17

Prosodic typology: by prominence type, word prosody, and macro-rhythm*

SUN-AH JUN

17.1 Introduction

In the autosegmental-metrical (AM) model of intonational phonology, the two major properties marked by intonation are prominence and phrasing (e.g. Beckman 1996; Shattuck-Hufnagel & Turk 1996; Ladd 1996/2008). Intonational tunes in this framework are composed of pitch accents and/or boundary tones.[1] Pitch accents are prominent pitch targets or movements over a stressed syllable (or the head of a word), and boundary tones are pitch targets or movements marking prosodic structure and phrasing, and are typically realized at the edge of a prosodic unit. In the first volume of *Prosodic Typology*, I had proposed a model of prosodic typology (Jun 2005b) based on various languages whose intonation was described in the AM model of intonational phonology. That model of typology, therefore, included two major parameters of prosody—prominence and phrasing—each of which was examined at both lexical and postlexical levels. This was so because the prosodic property of an utterance is a combination of prosody at the word level as well as the phrase

* I would like to thank those who provided me with answers to my questions related to this chapter. They are in alphabetical order: Mariapaola D'Imperio, Janet Fletcher, Shelome Gooden, Martine Grice, and Carlos Gussenhoven. I also appreciate the feedback I received from the audience, especially Dani Byrd, Patricia Keating, and Stefanie Shattuck-Hufnagel, in the Special Session on Speech Rhythm in Production, Perception, and Acquisition, at the Fall meeting of the Acoustical Society of America, in San Diego, in 2011, and from the audience of the ToBI Workshop, held in the Department of Linguistics, the Ohio State University, in April 2012.

[1] The original model of intonational phonology which was based on stress languages (e.g. English, German) and lexical pitch accent languages (e.g. Swedish, Japanese) proposed that intonation is composed of pitch accent and boundary tones, but as has been claimed in Jun (2005b) and in this chapter, not every language has a pitch accent. For example, tone languages without stress and languages that mark prominence purely by their "edges" (e.g. Korean, W. Greenlandic) do not have intonational pitch accents.

level, and both word- and phrase-level prosody marks prominence and phrasing. The prominence marking at the lexical/word level was categorized by the type of lexical prosody, i.e. whether a language has lexical pitch accent, stress, tone, some combination of these, or none of these. The prominence marking at the postlexical/phrase level was categorized as head vs. edge, i.e. whether the prominence is cued by the head of a phrase (e.g. a nuclear pitch accent), by a boundary tone at the phrase edge, or by both. The parameter of phrasing was categorized by the types of prosodic units a language has at the lexical and postlexical levels. Lexical prosodic units include morae, syllables, and feet, and reflect the traditional typology of speech rhythm, e.g. syllable-timed vs. stress-timed. Postlexical prosodic units include an Accentual Phrase (AP), an Intermediate Phrase (ip), and an Intonational Phrase (IP). In sum, the model combined two traditions of prosodic typology (i.e. typology of word prosody and speech rhythm) with phrasal prosody as defined in the framework of intonational phonology.

However, because the phrasal prosody represented in the typology was prominence types and prosodic units, the model did not have any way to compare purely *tonal* aspects of prosody. As noted in Jun (2005b, p. 447), this typology did not capture similarities or differences across languages based on the tonal pattern of intonation. Specifically, it could not distinguish languages that have different global tonal patterns of utterances but belong to the same type of prominence marking, nor does it capture the similarity of global tonal pattern of languages belonging to different types of prominence marking. For example, English and Greek (Arvaniti & Baltazani 2005) were categorized as having the same feature of prosodic typology, i.e. stress-based head-prominence languages having same types of prosodic phrases. This grouping could not capture the fact that Greek has more regular phrase-medial tonal patterns than English. Similarly, both Chickasaw (Gordon 2005) and French (Jun & Fougeron 1995, 2000, 2002) were categorized as the head/edge-prominence languages, having their phrasal prominence marked by both pitch accents and an AP boundary tone, but the model could not capture the fact that French has much more regular intonation patterns than Chickasaw. On the other hand, Spanish (Beckman, Díaz-Campos, McGory, & Morgan 2002; Prieto & Roseano 2010) and Bengali (Hayes & Lahiri 1991a; Khan 2008, this volume) are intonationally similar by having a sequence of phrase-medial rising tonal patterns, but they differ in the prominence type and in what the rising tone is composed of. Spanish, being a head-prominence language, has a rising pitch accent, while Bengali, being a head/edge-prominence language, has a low pitch accent and a high AP boundary tone. This suggests that the global tonal pattern of an utterance is another prosodic dimension that is orthogonal to the types of prominence marking, motivating the need to include the degree of regularity in phrase-medial tonal patterns as a parameter of prosodic typology.

