of the way different constraints interact in a system of priority-ranked constraints, where syntax–prosody mapping constraints are sometimes dominated by, and hence overruled by, purely phonological constraints. Optimality Theory (Prince and Smolensky, 1993 (2004)) is indispensible for a proper and precise understanding of this kind of interaction.

2. Background

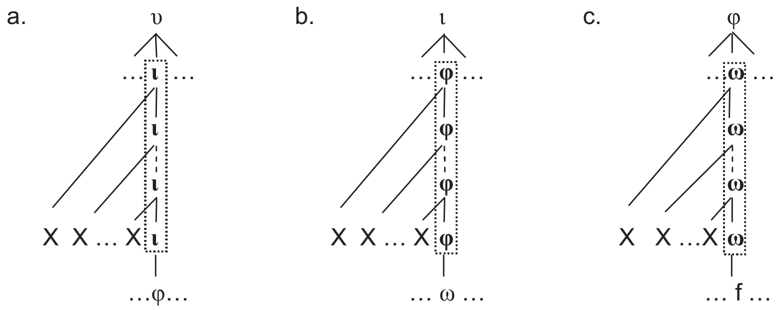
Prosodic Hierarchy Theory holds that speech is organized into a set of genuinely phonological domains that form a hierarchy of containment, with each non-terminal constituent made up of a sequence of smaller constituents at the next level down. Earlier work on the prosodic hierarchy included only the levels starting with the phonological word and above (see Hayes, 1989 for a summary), but with the development of Prosodic Morphology in McCarthy and Prince (1986 et sqq.) and related work dealing with morpho-phonological issues involving the rhythmic units of feet, syllables, and moras, it has become standard to include these smaller phonological constituents that make up the phonological word as part of the standard prosodic hierarchy.

The guiding principle (named the *Strict Layer Hypothesis*[1] in Selkirk, 1984:26) that these prosodic levels cannot be *skipped* or *repeated* was first seen as inviolable, and later reduced to a set of several violable constraints (Ito and Mester, 1992; Selkirk, 1996, the latter in the context of Optimality Theory (OT)). A constraint that bans the *skipping* of levels has been proposed under different names, such as Strict Layering, Exhaustivity, Parse-into-X, with slightly different definitions. Similarly, the constraint banning *repetition* of levels appears as Recursivity or NoRecursion, defined in various ways suited for the task at hand.

While the theory, and the field at large, moved towards violable OT constraints (Truckenbrodt, 1999, etc.), the prosodic categories themselves, that is, the building blocks that define the hierarchy, continued for the most part to be taken for granted. In Ito and Mester (2009a), we point out that the proliferation of prosodic categories, each empirically well-founded in specific cases, has resulted in a dissolution of the original tightly organized universal hierarchy into an ungainly collection of a large number of prosodic types, each instantiated here and there in different languages but never simultaneously realized within a single language. We go on to argue that the problematic proliferation can be avoided if many of the empirically necessary levels (such as Minor vs. Major Phrase) are understood *not* as additional categories existing in their own right, but rather as *prosodic subcategories* of recursively deployed basic categories (see Ito and Mester, 2007, 2009a, 2009b, and also Selkirk (2009, 2011a), where the essentials of this proposal are adopted).

The basic idea, which we take up in this paper, is that recursive prosodic structures as in (3) come into existence under the pressure of high-ranking constraints in the grammar. Such prosodic level repetition is necessitated by the syntax–prosody interface (whether conceived of as classical syntax–prosody alignment constraints, or as the Match constraints recently proposed by Selkirk), or by the phonology itself, for example, in the form of high-ranking parsing constraints requiring all segmental material, including functional (non-lexical) items, to be incorporated into prosodic structure without violating the interface constraint prohibiting prosodic constituents that are not grounded in lexical material. Evidence for such recursive structures appears, for example, in Ladd, 1986, 1988 (recursion of ι in nested coordinated structures in English), Gussenhoven, 1991, 2005 (recursion of φ in the parsing of Dutch and English noun phrases, as shown by the operation of the rhythm rule), and Ito and Mester, 2007, 2009a, 2009b (recursion of ω in Japanese compounds, and in English and German function word complexes).

(3) Recursion-based subcategories of ι, φ, and ω



In such recursive prosodic structures, the standard tree-structural notions apply, such as *head* vs. *non-head* and *maximal projection* vs. *minimal projection*: In $[_{\alpha}\ \alpha\ \beta]$ and $[\beta\ \alpha\ _{\alpha}]$, the smaller α is the (structural) *head* of the larger α; the topmost projection of α—i.e., α not dominated by α—is the *maximal* α, the lowest projection—i.e., α not dominating α—the *minimal* α.

(4) Definitions from Ito and Mester (2007, 2009a, 2009b)
- a. maximal (projection of) $\alpha =_{def} \alpha$ not dominated by α
- b. minimal (projection of) $\alpha =_{def} \alpha$ not dominating α

The binary projection features [$\pm$max] and [$\pm$min] proposed in Haider (1993:40, see also work cited there) and defined in exactly this way are a convenient means to represent, for any given category, the natural classes of recursive subcategories given in (5), using φ as an example.

[1] To our knowledge, it was first proposed as a constraint on prosodic constituent structure representation in Selkirk (1981:382), who observes that ‘‘[t]he combinatorial possibilities of prosodic categories are far more restricted than those of syntactic categories. [. . .] In general, it is true that a prosodic category of level *n* is defined as consisting of categories of level *n-1*.’’

(5) Projections of "phrase" (φ)

$$
\begin{array}{llll}
 & \iota & & \\
 & | & & \\
\varphi^{[-\text{min}]}\text{, the non-minimal projections} \left\{ \begin{array}{l} \\ \\ \end{array} \right. & \varphi & \} & \varphi^{[+\text{max},-\text{min}]}\text{, the maximal projection} \\
 & | \diagdown & & \\
\varphi^{[-\text{max}]}\text{, the non-maximal projections} \left\{ \begin{array}{l} \\ \\ \end{array} \right. & \varphi \ldots & \} & \varphi^{[-\text{max},-\text{min}]}\text{, an intermediate projection} \\
 & | \diagdown & & \\
 & \varphi \ldots & \} & \varphi^{[-\text{max},+\text{min}]}\text{, the minimal projection} \\
 & | & & \\
 & \omega & &
\end{array}
$$

Besides targeting (every instance of) a specific prosodic category, domain-sensitive processes can specifically target maximal and minimal projections within such recursion-based networks of subcategories, often removing the need, and thus the motivation, for further distinctions among basic categories. The system in (5) provides additional natural classes of subcategories, such as the non-minimal projections motivated by Elfner (2010) for Irish. In addition, as pointed out by Haider, $\varphi^{[+\text{max},\ +\text{min}]}$ denotes the trivial phrase consisting of a single projection, which is therefore both maximal (not dominated by another φ) and minimal (not dominating another φ). Exactly these types of projections were necessary to properly understand the different types of Japanese compounds in Ito and Mester (2007), e.g., $\omega^{[+\text{min}]}$ as target of rendaku-voice, $\omega^{[+\text{max}]}$ as the site of junctural accent, and other tonal restrictions (initial rise, one vs. two accents) captured through differences in φ-structure (monophrasal vs. biphrasal compounding).[2]

We see our proposal as complementing recent work by Selkirk (2009, 2011a), who has developed a vastly simplified approach to the syntax–prosody mapping which distinguishes only three levels (word, phrase, and clause) where syntactic constituents are systematically made to correspond to phonological domains (''Match Theory'').

This paper has several interrelated goals. Focusing on the Minor Phrase in Japanese, we show that the theory of recursion-based prosodic subcategories allows for a conceptually and empirically cleaner understanding of the facts and generalizations (section 2). This new understanding in turn makes it possible for the syntax–prosody mapping hypothesis (SPMH) to be maintained in its strong form, where all prosodic categories are grounded in syntactic categories (section 3). Finally, we present a system of OT-constraints that captures the interplay of the different factors involved in an illuminating way (sections 3 and 4).

3. Major Phrase and Minor Phrase: downstep and initial rise

Ito and Mester (2009b) propose that the standardly assumed Major Phrase and Minor Phrase are both instantiations of a single phonological phrase category φ.[3] The distinction had traditionally been viewed as irreducible (McCawley, 1968; Haraguchi, 1977; Poser, 1984; Kubozono, 1988; Pierrehumbert and Beckman, 1988; Selkirk and Tateishi, 1988, etc.) because each is a domain for a distinct phonetic process (6).

(6) Domain Requirements:
 (i) The Major Phrase (MaP) is the domain of downstep (lowering of the pitch register following each accented syllable).
 (ii) The Minor Phrase (MiP) is the domain of Initial Rise %LH (Low tone at left boundary followed by phrasal High tone).

Downstep is a compression of the pitch register after a lexical accent H*L (but not after other HL sequences) that affects all following tonal material, whether accented or unaccented, within a local domain, the MaP. At the beginning of the next MaP, the pitch register is reset. More precisely, the diagnosis of downstep requires the paradigmatic comparison of the pitch height of X, accented or unaccented, in AX (after accented A) and in UX (after unaccented U). If the pitch height of X is, ceteris paribus, significantly lower after A than after U, X is said to be downstepped after A. Below we reproduce