In this chapter, I propose a revised model of prosodic typology by considering the phrase-medial, global, tonal pattern of an utterance, called *macro-rhythm*, together

with prominence marking and word prosody. In my earlier model of prosodic typology (Jun 2005b), I used the term *macro-rhythm* to refer to the rhythm created by a prosodic unit larger than a word, to be in contrast with the traditional speech rhythm, which I called *micro-rhythm* as it is created by the repeated sequence of smaller units such as syllables or feet. Specifically, in the earlier model, *macro-rhythm* referred to the rhythm created by the regular tonal pattern of a sequence of small prosodic units (e.g. Accentual Phrase (AP)) as well as by a semi-regular tonal pattern of larger prosodic boundaries in an utterance (e.g. Intonational Phrase (IP)). But, in this revised model, I will narrow the definition of *macro-rhythm* to the phrase-medial tonal rhythm whose unit is equal to or slightly larger than a word, regardless of whether the tonal pattern is composed of edge tones (i.e. AP or word tones), head tones (i.e. pitch accents or lexical tones), or both. The tonal rhythm of phrases larger than an AP (e.g. IP or ip) is not included because the IP or ip tonal rhythm tends to vary more within a language (due to the variable size of IP or ip) and vary less across languages (for both size and the type of boundary tone[2]).

Perceived rhythm based on a regular pitch movement has also been noted in other studies (e.g. Barry 1981; Thomassen 1982; Lerdahl & Jackendoff 1983; Handel 1993; Dilley & Shattuck-Hufnagel 1999; Dilley & McAuley 2006, 2008; Andreeva, Barry, & Steiner 2007; Kohler 2008; Barry, Abdreeva, & Koreman 2009; Niebuhr 2009). Thomassen (1982) and Lerdahl & Jackendoff (1983) showed that repetitions of simple tonal sequences, such as a sequence of rising pitch contours or a sequence of falling pitch contours, affect perceived grouping of words and meter, and Barry and his colleagues (Barry 1981; Andreeva, Barry, & Steiner 2007; Barry, Abdreeva, & Koreman 2009) showed that fundamental frequency (Fo) contributes to the perception of rhythm as much as duration.

Furthermore, tonal rhythm has been shown to facilitate word segmentation. Recently, Dilley and McAuley (2006, 2008) showed that a sequence of three monosyllabic words, W1 W2 W3, with a H-L-H contour (i.e. H on the first syllable, L on the second syllable, and H on the third syllable), where any two adjacent words could be grouped to form a compound word, could be parsed as (W1 W2)(W3) or (W1)(W2 W3) (e.g. "*note book worm*" could be grouped either as (*notebook*)(*worm*) or (*note*)(*bookworm*)) depending on the tonal pattern of the preceding context. If the preceding context has a sequence of falling tones (<HL> <HL> <HL>) with the final L aligned with the syllable *before* W1, the last word the listeners heard was W3, suggesting that W1 and W2, having a H-L contour, were parsed as one compound word (e.g. (*notebook*)(*worm*)). On the other hand, if the tonal pattern of the preceding context is a sequence of rising tones (<LH> <LH> <LH>) with the final

[2] Boundary tones of IP or ip are often Low or High, and sometimes Mid (M), Falling (HL) or Rising (LH). The size of an IP is reported to be similar in English, German, and French (about 1.5 sec. Jun 2005b, p. 443).

H aligned with W1, W2 and W3 which form a L-H contour were perceived as the last word (e.g. (*note*)(*bookworm*)). This suggests that the tonal rhythm, a sequence of rising or falling pitch contour, facilitated the grouping of syllables, i.e. word segmentation. In this experiment, duration of the syllables was kept the same across the tonal conditions, so the grouping of syllables was purely due to the tonal rhythm. Dilley & McAuley (2008) also found that the effect of tonal rhythm on word segmentation was stronger than that of duration where the syllable duration was manipulated reflecting strong-weak stress-timed rhythm while Fo remained flat. Niebuhr (2009) found a similar result on German data regardless of whether the tonal context phrase comes before or after the target string.