---

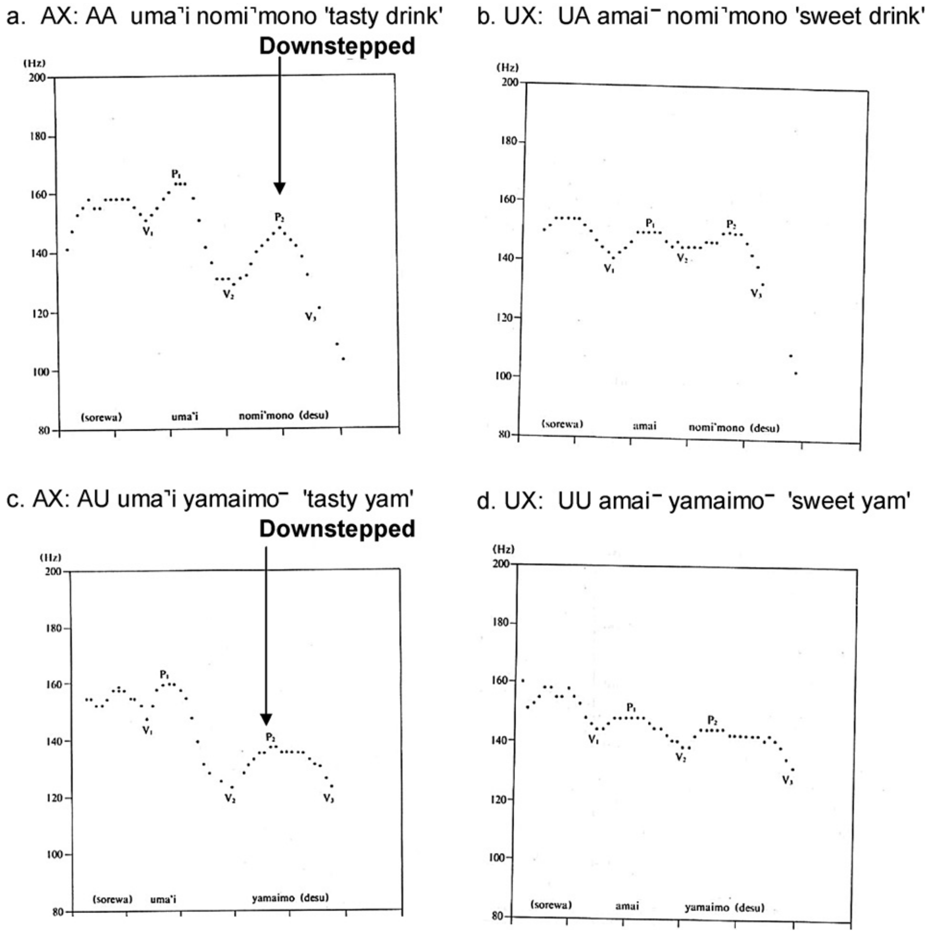
[2] A reviewer asks whether there is a hidden assumption such that if φ_1 immediately dominates one or more φ, exactly one of the dominated φ must be analyzed as $\varphi^{[+\text{min}]}$ of φ_1. There is no such assumption, and the definition of a minimal φ is no more and no less than that given above, i.e., $\varphi^{[+\text{min}]}$ is a ''φ not dominating another φ'', and there can potentially be many $\varphi^{[+\text{min}]}$ immediately dominated by a φ. For example, in the following structure, only φ_1 is [−min], and both φ_2 and φ_3 are [+min].

$$
\begin{array}{cc}
\multicolumn{2}{c}{\varphi_1} \\
/ & \backslash \\
\varphi_2 & \varphi_3 \\
| & | \\
\omega & \omega
\end{array}
$$

[3] In related work (Ito and Mester, 2009a), we argue that the distinction between prosodic word and clitic group should also be abolished in favor of the category prosodic word and its recursively adjoined structures.

comparative pitch tracks for two relevant pairs of examples from Kubozono (1988:261–262). We use the traditional accent corner (˥) after a vowel to indicate the accentual fall and a superscripted straight line at the end of a word (‾) to indicate that it is unaccented. In each case, downstep can be seen in the first example AX (AA, AU) (7a, c) in comparison with the corresponding second example UX (UA, UU) (7b, d).

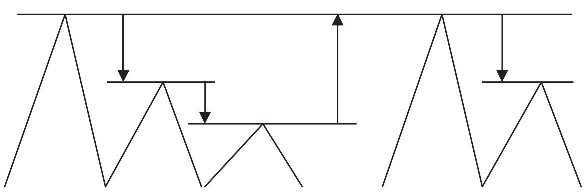
(7) Comparative pitch tracks from Kubozono (1988:261–262)



Kubozono (2007) makes a compelling case that this kind of paradigmatic understanding of downstep, which is the one introduced by Poser (1984) and used in Kubozono's own work and that of Pierrehumbert and Beckman, is superior to a syntagmatic understanding (as in Selkirk and Tateishi, 1991 and Nagahara, 1994) based on a direct comparison of the pitch of successive phrases in the same utterance (e.g., in (7a) or (7c)).

The schematic picture in (8) (after Ishihara, 2011:1872) also illustrates the chaining of downstep (within its domain), as experimentally shown by Poser (1984). Like (7), it also depicts the crucial fact that downstep occurs in a larger domain (MaP) than Initial Rise, which is found at the beginning of each MiP.

(8)



($_{\text{MaP}}$ <$_{\text{MiP}}$H*L> <$_{\text{MiP}}$ H*L> <$_{\text{MiP}}$ H*L>) ($_{\text{MaP}}$ <$_{\text{MiP}}$ H*L> <$_{\text{MiP}}$ H*L>)

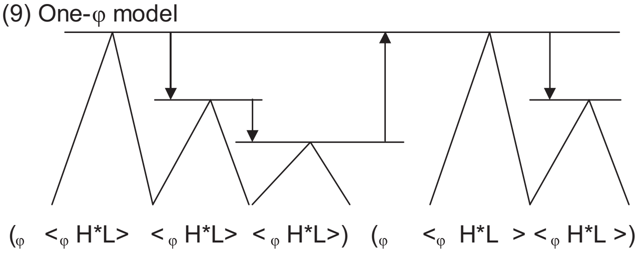
Since the post-accentual L is consistently and significantly lower than the pre-accentual L, Poser (1984:319–320), noting the limited nature of his evidence, suggests that the locus of downstep is within the accent itself, ''at the transition between the High and Low, so that the Low tone that plays a role in triggering catathesis[4] would be the first tone to undergo it.'' This proposal, tentatively endorsed by Pierrehumbert and Beckman (1988:88), strikes us as perhaps the simplest and most succinct statement of the facts, provided it can be empirically supported on a broader basis.

While (8) represents the classic understanding of Japanese downstep, recent investigations have painted a more complex picture, raising the possibility that not all structure-dependent f0-downtrends (i.e., downtrends not reducible to automatic phonetic declination) can be subsumed under one coherent notion of 'downstep'. On the one hand, work by Selkirk et al. (2003) and Maekawa (2009)[5] has shown that structural f0-downtrends can also affect sequences of all-unaccented phrases, where they are obviously not triggered by an accent and hence cannot be reduced to Poserian downstep, whose locus is the accent itself. On the other hand, structural downtrends have been observed in domains larger than the MaP. Some of the evidence (see especially Ishihara, 2011) concerns cases where focused material is lowered by a preceding accent, calling into question Pierrehumbert and Beckman (1988)'s claim (against Poser, 1984) that left focus edges in Japanese always correspond to left MaP-edges. Ishihara (2011) has shown that this kind of evidence can be made compatible with the classical understanding of downstep by dissociating focus from phrasing and modeling its characteristic prosody by means of other phonological mechanisms. Other ostensibly non-MaP-bounded downtrends, such as the examples in Kubozono (2007), where *wh*-phrases involve bona fide MaP-boundaries, also call for a reappraisal of the theoretical approach.

This paper is not the place to delve into this complex area of questions. We will therefore remain within the limits of classical downstep and assume, for present purposes, that it is a coherent and unified concept and that the new evidence involves independent mechanisms interacting with it. The question we want to address is the following: What does the obvious difference between the domains of downstep and initial rise depicted in (8) actually reveal about prosodic constituency? Does it conclusively show that there are two different kinds of phrases, MaP (with downstep) and MiP (with initial rise)? Departing from the virtual consensus of past work on the prosodic structure of Japanese, we think not. As shown in detail in Ito and Mester (2009b), while an upper bound corresponding to MaP in (8) needs to be imposed on downstep because of the observed pitch reset (setting aside the complicating factors mentioned in the previous paragraph), there is no reason to impose a lower bound and to exclude it from MiP. The reason is that downstep cannot happen within MiP for intrinsic reasons. Downstep requires either a sequence AA (accented word + accented word) or a sequence AU (accented word + unaccented word). But MiP contains maximally one A, so the first cannot occur; and while a U can be phrased together with a following A into a single MiP (UA), it cannot be phrased together with a preceding A—instead of *(AU), we find (A)(U) (see Kubozono, 1988:150–154 and (11)–(13) below). The upshot is that the structural description of downstep cannot be met within the MiP domain. The domain restriction (6i) limiting downstep to MaPs is therefore superfluous. As for the initial rise (%L boundary tone followed by phrasal H), because MaP-initiality usually entails MiP-initiality, it is also found MaP-initially (there is an additional argument from cumulativity since MaP-initial rises are significantly larger than MiP-initial rises). Again, this means that there is no need for the domain restriction (6ii) limiting the initial rise to MiP.

Given the lack of domain arguments from the empirical phonetic side, is there theoretical grounding for the distinction between MiPs and MaPs? Because of the Strict Layering Hypothesis, where repetition of levels was a priori banned in prosodic representations, the question was not seriously raised whether it would be problematic to consider MaP and MiP to be recursive φ's, as in (9).

(9) One-φ model



[4] ''Catathesis'' is a term of art used by Poser, Pierrehumbert and Beckman, and others in order to remain neutral on the question whether or not the downtrend observed in Japanese is of the same kind as the well-established downtrend labeled ''downstep'' found in many African tone languages. Subsequent research again reverted to the more familiar term ''downstep'' for Japanese, without necessarily implying such an identification.

[5] Maekawa's evidence is also noteworthy because it is based on a corpus of recorded actual utterances and not on examples hand-crafted by the experimenter for an artificial lab setting.

A moment's reflection shows that initial rise applies to all φ-phrases, and downstep applies to all φ-phrases, albeit vacuously on some. Ito and Mester (2009b) point out further empirical and theoretical advantages for reducing the distinction between MaP and MiP to a distinction between recursive subcategories of φ.

One defining property of the Minor Phrase not shared with the Major Phrase is accent culminativity: A Minor Phrase can have at most one H*L accent (hence the alternative descriptive name ''Accentual Phrase''). This also does not warrant positing distinct categories. Given the distinction between minimal and maximal subcategories introduced in (5), accent culminativity is simply a property of the minimal φ (a φ that does not dominate any other φ). Minimal and maximal subcategories are even sufficient to deal with the rich descriptive typology of Japanese complex compounds (Ito and Mester, 2007), and the system is flexible enough to avoid further problematic proliferation of categories like the ''Superordinate Minor Phrase'' (Shinya et al., 2004) between Minor Phrase and Major Phrase necessitated for certain syntactic structures (Ito and Mester, 2010).