A facilitating role of tonal rhythm on word segmentation was also shown in studies that examined non-stress languages. In languages that do not have lexical stress, each content word tends to form a small prosodic unit such as an Accentual Phrase (AP), defined by a regular tonal pattern. In French (Welby 2007), Japanese (Warner et al. 2010), and Korean (Kim 2004; Kim & Cho 2009), where an AP-initial boundary is marked by a rising tone, a word onset with a rising tonal pattern facilitated word segmentation. This supports the view that tonal rhythm (i.e. phrasing marked by tones) in edge-prominence languages functions like stress/pitch accent in stress-accent languages, by marking prominence of words and delivering information of syntactic, semantic, and information structure of a phrase (Jun 1993, 1998; Venditti, Jun, & Beckman 1996; Ladd 1996/2008). That is, focus and syntactic/semantic groupings are typically marked by changes in accentual phrasing in edge-prominence languages, instead of pitch accenting and/or deaccenting as in stress-accent languages.

In sum, languages differ prosodically by the type of prominence marking as well as macro-rhythm. They also differ by the way they mark head prominence and this can be captured by word prosody. Since word prosody is important in developing intonational phonology but is not predictable from the type of prominence or macro-rhythm, it is included as a third parameter of prosodic typology. That is, the revised model of typology includes three parameters: type of prominence, degree of macro-rhythm, and word prosody. The first two are phrase level parameters and the last one is a word level parameter. The traditional speech rhythm, i.e. micro-rhythm (e.g. stress-timed or syllable-timed), is not included in the revised model because it does not play a role in intonational phonology. (See the Discussion section for the relationship between macro-rhythm and micro-rhythm and the other two parameters).

The organization of this chapter is as follows. Section 17.2 describes the criteria to determine and predict the degree of macro-rhythm of languages analyzed in the AM model of intonational phonology. Section 17.3 shows how languages in each prominence type can be divided by their different degrees of macro-rhythm group based on their tonal inventory and the frequency and domain of pitch accents/boundary tones.

Section 17.4 shows how languages are classified in the new model of prosodic typology, and section 17.5 discusses the proposed model and how the degree of macro-rhythm is related to the phonetic realization of stress and micro-rhythm. It also provides a preliminary method of quantifying the degree of macro-rhythm and ends with a conclusion.

17.2 New proposal: macro-rhythm

Rhythm is the temporal organization of speech perceived by a regular occurrence of events, whether the event is auditory or visual and whether the acoustic medium is timing, fundamental frequency (Fo), or amplitude (e.g. Lerdahl & Jackendoff 1983; Handel 1993; Barry et al. 2009; Niebuhr 2009). As mentioned earlier, the traditional speech rhythm (micro-rhythm) is formed by a sequence of syllables, morae, or a sequence of alternating strong and weak syllables. Macro-rhythm is a tonal rhythm, a rhythm perceived by changes in Fo. Therefore, a stronger degree of macro-rhythm would be created by a sequence of alternating low and high tones (L-H-L-H-L-H). Here, a tone is not necessarily associated with a stressed syllable (i.e. a pitch accent) or the edge of a prosodic unit (i.e. a boundary tone). A subunit of tonal rhythm (i.e. LH or HL) can cover materials larger or smaller than a word, and unlike the Tonal Unit in Hirst & Di Cristo's (1996) model of French intonation, it does not necessarily form a prosodic unit.

Both types of rhythm are shown in (1): a schematic Fo contour representing (tonal) macro-rhythm and the metrical grid representing micro-rhythm are provided for the sentence, *Mariana loves marmalade*. Grid marks ("x") represent the head of a metrical unit (Prince 1983). In the bottom row, each syllable gets one grid mark, and at the top row, the head of a foot gets one grid mark. ( ) represents a foot.

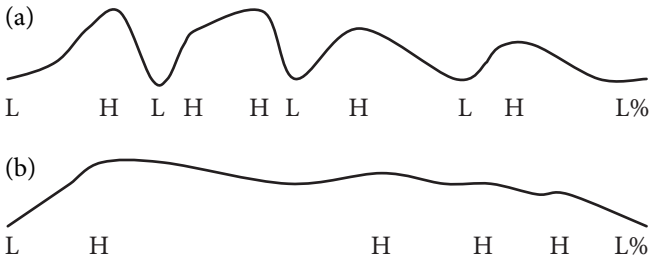
(1)

<= macro-rhythm

(x) (x) (x) (x) (x) <= micro-rhythm
x x x x x x x x
Mari a na loves marmalade.