4. Maintaining the SPMH

Our proposal thus amounts to a prosodic hierarchy with clear correspondence between syntactic and prosodic categories. Concretely, we adopt a version of Selkirk's Match Theory (Selkirk, 2009, 2011a) where the basic principle mapping syntactic structure to prosodic structure is very simple, as in (10).

(10) Match Theory (cf. Selkirk, 2011a)

Syntactic category		**Matching phonological category**
Clause (CP)	→	PClause[6] (usually called ''intonational phrase'', ι)
Phrase (XP)	→	PPhrase (=''Phonological Phrase'', φ)
Word (X: N, V, A)	→	PWord (=''Phonological Word'', ω)

In this view, the Minor Phrase is not its own phonological category, a welcome result for the SMPH, since, as noted by Selkirk and Tateishi (1988:332) and elaborated in the introduction above, the Minor Phrase has no syntactic correspondent like the other categories. In our conception, and that of Selkirk (2011a), the Minor Phrase emerges in certain cases as a recursive subcategory (i) because of syntactic embedding, and (ii) because phonological constraints require a more articulated structure. The goal of this section is to provide an explicit optimality-theoretic analysis of syntax–prosody mapping in Japanese that successfully derives appropriate phonological phrasings, which in turn predict the relevant domain-sensitive phonetic phenomena (initial rise, patterns of pitch scaling, etc.).

We illustrate with some examples from Vance (2008:181). Four instances of the syntactic structure [[NP-Poss]N] are given in (11) which differ in terms of the accentedness of the two nouns (we group the final conjunction *–to* with the head noun here, for simplicity; unless otherwise noted, we follow the standard convention with square brackets [. . .] indicating syntactic constituency and parentheses (. . .) indicating prosodic φ-constituency).

(11) A = accented ω, U = unaccented ω

a. [[U]U] [[hiroshima no¯] sakana to¯]
Hiroshima fish and

b. [[U]A] [[hiroshima no¯] tama ˥go to]
Hiroshima eggs and

c. [[A]A] [[oka˥yama no] tama˥ go to]
Okayama eggs and

d. [[A]U] [[oka˥yama no] sakana to¯]
Okayama fish and

It is always possible, in a diligent pronunciation, to parse each word as a separate φ, with its own initial rise. But in usual speech, this is not what happens. Rather, unaccented words are often phrased together with adjacent (accented or unaccented) words. In this context, Kubozono (1988:150–154) has demonstrated, for a large variety of examples like (11), a pronounced directional asymmetry. While U is typically phrased together with a following A (or U), it is not phrased together with a preceding A. Rather, each word is its own φ. So there is a consistent initial rise at the beginning of the second word in (A)(U), but not in (UA). Kubozono's data are reproduced in (12).

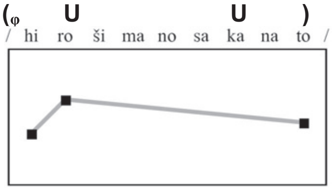
[6] We find the term ''PClause'', as part of a system of phonetically neutral structural terms, more appropriate in a conception that maintains the strong SPMH than a term like ''intonational phrase'', which is carried over from work on languages with a clear intonational structure. The fact that Japanese might not have an intonational system as elaborate as that of English does not mean that it lacks phonological constituent structure above the PPhrase, see Kawahara (2012) for discussion.

(12)

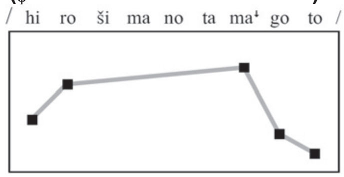
accent profile	% initial rise in second word	N=
UU	39%	44
UA	45%	55
AA	100%	121
AU	100%	77

The schematic phonetic pitch tracks in (13) (from Vance, 2008:181–182) illustrate the basic difference.

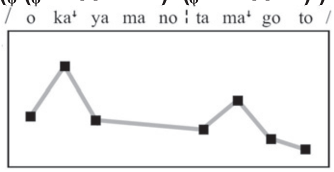
(13) a. [[hiroshima-no¯] sakana-to¯]
($_{\varphi}$ U U)



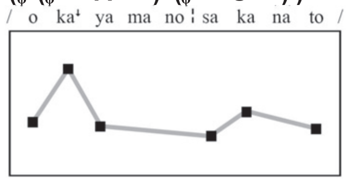
b. [[hiroshima-no¯] tama˺go-to]
($_{\varphi}$ U A)



c. [[oka˺yama-no] tama˺go-to]
($_{\varphi}$ ($_{\varphi}$ A) ($_{\varphi}$ A))



d. [[oka˺yama-no] sakana-to¯]
($_{\varphi}$ ($_{\varphi}$ A) ($_{\varphi}$ U))



When the first word is unaccented (13ab), there is often no initial rise in the second word, whether it is itself accented (13b) or not (13a). On the other hand, when the first word is accented (13cd), the second word shows an initial rise, whether accented (13c) or not (13d). Although the syntactic structure is the same for all four collocations, namely, [[NP-Poss]N], the difference in their accentual profiles results in distinct prosodic structures: monophrasal for U-initial collocations, ($_{\varphi}$UU) and ($_{\varphi}$UA), and biphrasal and adjoined for A-initial collocations, ($_{\varphi}$($_{\varphi}$A)($_{\varphi}$A)) and ($_{\varphi}$($_{\varphi}$A)($_{\varphi}$U)). As illustrated in (14), all initial φ-boundaries coincide with a rising %LH-boundary tone, all accented words are accompanied by an accentual fall H*L, and downstep (indicated by italicization) is seen on φ's following another φ within a larger φ-domain.

(14) Syntax–prosody mapping Tonal assignment/implementation/interpolation

	Syntax–prosody mapping	Tonal assignment/implementation/interpolation
a.	[[U]U] → ($_{\varphi}$UU)	→ ($_{\varphi\%LH}$ U U)
b.	[[U]A] → ($_{\varphi}$UA)	→ ($_{\varphi\%LH}$ U A^{H*L})
c.	[[A]A] → ($_{\varphi}$($_{\varphi}$A)($_{\varphi}$A))	→ ($_{\varphi}$ ($_{\varphi\%LH}$ A^{H*L}) ($_{\varphi\%LH}$ A^{H*L}))
d.	[[A]U] → ($_{\varphi}$($_{\varphi}$A)($_{\varphi}$U))	→ ($_{\varphi}$ ($_{\varphi\%LH}$ A^{H*L}) ($_{\varphi\%LH}$ *U*))

Unless warranted, we will henceforth omit the tonal details from the phonological representations, since they can generally be deduced from the relevant prosodic positions (e.g., initial rise at initial φ-boundaries, accentual fall, and downstep).

Previous accounts of minor phrasing in Japanese have appealed to various phonological conditions (on accentedness, branchingness, peripherality, etc.), and although not defined identically, these are broadly equivalent to well-known constraints found in many prosodic systems cross-linguistically and clearly play an important role in determining prosodic phrasing.

The overall ranked constraint system (to be argued for below) is schematically summarized in the Hasse diagram in (15).

(15)
ACCENT-AS-HEAD, LAPSE(φ) } prosodic constraints
|
MINIMAL BINARITY(φ) }
|
MATCH-XP-TO-φ syntax-prosody interface constraint
|
NORECURSION prosodic constraint

A notable characteristic of this constraint system is that the syntax–prosody interface constraint (MATCH-XP-TO-φ) is sandwiched between different prosodic constraints, so that some prosodic constraints (ACCENT-AS-HEAD, LAPSE(φ), MINIMALBINARITY(φ)) are ranked higher than the interface constraint, which, in turn, is ranked higher than the prosodic constraint on recursive structures. This configuration of constraints allows for an interesting set of predictions. We take these up in turn.

The interface constraint MATCH-XP-TO-φ (16) requires that each syntactic constituent XP correspond to a phonological phrase φ. More precisely, we adopt here the formalization of the MATCH constraint developed in Elfner (2012).

(16) MATCH-XP-TO-φ: Suppose there is a syntactic phrase XP in the syntactic representation that exhaustively dominates a set of one or more terminal nodes α. Assign one violation mark if there is no phonological phrase φ in the phonological representation that exhaustively dominates all and only the phonological exponents of the terminal nodes in α (after Elfner, 2012:28; cf. also the earlier version of the constraint in Selkirk, 2011a).

A syntactic node α exhaustively dominates a set of terminal nodes β if and only if α dominates all and only the terminal nodes in β. Departing from earlier conceptions (such as Truckenbrodt, 1999) where functional categories are invisible for the syntax–prosody mapping, (16) requires matching both for lexical and functional categories. It is rather phonologically null elements that do not count for purposes of prosodic matching (since the constraint is defined in terms of sets of terminal nodes which have phonological exponents). A syntactic configuration $[_{YP}[_{XP}\, a]]$, where the category YP dominates no separate phonologically overt material (i.e., material not dominated by XP), is fully matched by a single phonological phrase $(_{\varphi}\, a)$. On the other hand, for a recursive syntactic structure $[_{XP}[_{XP}\ldots]\ldots]$, where both instantiations of XP dominate overt material, MATCH-XP-TO-φ requires a matching recursive prosodic structure $(_{\varphi}(_{\varphi}\ldots)\ldots)$, as indicated by arrows in (17).

(17)
$$\begin{array}{cccc} [_{XP} & [_{XP}\ldots] & \ldots &] \\ \downarrow & \downarrow\quad\downarrow & & \downarrow \\ (_{\varphi} & (_{\varphi}\ldots) & \ldots &) \end{array}$$

The prosodic representation in (17), however, contains a violation of the NORECURSION constraint (18), which disallows (prosodic) level repetition.

(18) NORECURSION No recursive structures. Assign one violation for each node of category α immediately dominated by another node of category α.

With an input recursive syntactic structure like (17), there are no prosodic representations that can fulfill both constraints. The constraint system in (15) is such that the interface constraint is more highly ranked than the phonological constraint on recursive structure, so the recursive structure is tolerated and emerges as the winning candidate in (19a) (indicated by ▶).

(19)	$[_{XP}[_{XP}\, a\ ]\ \ b\ ]$	MATCH-XP-to-φ	NORECURSION
a. ▶	$(_{\varphi}\ (_{\varphi}\, a\ )\ \ b\ )$		*
b.	$(_{\varphi}\ \ a\ )(_{\varphi} b\ )$	*!	

Candidate (19a) fulfills MATCH-XP-TO-φ but violates NORECURSION, while candidate (19b) has the opposite violation profile (the higher XP violates MATCH-XP-TO-φ). NORECURSION still has a role to play in the grammar, because, all else being equal, the candidate with the fewest violations of NORECURSION is the chosen surface candidate. For example, comparing candidates that all fulfill the higher-ranking MATCH-XP-TO-φ in (20), the evaluation is left to the NORECURSION constraint, which chooses the candidate with the least violations of NORECURSION.[7]

(20)	$[_{XP}\ [_{XP}\, a\ ]\quad b\ ]$	MATCH-XP-TO-φ	NORECURSION
a. ▶	$(_{\varphi}\ \ (_{\varphi}\ a\ )\quad b\ )$		*
b.	$(_{\varphi}\ \ (_{\varphi}\ a\ )\ \ (_{\varphi} b)\ )$		**!
c.	$(_{\varphi}\, (_{\varphi}\ (_{\varphi} a\ ))\ (_{\varphi} b)\ )$		***!

[7] Although there are φ-constituents that are not XPs, MATCH-XP-TO-φ is still fulfilled in (19b) and (19c) because the interface constraint MATCH-XP-TO-φ is unidirectional, from the syntactic structure to the prosodic structure, and requires only that for every XP there is a corresponding φ.

A reviewer asks whether the recursive model may be too unconstrained if there is no limit on the legitimate number/ degree of recursions. The NoRecursion constraint, even if low-ranking, will always militate against superfluous recursive structures not grounded in syntactic structure.[8]

It is not the case, however, that the optimal prosodic structure is always a simple reflection of syntactic structure. Prosodic constituents not due to the existence of a particular syntactic category arise because certain prosodic markedness constraints, as shown in the overall constraint hierarchy in (15), override the matching constraint requiring identity between syntactic and prosodic constituent structure. This is a key theme of the prosodic theory espoused in Selkirk (2011a), Selkirk and Elordieta (2010), Elfner (2010), and Ito and Mester (2010), among others, and what follows can be seen as complementing and strengthening this theme by an explicit proposal concerning the interaction between the constraints governing Japanese syntax–prosody mapping.

As shown in (13)–(14) above, Japanese XPs of the type [[a] b] are prosodically realized as either the flat single φ-structure (a b) or the more articulated recursive φ-structure ((a)(b)), depending on the accent profile of the items involved. Neither are syntactically matched prosodic structures: the former lacks a matching prosodic category (a), and the latter has an extra (b) with no matching syntactic category.

Following up on a proposal in Selkirk and Elordieta (2010), we hypothesize that the flat single φ-structure arises because ($_{\varphi}$a b) fulfills the structural binarity constraint requiring φ to dominate two phonological words. Such binarity constraints are arguably universal: Most well-known are binarity requirements for foot structure (FtBin, Prince, 1980), others hold for words (WdBin, Ito and Mester, 1992) and for phrases (φ-Bin, Kubozono, 1988; Selkirk, 2000). Binarity constraints usually come in separate versions for maximal and minimal binarity (see Mester, 1994; Selkirk, 2000, and Ito and Mester, 2007). The constraint involved here is the minimal version.

(21) MinimalBinarity(φ): φ is minimally binary.
Assign one violation for each φ that does not dominate two ω.

The relevant tableaus are given in (22) and (23). Initial edges of φ will always carry a %LH, and accented words are accompanied by an accentual fall H*L (see (14) above).

(22) Accent profile: UU [[hiroshima no⁻] sakana to⁻] = (13a)

	[[U] U]	MinBin(φ)	Match-XP-to-φ
a. ▶	(U U)		*
b.	((U) U)	*!	
c.	((U) (U))	*!*	

(23) Accent profile: UA [[hiroshima no⁻] tama˺go to] = (13b)

	[[U] A]	MinBin(φ)	Match-XP-to-φ
a. ▶	(U A)		*
b.	((U) A)	*!	
c.	((U) (A))	*!*	

With MinBin ranked higher than Match-XP-to-φ (henceforth, referred to as Match-XP), the a-candidate is correctly chosen as optimal for both UU-profiles (22) and UA-profiles (23).[9] The distinction between the winning a-candidate and

[8] Selkirk has proposed a prosody-to-syntax mapping constraint, Match-φ-to-XP, which penalizes every φ without a correspondent XP. Such a constraint does much of the same work as NoRecursion, because it too penalizes unmotivated prosodic constituents. We here pursue the possibility of using only the interface constraints mapping from syntax to prosody, and not the mapping from prosody to syntax. Further investigation is necessary to evaluate the ramifications of either approach.

[9] We have followed Vance (2008) in assuming a traditional noun phrase structure, which simplifies the analysis. Little of substance seems to change, however, when we adopt a DP structure such as [$_{DP}$[$_{DP}$[$_{NP}$ *hiroshima–no*]]][$_{D'}$[$_{NP}$ *sakana*] $_{D}$[Ø]]], whose D-head remains unfortunately empty (see Saito and Murasugi, 1990; Saito et al., 2008, and Furuya, 2009, among others, on Japanese DP structure; there is disagreement about the position of *-no*, which is arguably not a D-head like *-s* in English). The main difference lies in the larger number of one-word XPs whose parsing as separate φ-phrases is ruled out by MinBin, just as in (21) and (22). We will therefore continue to work with simplified noun phrase structures, noting that the analysis is likely to also be compatible, mutatis mutandis, with a more elaborate syntax. At the other end of the spectrum, a parse in terms of Bare Phrase Structure (Chomsky, 1995; Collins, 2002) would result in structures similar to the ones we are positing (but without category labels). Thanks to John Whitman and Tomo Yoshida for illuminating discussion of these syntactic issues, who should however not be held responsible for the shortcomings of our exposition.

the losing b-candidate is a formal one since there is no empirical criterion diagnosing right φ-edges in Japanese. The losing c-candidate, on the other hand, with two MinBin violations, predicts an initial rise on both words, which is not the tonal contour found for UU and UA combinations.

However, exactly this c-type candidate, with initial rises on both words and hence with two internal φ's (($_{\%LH}$. . .)($_{\%LH}$. . .)), is optimal for the other two accentual profiles, AA and AU: The overall contrast is between (UX) and ((A)(X)). This does not mean that for these examples the constraint rankings are different, in fact, reversing MinBin and Match-XP would make the a-candidate the winner, which is of no help, since it also has only one initial rise. The ranking MinBin over Match-XP is fixed, but there are constraints dominating MinBin that make the c-candidate win when the input has an AX-collocation. The first such constraint concerns accent culminativity, which is violated by the ungrammatical monophrasal parse *(AA). Taking up an idea in Ito and Mester (2007), we subsume accent culminativity under a constraint demanding that an accent is a *head feature*: It must function as the tonal head of some phonological phrase φ (24).

(24) Accent-as-Head: Every accent is the head of a $\varphi^{[+min]}$.
Assign one violation for each accent that is not the head of a minimal phonological phrase φ.

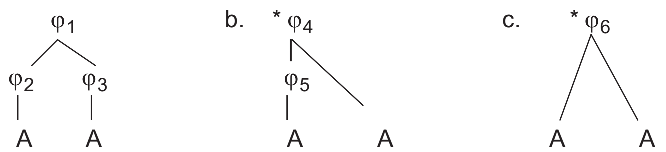
For a minimal (and hence non-recursive) phrase φ, this immediately ensures that it has at most one accent (i.e., no accent or one accent): A second accent, as in a parse like *(AA) in (25c), would not be a head, which is unique by definition.

(25) Accent profile: AA [[oka˺yama no] tama˺go to] = (13c)

	[[A] A]	Accent-as-Hd	MinBin	Match-XP
a. ►	((A) (A))		**	
b.	((A) A)	*!	*	
c.	(A A)	*!		*

The (a–c) candidates in (25) are illustrated in tree form below. The minimal φ's (i.e., φ's not dominating other φ's) are φ_2, φ_3, φ_5, and φ_6. φ_1 and φ_4 are non-minimal since they dominate other φ's.

(26) a. b. c.



Accent-as-Hd is violated in (26bc), because φ_4 dominating A is nonminimal and φ_6 hosts two accents, but (26a) fulfills Accent-as-Hd because both A's are immediately dominated by minimal φ's.

This one-accent limitation is inherited by every category dominated by φ (ω, Ft, σ, and μ).[10] In a recursive (and hence non-minimal) phrase φ, (24) restricts the distribution of accents very severely, ruling out the italicized A's in structures like *(*A*(U)), *((U)*A*), *(*A*(A)), and *((A)*A*)—in a more explicit notation, *($_{\varphi[-\min]}$*A*($_{\varphi[+\min]}$U)), etc.— since they are not dominated by $\varphi^{[+min]}$. ((A)(A)), on the other hand, is possible since each accent heads its own $\varphi^{[+min]}$, and is the parse assigned to [[A]A], as shown in (25a), where the fully isomorphic (b)-candidate is ruled out by (24).

The second relevant constraint is Lapse(φ), which militates against tonal lapses: stretches of low-toned material after the accent.

(27) Lapse(φ): No accentual lapse.
Assign one violation for each fully L-toned ω in φ.

Ranked above MinBin, Lapse(φ) determines the outcome in AU-collocations with a separately phrased U, as shown in (28).

10 It is important to note, in answer to a reviewer's comment, that the requirement is not bidirectional: An accent must be the head of a $\varphi^{[+min]}$, but a $\varphi^{[+min]}$ need not be headed by an accent. Since heads (in Linguistics) are elements that project some of their essential properties (such as category membership, gender, etc.) onto the whole constituent, it may be possible to understand the Accent-as-Hd constraint as simply requiring a separate φ, and not specifically a minimal φ, to host each accent. It would then be up to general conditions on head projection and head prominence to ensure that only a minimal φ qualifies as an accentual host.

(28) Accent profile: AU [[oka˺yama no] sakana to¯] = (13d)

	[[A] U]	LAPSE(φ)	MINBIN(φ)	MATCH-XP
a. ▶	((A)(U))		**	
b.	((A) U)	*!	*	
c.	(A U)	*!		*

The tonal profiles of the candidates in (28) are depicted in more detail in (29), with ω-structure indicated by angled brackets < >. LAPSE(φ) violations arise when unaccented words follow an accentual fall without an intervening left φ-boundary, as in (28bc)/(29bc), where the U-words are fully low-toned. In (28a)/(29a), on the other hand, the U-word, constituting its own φ-phrase, is high after the initial rise.

(29) = tonal profiles of candidates in (28)

a. $(_{\varphi}(_{\varphi}$ %LH $<_{\omega}$ A^{H*L} >) $((_{\varphi}$ %LH $<_{\omega}$ U >) H-toned word U (→No LAPSE(φ) violation)

b. $(_{\varphi}(_{\varphi}$ %LH $<_{\omega}$ A^{H*L} >) $<_{\omega}$ U >) L-toned word U (→LAPSE(φ) violation)

c. $(_{\varphi}$ %LH $<_{\omega}$ A^{H*L} > $<_{\omega}$ U >) L-toned word U (→LAPSE(φ) violation)

It is important that LAPSE(φ) is a constraint *not* against unaccented material, but against fully L-toned words. U-words that make up their own φ, as in (28a)/(29a), or precede another word in φ, as in (30), are H-toned after the initial rise.

(30)
a. $(_{\varphi}$ %LH $<_{\omega}$ U > $<_{\omega}$ A^{H*L} >) H-toned U (→No LAPSE(φ) violation)

b. $(_{\varphi}$ %LH $<_{\omega}$ U > $<_{\omega}$ U >) H-toned U's (→No LAPSE(φ) violation)

LAPSE(φ) is part of a family of related constraints that also includes the much more stringent law governing the pitch accent of Ancient Greek, according to which the ''contonation'' [HL] (i.e., the unit consisting of the accentual Rise + Fall) may be followed by no more than one low-toned vowel mora at the end of the word. In other words, the low pitch L% at the end of the word cannot occupy more than one mora, ''there must not be any of the word left after the low pitch has been reached'' (Probert, 2006:109). This is illustrated in (31), where the ungrammatical forms all have L% associated with more than one mora (bolded).

(31)
a. [H L] L%
μ μμ μ
antʰrɔɔpos

b. [H L]L%
μ μμ μμ
antʰrɔɔpuu

*[H L] L%
μ μμ **μμ**
antʰrɔɔpuu

* [HL] **L%**
μ μμ **μμ**
antʰrɔɔpuu

'human being',
a. nom.
b. gen.sg.

c. [HL] L%
μμ μ
dɔɔ ron

d. [H L] L%
μμ μμ
dɔɔ rois

* [HL] **L%**
μμ **μμ**
dɔɔ rois

'gift',
c. nom.sg.
d. dat.pl.

e. [H L] L%
μμ μ μ
ɔɔpʰelon

*[HL] **L%**
μμ **μ μ**
ɔɔpʰelon

e. 'owe', aorist

This is the Misteli–Allen Law of Limitation (Misteli, 1868; Allen, 1966, 1973, 1987, see Probert, 2003, 2006 for recent discussion) that defines the window at the end of the word which must contain the accent. Certain word classes, including finite verbs, have recessive accent, meaning that the accent must recede as far to the left as possible within this window (see Wackernagel, 1877 for an explanation of recession as optimization).

For completeness, we give some more candidates for the Japanese examples to show how they work under this analysis. With all the relevant constraints in place, our analysis appears in (32)–(35), where the bolded box highlights the crucial interactions determining the winning prosodic candidate.[11]

(32) UU: [[hiroshima no^{-}] sakana to^{-}] = (13a)[12]

	[[U] U]	LAPSE(φ)	ACCENT -AS-HD	MINBIN(φ)	MATCH-XP	NORECURSION
a.	((U) U)			*!		*
b. ▶	(U U)				*	
c.	((U) (U))			*!*		**
d.	(U) (U)			*!*	*	
e.	(U (U))			*!	*	*

(33) UA: [[hiroshima no^{-}] tama⌝go to] = (13b)

	[[U] A]	LAPSE(φ)	ACCENT-AS-HD	MINBIN (φ)	MATCH-XP	NORECURSION
a.	((U) A)		*!	*		*
b. ▶	(U A)				*	
c.	((U) (A))			*!*		**
d.	(U) (A)			*!*	*	
e.	(U (A))			*!	*	*

(34) AA: [[oka⌝yama no] tama⌝go to] = (13c)

	[[A] A]	LAPSE (φ)	ACCENT-AS-HD	MINBIN(φ)	MATCH-XP	NORECURSION
a.	((A) A)		*!	*		*
b.	(A A)		*!		*	
c. ▶	((A) (A))			**		**
d.	(A) (A)			**	*!	
e.	(A (A))		*!	*	*	*

(35) AU: [[oka⌝yama no] sakana to^{-}] = (13d)

	[[A] U]	LAPSE(φ)	ACCENT-AS-HD	MINBIN(φ)	MATCH-XP	NORECURSION
a.	((A) U)	*!		*		*
b.	(A U)	*!			*	
c. ▶	((A) (U))			**		**
d.	(A) (U)			**	*!	
e.	(A (U))		*!	*	*	

The type of prosodic phrasing under discussion does not just hold for noun phrases of a specific structure. Other syntactic combinations with the same behavior are given in Vance, 2008:183, confirming the general XP mapping (irrespective of the syntactic category of X).

(36)	a.		[[Adj] N particle]	
		((A) (A))	((taka⌝i) (omo⌝ča mo))	'expensive toy too'
		((A) (U))	((taka⌝i) (kuruma mo^{-}))	'expensive car too'
		(U A)	(omoi^{-} omo⌝ča mo)	'heavy toy too'
		(U U)	(omoi^{-} kuruma mo^{-})	'heavy car too'

[11] Selkirk (2009, 2011a) proposes that the lowest-ranked NORECURSION constraint might not exist, and this is indeed a possibility. We retain it here mainly to indicate the minimal degree of recursion that the new conception of prosodic parsing predicts. Another consideration is that flatter modes of parsing without recursion are often alternatively possible (for example, for the phrases in (10)), and a higher ranking of NORECURSION is one of the conceivable ways of accounting for them.

[12] The constraint ACCENT-AS-HD is not applicable to unaccented phrases, since there is no accent that needs to head a $\varphi^{[+min]}$.

b.		[[Adj Suffix] NEG]	
		[[Adj /ku/] /na⌝i/]	
	((A) (A))	((hi⌝roku) (na⌝i))[13]	'not wide'
	(U A)	(osoku¯ na⌝i)	'not late'
c.		[[N particle] V]	
	((A) (A))	((kuda⌝mono mo) (tabe⌝ru))	'eat fruit too'
	((A) (U))	((kuda⌝mono mo) (sagasu¯))	'look for fruit too'
	(U A)	(kamaboko mo¯ tabe⌝ru)	'eat fish paste too'
	(U U)	(kamaboko mo¯ sagasu¯)	'look for fish paste too'

Japanese names (sequences of family name + first name) are treated in the same way in fluent speech, as illustrated in (37).[14]

(37) a. ((A) (A)) ((Nishi⌝mura) (Yasu⌝aki))
b. ((A) (U)) ((Nishi⌝mura) (Ichiroo¯))
c. (U A) (Uemura¯ Yasu⌝aki)
d. (U U) (Uemura¯ Ichiroo¯)

The crucial ranking relations between the constraints involved in the analysis are summarized in (38).

(38) Accent-as-Head>>MinBin(φ)
Lapse(φ)>>MinBin(φ)
MinBin(φ)>>Match-XP
Match-XP>>NoRecursion
Accent-as-Head>>NoRecursion

Accent-as-Head Lapse(φ)
MinBin(φ)
Match-XP
NoRecursion

The gist of the account is that the strict homomorphism favored by Match-XP is offset by MinBin(φ), which favors less parsing, and by Lapse(φ) and Accent-as-Head, which each favor more parsing; this is why preference-ranked constraints, as provided by OT, are indispensable.

5. Further consequences

Pursuing the formal constraint-based analysis one step further, we take up three-member syntactic combinations, using the accented and unaccented words in (39). This leads to an interesting prediction about their prosodic structure, depending on the nature of the syntactic input (left- vs. right-branching).

(39)

A(ccented) words		U(naccented)-words	
isura⌝eru	'Israel'	amerika¯	'America'
kurasume⌝eto	'classmate'	tomodachi¯	'friend'
rapputo⌝ppu	'laptop'	pasokon¯	'personal computer'
atarashi⌝i	'new'	omoi¯	'heavy'

All possible combinations of accented and unaccented words in three-word collocations are listed below for left-branching syntax (40) and right-branching syntax (41), together with their prosodic structures. We are here dealing only

[13] Another variant is ((*hiro⌝ku*) (*na⌝i*), with penult accent on the adjective (Vance, 2008).
[14] Different from foreign names, which are treated as compounds, with systematic deaccentuation of their first part (see Zamma, 2001 for discussion).

with the basic patterns for these prosodic structures established in the work of Kubozono (1989), not with possible other variants found in other studies, in particular, Selkirk and Tateishi (1988).

(40)	Left-branching syntax: [[[α] β] γ]		*Prosodic structure*
a.	[[[U]U]U]:	[[[amerika-no¯] tomodachi-no¯] pasokon¯] America -$_{\text{GEN}}$ friend -$_{\text{GEN}}$ PC '(my) American friend's PC'	((*UU*) *U*)
b.	[[[U]A]A]:	[[[amerika-no¯] kurasume⌝eto-no] rapputo⌝ppu]	((*UA*) (*A*))
c.	[[[U]A]U]:	[[[amerika-no¯] kurasume⌝eto-no] pasokon¯]	((*UA*) (*U*))
d.	[[[U]U]A]:	[[[amerika-no¯] tomodachi-no¯] rapputo⌝ppu]	((*UU*) (*A*))
e.	[[[A]A]A]:	[[[isura⌝eru-no] kurasume⌝eto-no] rapputo⌝ppu]	(((*A*)(*A*)) (*A*))
f.	[[[A]A]U]:	[[[isura⌝eru-no] kurasume⌝eto-no] pasokon¯]	(((*A*)(*A*)) (*U*))
g.	[[[A]U]A]:	[[[isura⌝eru-no] tomodachi-no¯] rapputo⌝ppu]	(((*A*)(*U*)) (*A*))
h.	[[[A]U]U]:	[[[isura⌝eru-no] tomodachi-no¯] pasokon¯]	(((*A*)(*U*)) (*U*))

(41)	Right-branching syntax: [[α] [[β] γ]]			*Prosodic structure*
a.	[[U] [[U]U]]:	[[tomodachi-no¯] [[omoi¯] pasokon¯]] friend-$_{\text{GEN}}$ heavy PC '(my) friend's heavy PC'		(*U* (*U U*))
b.	[[U] [[A]A]]:	[tomodachi-no¯]	[[atarashi⌝i] rapputo⌝ppu]]	(*U* ((*A*)(*A*)))
c.	[[U] [[A]U]]:	[tomodachi-no¯]	[[atarashi⌝i] pasokon¯]]	(*U* ((*A*)(*U*)))
d.	[[U] [[U]A]]:	[tomodachi-no¯]	[[omoi¯] rapputo⌝ppu]]	(*U* (*UA*))
e.	[[A] [[A]A]]:	[kurasume⌝eto-no]	[[atarashi⌝i] rapputo⌝ppu]]	((*A*) ((*A*)(*A*)))
f.	[[A] [[A]U]]:	[kurasume⌝eto-no]	[[atarashi⌝i] pasokon¯]]	((*A*) ((*A*)(*U*)))
g.	[[A] [[U]A]]:	[kurasume⌝eto-no]	[[omoi¯] rapputo⌝ppu]]	((*A*) (*UA*))
h.	[[A] [[U]U]]:	[kurasume⌝eto-no]	[[omoi¯] pasokon¯]]	((*A*) (*UU*))

In what follows, we show how the constraint-based analysis (38) developed in the previous section carries over straightforwardly for these longer sequences, predicting the correct prosodic structures (and concomitantly, the correct tonal contours associated with them). In addition, the close formal investigation of cases involving combinations of accented and unaccented words reveals the need for additional development of the syntax–prosody matching constraint.

We start with all-unaccented three-member collocations in tableaux (42) (=left-branching, (40a)) and (43) (=right-branching, (41a)) with the relevant constraints. Note that there are three relevant syntactic constituents in (40a) [$_1$ [$_2$ [$_3$ *amerika-no* $_3$] *tomodachi-no* $_2$] *pasokon* $_1$], namely, 1 = [*amerika-no tomodachi-no pasokon*], 2 = [*amerika-no tomodachi-no*], and 3 = [*amerika-no*]. Mutatis mutandis, the same holds for the other examples.

The prosody of the first candidate (42a) is fully matched with the left-branching syntax, i.e., each syntactic constituent is replicated by the prosodic constituent, but it violates the high-ranking binarity constraint MinBin(φ), which discourages phrasing each word as a separate phrase. Among the three MinBin(φ)-fulfilling candidates (42c,d,e), candidate (42c), with the least number of Match-XP violations, emerges as the winner.

(42) Left-branching UUU: = (40a)
[[[amerika-no¯] tomodachi-no¯] pasokon¯]
(($_{\%\text{LH}}$ U U) U) → Initial rise only on *amerika*¯

	[[[U] U] U]	MinBin(φ)	Match-XP	NoRecursion
a.	(((U) U) U)	*!		**
b.	(U (U (U)))	*!	**	**
c. ►	((UU) U)		*	*
d.	(U (UU))		**!	*
e.	(UUU)		**!	
f.	((UU) (U))	*!	*	**
g.	((U) (UU))	*!	*	**
h.	(((U)(U)) (U))	*!**		****
i.	((U) ((U)(U)))	*!**	*	****
j.	(U)(U)(U)	*!**	**	

(43) Right-branching UUU: = (41a)
[[tomodachi-no⁻] [[omoi⁻] pasokon⁻]]
(%LH U (%LH U U)) →Initial rise on *tomodachi⁻* and *omoi⁻*, but not on *pasokon⁻*

	[[U] [[U] U]]	MinBin(φ)	Match-XP	NoRecursion
a.	(((U) U) U)	*!	**	**
b.	(U (U (U)))	*!	**	**
c.	((UU) U)		***!	*
d. ▶	(U (UU))		**	*
e.	(UUU)		***!	
f.	((U) (UU))	*!	*	**
g.	((UU) (U))	*!	***	**
h.	(((U)(U)) (U))	*!**	*	****
i.	((U) ((U)(U)))	*!**		****
j.	(U)(U)(U)	*!**	**	

For both left-branching and right-branching unaccented sequences, MinBin(φ) makes the first cut, followed by Match-XP, in determining their respective correct prosodic structures. This in turn leads to the difference in surface phonetic outcome, with a single initial rise for left-branching cases (($_{\%LH}$UU)U) and two initial rises for right-branching cases ($_{\%LH}$U ($_{\%LH}$UU)).

A basic observation here is that satisfying MinBin(φ) and Match-XP requires recursion-based subcategories, in violation of low-ranking NoRecursion (see footnote 11 above for related discussion). Recent work by Maekawa highlights the importance of the exact phrasing assigned to such longer strings of unaccented words. Diverging from the received wisdom of earlier work (such as Poser, 1984 and Pierrehumbert and Beckman, 1988) that viewed downstep as an exclusive property of material following an accent within the same maximal phonological phrase (in our terms), Maekawa (2009) found downstep-like phenomena in unaccented sequences (i.e., structure-dependent downtrends not explicable as automatic downdrift).

Turning now to all-accented sequences, Accent-as-Head militates against grouping any two accented words into a single minimal phrase, overriding any pressure from MinBin(φ), and Match-XP insists on further articulation into recursive subcategories of φ. As illustrated in the two tableaux below, candidates (h, i, j) fulfill Accent-as-Head and have the same number of MinBin(φ) violations, so the Match-XP constraint chooses the prosodically left-branching h-candidate for left-branching syntax (44) and the prosodically right-branching i-candidate for right-branching syntax (45).

(44) Left-branching AAA: = (40e)
[[[isura˥eru-no] kurasume˥eto-no] rapputo˥ppu]
(((%LH A^{H*L}) (%LH $A^{!H*L}$)) (%LH $A^{!H*L}$)) → Downstepped !H*L, initial rise on every word
Israel -$_{GEN}$ classmate -$_{GEN}$ laptop
'(my) Israeli classmate's laptop'

	[[[A] A] A]	Accent-as-Hd	MinBin(φ)	Match-XP	NoRecursion
a.	(((A) A) A)	*!*	*		**
b.	(A (A (A)))	*!*	*	**	**
c.	((AA) A)	*!*		*	*
d.	(A (AA))	*!*		**	*
e.	(AAA)	*!*		**	
f.	((AA) (A))	*!	*		**
g.	((A) (AA))	*!	*	*	**
h. ▶	(((A)(A)) (A))		***		****
i.	((A) ((A)(A)))		***	*!	****
j.	(A)(A)(A)		***	*!*	

(45) Right-branching AAA: = (41e)

[[kurasume⌝eto-no] [[atarashi⌝i] rapputo⌝ppu]] → Initial rise on every word, boost (↑) & downstep(!) on *atarashii*, downstep on *rapputoppu*

$((_{\%LH}\ A^{H*L}\ )\uparrow((_{\%LH}\ A^{!H*L})(_{\%LH}\ A^{!H*L})))$

'(my) classmate's new laptop'

	[[A] [[A] A]]	Accent-as-Hd	MinBin (φ)	Match-XP	NoRecursion
a.	(((A) A) A)	*!*	*	**	**
b.	(A (A (A)))	*!*	*		**
c.	((AA) A)	*!*		*	*
d.	(A (AA))	*!*		**	*
e.	(AAA)	*!*		***	
f.	((AA) (A))	*!	*	*	**
g.	((A) (AA))	*!	*		**
h.	(((A)(A)) (A))		***	*!	****
i. ▶	((A) ((A)(A)))		***		****
j.	(A)(A)(A)		***	*!*	

Besides downstep (!) found in both structures, the right-branching structures receive a metrical boost (↑) (Kubozono, 1988:205–219) at their internal juncture. Comparing the absence of a boost in left-branching (((A)(A))_(A)) with its presence in right-branching ((A)_((A)(A))), we conclude that Kubozono's metrical boost is a pitch range reset localized at the beginning of a recursive constituent ($(_{\varphi}\ (_{\varphi}\ \ldots))$), i.e., at the point where both the smaller constituent and the larger constituent begin.

Our OT constraints and their ranking thus ensure the correct prosodic outcome for all-unaccented and all-accented collocations in both left-branching and right-branching syntax. However, the analysis as it stands is still lacking a crucial ingredient needed to capture the syntax–prosody correspondences[15] for combinations of accented and unaccented words, where both Accent-as-Head and Lapse(φ) play a crucial role. The central issue here is the following. Even though, as we have just seen, Match-XP is active in these parses and often results in syntax–prosody homomorphism, significant instances of non-homomorphism arise because Match-XP is itself dominated by Lapse(φ) and MinBin(φ). More precisely, the left-branching syntactic bracketing for [[[A]U]A] (46) (=(40g)) is wrongly parsed as the prosodically right-branching ((A)(UA)) (46g), instead of the correct prosodically left-branching (((A)(U))(A)) (46h).

(46) [[[A]U]A]: [[[isura⌝eru-no] tomodachi-no⁻] rapputo⌝ppu] = (40g)

	[[[A] U] A]	Accent -as-Hd	Lapse (φ)	MinBin(φ)	Match-XP
a.	(((A) U) A)	*!	**	*	
b.	(A (U (A)))	*!		*	**
c.	((AU) A)	*!	**		*
d.	(A (UA))	*!			**
e.	(AUA)	*!	**		**
f.	((AU) (A))		*!	*	*
g. ▶wrong winner	((A) (UA))			*	*
h. actual output	(((A)(U)) (A))			**!*	
i.	((A) ((U)(A)))			**!*	*
j.	(A)(U)(A)			**!*	**

Research on right- vs. left-branching structures (Kubozono, 1989) has conclusively shown that such structures are always prosodically distinct and never neutralized (here, to avoid a violation of MinBin(φ)). In other words, [[[A]U]A] must be parsed as (46h) (((A)(U))_(A)), with an internal juncture after AU, and not as (46f) ((A)_(UA)), with a juncture after A. The medial U cannot be rebracketed with the following A to form a prosodic constituent. Said in another way, the intermediate, non-minimal syntactic phrase [[A]U] must have a prosodic correspondent in the parse of [[[A]U]A]. The contrasting recursive structures for right- and left-branching syntactic inputs here correspond directly to the recursive MiP-structures argued for by Kubozono (1989:59).

[15] We are very grateful to an anonymous reviewer, whose comments led to significant improvements in this section.

The same issue arises for the syntactic bracketing [[[A]U]U], wrongly parsed as (47g) ((A)(UU)) instead of the more homomorphous parsings (47h) (((A)(U))(U)) because MINBIN(φ)>>MATCH-XP, or (47f) ((AU)(U)) because of LAPSE(φ).

(47) [[[isuraˈeru-no] tomodachi-no⁻] pasokon⁻] =(40h)

	[[[A] U] U]	ACCENT -AS-HD	LAPSE (φ)	MINBIN (φ)	MATCH-XP
a.	(((A) U) U)		*!*	*	
b.	(A (U (U)))	*!		*	**
c.	((AU) U)		*!*		*
d.	(A (UU))	*!			**
e.	(AUU)		*!*		**
f.	((AU) (U))		*!	*	*
g. ▶wrong winner	((A) (UU))			*	*
h. actual output	(((A)(U)) (U))			**!*	
i.	((A) ((U)(U)))			**!*	*
j.	(A)(U)(U)			**!*	**

A first approach to the issue here is the obvious cyclic one: since [[A]U] is parsed as ((A)(U)) in isolation, it will also be parsed in this way in [[[A]U]A] and [[[A]U]U], provided (i) the prosodic structure is computed cyclically and (ii) faithfulness to cyclically established structure outranks the relevant markedness constraints (here, MINBIN(φ)). This would be a straightforward and theoretically attractive solution, were it not for the fact that Selkirk and Tateishi (1988) have argued in detail that the general patterns, and variants, of minor phrasing in Japanese cannot be obtained in a cyclic manner, but need to be derived by wellformedness constraints applying in parallel to the whole structure.

Noting the option of the cyclic solution, we do not pursue it here on the strength of the arguments in Selkirk and Tateishi (1988), and approach the issue in a different manner. The way the problem presented itself already suggests a line of attack, which lies in treating the lowest level of XP-structure prosodically in a different way from higher levels of XP-structure. Let us refer to these levels as $XP^{[+min]}$ (XP not dominating another XP) and $XP^{[-min]}$ (XP dominating another XP), where potential differences in lexical category X are immaterial.[16] The central point is the following: $XP^{[+min]}$ can fail to be parsed as a φ in order to avoid a prosodically non-binary φ. This happens, for example, when [[U]A] is parsed as (UA) because of MINBIN(φ). Here the syntactic phrase [U] has no correspondent, but nothing has been rebracketed. But when [[[A]U]A] is parsed as ((*A*)(*UA*)), something else has happened, namely, a kind of rebracketing: Instead of being grouped with the first A, the medial U is now grouped with the second A. This is obviously a different kind of failure in the syntax-phonology mapping from the simple lack of a prosodic instantiation of an $XP^{[+min]}$ constituent. Several ways of formalizing this are conceivable, we will here choose what looks like the most straightforward one, namely, to postulate a separate MATCH constraint for $XP^{[-min]}$, i.e., for constituents that encode syntactic branchingness involving at least one other XP.[17] The crucial point is that $XP^{[-min]}$ must correspond to a φ, even when this results in a non-binary φ, namely, the φ corresponding to the medial U in (((A)(U))(A)) and (((A)(U))(U)) because of LAPSE(φ).[18] The general point here harks back to the central argument in Kubozono (1989:59), namely, that "that the ups and downs of Fo contours in Japanese are heavily constrained by the syntactic constituency of the utterances, much more so than has previously been thought". In particular, left- vs. right-branchingness is of prime importance. In our analysis, we posit MATCH-$XP^{[-min]}$ as a constraint separate from the general MATCH-XP, and higher-ranking, as in (49).

(48) MATCH-$XP^{[-min]}$-TO-φ: Suppose there is a syntactic phrase $XP^{[-min]}$ in the syntactic representation that exhaustively dominates a set of one or more terminal nodes α. Assign one violation mark if there is no phonological phrase φ in the phonological representation that exhaustively dominates all and only the phonological exponents of the terminal nodes in α (after Elfner, 2012:28; cf. also the earlier version of the constraint in Selkirk, 2011a).

[16] One way of implementing this is derivational, taking the first step in the syntax-phonology mapping to be a neutralization of all syntactic category distinctions, which are assumed to be invisible at this stage (Truckenbrodt, 1999), into a generic X/XP pattern.

[17] An alternative would be Match-$\varphi^{[-min]}$, a prosody-syntax MATCH constraint requiring each φ to correspond to an XP, as in Selkirk, 2011b.

[18] We take the correct structures to be (((A)(U))(A))/(((A)(U))(U)) not ((AU)A)/((AU)U) because the information about the pitch curves of such utterances in Kubozono, 1989 clearly indicate the presence of an initial rise in the medial U, i.e., its beginning must correspond to the beginning of a φ.

(49) Match-XP$^{[\text{-min}]}$ >>MinBin(φ)
Accent-as-Head>>MinBin(φ)
Lapse(φ)>>MinBin(φ)
MinBin(φ)>>Match-XP
Match-XP>>NoRecursion
AccentHead>>NoRecursion

Accent-as-Head Lapse(φ) Match-XP$^{[\text{-min}]}$
MinBin(φ)
Match-XP
NoRecursion

The revised analysis selects the correct winners, as shown in (50) and (51).

(50) [[[isura˺eru-no] tomodachi-no¯] rapputo˺ppu] = (40g)

	[[[A] U] A]	Accent -as-Hd	Lapse (φ)	Match -XP$^{[\text{-min}]}$	MinBin (φ)	Match -XP
a.	(((A) U) A)	*!	**		*	
b.	(A (U(A)))	*!		*	*	**
c.	((AU) A)	*!	**			*
d.	(A (UA))	*!		*	*	**
e.	(AUA)	**!	**	*		**
f.	((AU) (A))		*!		*	*
g.	((A) (UA))			*!	*	*
h. ►correct winner	(((A)(U)) (A))				***	
i.	((A) ((U)(A)))			*!	***	*
j.	(A)(U)(A)			*!	***	**

(51) [[[isura˺eru-no] tomodachi-no¯] pasokon¯] = (40h)

	[[[A] U] U]	Accent -as-Hd	Lapse (φ)	Match -XP$^{[\text{-min}]}$	MinBin (φ)	Match -XP
a.	(((A) U) U)		*!*		*	
b.	(A (U (U)))	*!		*	*	**
c.	((AU) U)		*!*			*
d.	(A (UU))	*!		*		**
e.	(AUU)		*!*	*		**
f.	((AU) (U))		*!		*	*
g.	((A) (UU))			*!	*	*
h. ►correct winner	(((A)(U)) (U))				***	
i.	((A) ((U)(U)))			*!	***	*
j.	(A)(U)(U)			*!	***	**

6. Conclusion

In Prosodic Hierarchy Theory, a large number of distinct prosodic categories has been proposed in order to provide enough separate domains for different processes. While empirically well-founded, this research program has resulted in an embarrassment of riches: These categories have never been instantiated in a single language in their totality, and their crosslinguistic identification has remained a largely unsolved problem. On the theory side, one root of the problem is the doctrine of strict layering. Even within a single language, it has led to a multiplication of categories. Whenever a process is found to operate in a slightly different domain than some other process, the model required setting up two separate categories.

Once repetition of levels (recursion) becomes an option, ‘‘constituent domain’’ no longer equals ‘‘category’’. Some of the categories proposed in the earlier prosodic literature are in reality only larger and smaller subcategories in a recursive structure 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. This is possible because in our proposal, relational notions, in particular, maximal and minimal projections of categories, play a crucial and natural role in this system. The strict syntax–prosody homomorphism suggested by Match-XP is overruled by MinBin(φ), leading to less parsing, and by Lapse(φ) and

ACCENT-AS-HEAD, leading to more parsing. MINBIN(φ) is in turn overruled by MATCH-XP$^{[-\min]}$, which prohibits any rebracketing in order to produce structures conforming better to binarity. For Japanese, one of the best-studied accentual systems, the evidence overall favors a model with a single category '*phonological phrase*', with the option of recursion.

Acknowledgements

Earlier versions of this work have been presented in the NINJAL (National Institute of Japanese Language and Linguistics) colloquium series, and in phonology seminars at UC Santa Cruz and International Christian University. We are grateful to the participants at these venues, who provided valuable comments and helpful advice, in particular, Ryan Bennett, Kendra Buchanan, Shin Ishihara, Haruo Kubozono, Lisa Selkirk, Tim Vance, John Whitman, and Tomo Yoshida. Many thanks also go to the *Lingua* reviewers as well as the editors, Yoshi Kitagawa and Satoshi Tomioka, for their patience and their careful reading and critique of a previous version of this paper, which led to significant improvements in both analysis and theory.

References

Allen, W.S., 1966. A problem of Greek accentuation. In: Bazell, C.E., Catford, J.C., Halliday, M.A.K., Robins, R.H. (Eds.), In Memory of J.R. Firth. Longmans, London, pp. 8–14.

Allen, W.S., 1973. Accent and Rhythm. Cambridge University Press, Cambridge.

Allen, W.S., 1987. Vox Graeca. Cambridge University Press, Cambridge.

Bierwisch, M., 1966. Regeln für die Intonation deutscher Sätze. In: Bierwisch, M. (Ed.), Untersuchungen über Akzent und Intonation im Deutschen. Studia grammatica, vol. 7. Akademie Verlag, Berlin, pp. 99–201.

Chomsky, N., 1995. Bare phrase structure. In: Webelhuth, G. (Ed.), Government and Binding Theory and the Minimalist Program (Generative Syntax 1). Blackwell, Cambridge, pp. 383–439.

Collins, C., 2002. Eliminating labels. In: Epstein, S., Seely, D. (Eds.), Derivation and Explanation in the Minimalist Program. Blackwell, Malden, MA, pp. 42–64.

Elfner, E., 2010. Recursivity in prosodic phrasing: evidence from Conamara Irish. Ms. University of Massachusetts, Amherst. To appear in Kan, Seda, Moore-Cantwell, Claire, Staubs, Robert (Eds.), Proceedings of the 40th Annual Meeting of the North-East Linguistic Society. GLSA Publications, Amherst, MA.

Elfner, E., 2012. Syntax–Prosody Interactions in Irish. Doctoral dissertation, Department of Linguistics, University of Massachusetts, Amherst.

Furuya, K., 2009. The DP Hypothesis through the Lens of Japanese Nominal Collocation Constructions. Doctoral dissertation, Department of Linguistics, City University of New York.

Gussenhoven, C., 1991. The English rhythm rule as an accent deletion rule. Phonology 8, 1–35.

Gussenhoven, C., 2005. Procliticized phonological phrases in English: evidence from rhythm. Studia Linguistica 59, 174–193.

Haider, H., 1993. Deutsche Syntax – generativ. Vorstudien zur Theorie einer projektiven Grammatik. Gunter Narr, Tübingen.

Haraguchi, S., 1977. The Tone Pattern of Japanese: An Autosegmental Theory of Tonology. Kaitakusha, Tokyo.

Hayes, B., 1989. The prosodic hierarchy in meter. In: Kiparsky, P., Youmans, G. (Eds.), Rhythm and Meter. Academic Press, Orlando, pp. 201–260.

Hirst, D., 1993. Detaching intonational phrases from syntactic structure. LI 21, 781–788.

Ishihara, S., 2011. Japanese focus prosody revisited: freeing focus from prosodic phrasing. Lingua 121, 1870–1888.

Ito, J., Mester, A., 1992. Weak layering and word binarity. Linguistic Research Center Working Paper LRC-92-09. UC Santa Cruz. Published in Honma, Takeru, Okazaki, Masao, Tabata, Toshiyuki, Tanaka, Shin-ichi (Eds.), 2003. A New Century of Phonology and Phonological Theory. A Festschrift for Professor Shosuke Haraguchi on the Occasion of His Sixtieth Birthday. Kaitakusha, Tokyo, pp. 26–65.

Ito, J., Mester, A., 2007. Prosodic adjunction in Japanese compounds. In: Miyamoto, Y., Ochi, M. (Eds.), Formal Approaches to Japanese Linguistics: Proceedings of FAJL 4. MIT Working Papers in Linguistics, vol. 55. MIT Department of Linguistics and Philosophy, Cambridge, Massachusetts, pp. 97–111.

Ito, J., Mester, A., 2009a. The extended prosodic word. In: Grijzenhout, J., Kabak, B. (Eds.), Phonological Domains: Universals and Deviations. Mouton de Gruyter, Berlin, pp. 135–194.

Ito, J., Mester, A., 2009b. The onset of the prosodic word. In: Parker, S. (Ed.), Phonological Argumentation: Essays on Evidence and Motivation. Equinox, London, pp. 227–260.

Ito, J., Mester, A., 2010. Recursive prosodic phrasing in Japanese. In: Proceedings of the 18th Japanese/Korean Conference. CSLI, Stanford, CA, pp. 147–164.

Jun, S.-A., 2005. Prosodic typology. In: Jun, S.-A. (Ed.), Prosodic Typology: The Phonology of Intonation and Phrasing. Oxford University Press, Oxford, pp. 430–458.

Kawahara, S., 2012. The intonation of nominal parentheticals in Japanese. In: Borowsky, T., Kawahara, S., Shinya, T., Sugahara, M. (Eds.), Prosody Matters: Essays in Honor of Elisabeth Selkirk. Equinox, London, pp. 304–340.

Kubozono, H., 1988. The Organization of Japanese Prosody. Doctoral dissertation, University of Edinburgh. Kurosio Publishers, Tokyo, 1993.

Kubozono, H., 1989. Syntactic and rhythmic effects on downstep in Japanese. Phonology 6, 39–67.

Kubozono, H., 2007. Focus and intonation in Japanese: does focus trigger pitch reset? In: Ishihara, S. (Ed.), Proceedings of the 2nd Workshop on Prosody, Syntax, and Information Structure (WPSI 2). Interdisciplinary Studies on Information Structure, vol. 9. Universitätsverlag Potsdam, Potsdam, pp. 1–27.

Ladd, D.R., 1986. Intonational phrasing: the case for recursive prosodic structure. Phonology 3, 311–340.

Ladd, D.R., 1988. Declination ''reset'' and the hierarchical organization of utterances. JASA 84, 530–544.

Maekawa, K., 2009. Contributions to Corpus Phonetics, Part 2: Validation of F0 model, vol. 22. National Institute for Japanese Language and Linguistics, Tachikawa, Japan, Handout of SALON presentation, December 22, 2009.
Martin, S.E., 1952. Morphophonemics of Standard Colloquial Japanese. Linguistic Society of America, Baltimore.
McCarthy, J.J., Prince, A.S., 1986. Prosodic Morphology. Ms. University of Massachusetts at Amherst and Brandeis University. Reissued in 1996 as ''Prosodic Morphology 1986.'' Technical Report #32, Rutgers University Center for Cognitive Science.
McCawley, J.D., 1968. The Phonological Component of a Grammar of Japanese. Mouton, The Hague, The Netherlands.
Mester, A., 1994. The quantitative trochee in Latin. NLLT 12, 1–61.
Misteli, F., 1868. Über die Accentuation des Griechischen. Zeitschrift für vergleichende Sprachforschung 17, 81–134, 161–194.
Nagahara, H., 1994. Phonological Phrasing in Japanese. Doctoral dissertation, Department of Linguistics, UCLA.
Nespor, M., Vogel, I., 1983. Prosodic structure above the word. In: Cutler, A., Ladd, D.R. (Eds.), Prosody: Models and Measurements. Springer, Berlin/Heidelberg/New York/Tokyo, pp. 123–140.
Nespor, M., Vogel, I., 1986. Prosodic Phonology. Foris, Dordrecht.
Pierrehumbert, J., Beckman, M., 1988. Japanese Tone Structure. MIT Press, Cambridge, MA.
Poser, W.J., 1984. The Phonetics and Phonology of Tone and Intonation in Japanese. Doctoral dissertation, MIT.
Prince, A.S., 1980. A metrical theory for Estonian quantity. LI 11, 511–562.
Prince, A.S., Smolensky, P. 1993 (2004). Optimality Theory: Constraint Interaction in Generative Grammar. RuCCS-TR-2. Ms. Rutgers University and University of Colorado, Boulder, Brunswick, New Jersey, and Boulder, Colorado [Published 2004, Blackwell, Malden, Mass.].
Probert, P., 2003. A New Short Guide to the Accentuation of Ancient Greek. Bristol Classical Press, London.
Probert, P., 2006. Ancient Greek Accentuation. Synchronic Patterns, Frequency Effects, and Prehistory. Oxford University Press, Oxford.
Saito, M., Murasugi, K., 1990. N'-deletion in Japanese: a preliminary study, Japanese/Korean Linguistics, vol. 1. CSLI Publications, Stanford, pp. 258–301.
Saito, M., Lin, T.-H.J., Murasugi, K., 2008. N'-ellipsis and the structure of noun phrases in Chinese and Japanese. Journal of East Asian Linguistics 17, 247–271.
Selkirk, E., 1978. On prosodic structure and its relation to syntactic structure. In: Fretheim, T. (Ed.), Nordic Prosody. TAPIR, Trondheim, pp. 111–140.
Selkirk, E., 1980. Prosodic domains in phonology: Sanskrit revisited. In: Aronoff, M., Kean, M.M.-L. (Eds.), Juncture. Anma Libri, Saratoga, CA, pp. 107–129.
Selkirk, E., 1981. On the nature of phonological representation. In: Anderson, J., Laver, J., Meyers, T. (Eds.), The Cognitive Representation of Speech. North Holland, Amsterdam, pp. 379–388.
Selkirk, E., 1984. Phonology and Syntax: The Relation between Sound and Structure. MIT Press, Cambridge, MA.
Selkirk, E., 1996. The prosodic structure of function words. In: Morgan, J.L., Demuth, K. (Eds.), Signal to Syntax. Lawrence Erlbaum, Mahwah, NJ, pp. 187–213.
Selkirk, E., 2000. The interaction of constraints on prosodic phrasing. In: Home, M. (Ed.), Prosody, Theory and Experiment: Studies Presented to Gösta Bruce. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 231–261.
Selkirk, E., 2009. On clause and intonational phrase in Japanese: the syntactic grounding of prosodic constituent structure. Gengo Kenkyuu 136, 35–73.
Selkirk, E., 2011a. The syntax–phonology interface. In: Goldsmith, J., Riggle, J., Yu, A. (Eds.), The Handbook of Phonological Theory. 2nd edition. Blackwell Publishing, Oxford, pp. 435–485.
Selkirk, E., 2011b. The syntax–phonology interface. In: Goldsmith, J., Riggle, J., Yu, A. (Eds.), The Handbook of Phonological Theory. 2nd edition. Blackwell Publishing, Oxford.
Selkirk, E., Elordieta, G., 2010. The role for prosodic markedness constraints in phonological phrase formation in two pitch accent languages. In: Handout of paper presented at Tone and Intonation in Europe (TIE), vol. 4. Stockholm University, Department of Scandinavian Languages http://www.hum.su.se/pub/jsp/polopoly.jsp?d=13236&a=73666, , retrieved 11/29/2011.
Selkirk, E., Tateishi, K., 1988. Constraints on minor phrase formation in Japanese. In: Papers from the 24th Annual Regional Meeting of the Chicago Linguistic Society. Chicago Linguistic Society, Chicago, pp. 316–336.
Selkirk, E., Tateishi, K., 1991. Syntax and downstep in Japanese. In: Georgopoulos, C., Ishihara, R. (Eds.), Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda. Kluwer, Dordrecht, pp. 519–544.
Selkirk, E., Shinya, T., Sugahara, M., 2003. Degree of Initial Lowering in Japanese as a Reflex of Prosodic Structure Organization. In: Proc.15th ICPhS, Barcelona, Spain, pp. 491–494.
Shinya, T., Selkirk, E., Kawahara, S., 2004. Rhythmic boost and recursive minor phrase in Japanese. In: Proceedings of the second international conference on speech prosody. pp. 345–348.
Truckenbrodt, H., 1999. On the relation between syntactic phrases and phonological phrases. LI 30, 219–255.
Vance, T.J., 2008. The Sounds of Japanese. Cambridge University Press, Cambridge, U.K..
Venditti, J.J., Maekawa, K., Beckman, M.E., 2008. Prominence marking in the Japanese intonational system. In: Miyagawa, S., Saito, M. (Eds.), The Oxford Handbook of Japanese Linguistics. Oxford University Press, Oxford, pp. 456–512.
Wackernagel, J., 1877. Der griechische Verbalaccent. Zeitschrift für vergleichende Sprachforschung 23, 457–470.
Zamma, H., 2001. Accentuation of person names in Japanese and its theoretical implications. Tsukuba English Studies 20, 1–18.