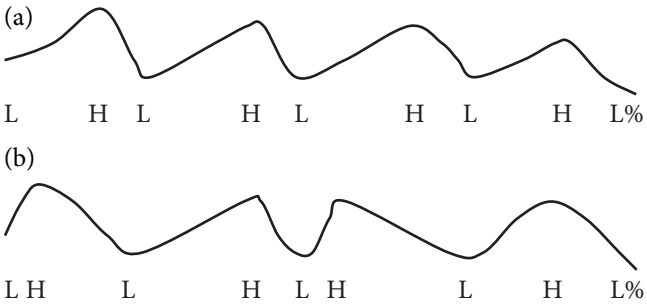
The degree of macro-rhythm can be evaluated based on a few rhythm rules. First, a pitch contour with a sequence of level tones is less macro-rhythmic than a contour with a sequence of rising or falling tones. That is, an alternation of Fo peaks (H) and valleys (L) is more macro-rhythmic than lack of alternation. (2) shows a schematic pitch contour illustrating the point. (2a), where L and H are alternating, is more macro-rhythmic than (2b).

(2) Stronger macro-rhythm, rule 1: Low/High alternation



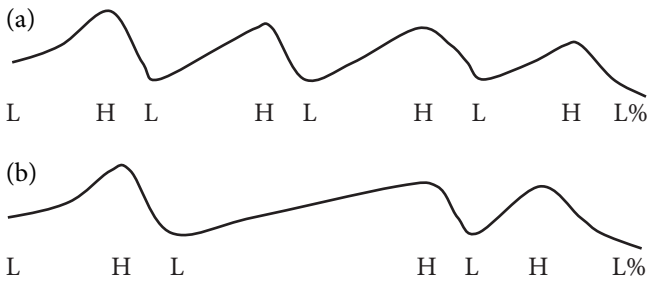
Second, a pitch contour with a sequence of similar sub-tonal units is more macro-rhythmic than that with less similar ones. In (3), both contours have a sequence of four rise-falls but (3a) is more macro-rhythmic than (3b) because the shape of each rise-fall in (3a) is more similar than that in (3b). That is, similarity of sub-tonal unit matters.

(3) Stronger macro-rhythm, rule 2: Similarity of sub-tonal units



Third, a pitch contour with a regular interval of sub-tonal unit is more macro-rhythmic than that of irregular intervals. (4a) is more macro-rhythmic than (4b) because the interval of each rise-fall is more regular in (4a) than in (4b). Thus, the regular timing or interval of sub-tonal unit matters.

(4) Stronger macro-rhythm, rule 3: Regular interval of sub-tonal units



In sum, a pitch contour has a stronger degree of macro-rhythm if the contour is composed of alternating Low and High tones and if the sub-tonal units (e.g. LH or

HL) of the pitch contour are more similar to one another and occur at regular intervals. These rules can be converted into the three criteria described in (5) by which we can predict and compare the degree of macro-rhythm of languages analyzed in the AM model of intonational phonology. Specifically, we can consider the number and the type of phrase-medial pitch accents/AP/word tones based on (3) and (2), respectively, and the frequency/domain of pitch accents/AP/word tones based on (4).

(5) Three criteria for predicting the degree of macro-rhythm of a language analyzed in intonational phonology

(i) *The **number** of possible phrase-medial pitch accent/AP/word tones*
Among the languages that have pitch accents, those with more types of phrase-medial pitch accents would have more variable pitch contour and thus be less macro-rhythmic than those with fewer types. Similarly, among languages that have an accentual phrase (AP) or word tones, those with more types of AP/word tones are less macro-rhythmic than those with fewer AP/word tones.

(ii) *The **type** of most common phrase-medial pitch accent/AP/word tones*
Languages employing a rising (e.g. L+H* or L*+H) or a falling (e.g. H+L* or H*+L) tone as the most common phrase-medial pitch accent are more macro-rhythmic than those employing a level pitch accent (i.e. H* or L*). Similarly, languages that most often employ rising or falling AP/word tones are more macro-rhythmic than those with level AP/word tones. Head/edge-prominence languages that have a rising or falling contour formed by a sequence of a pitch accent and AP tone of opposing tonal targets are more macro-rhythmic than those that have a level contour by having the same type of tones for pitch accent and AP tone.

(iii) *The **frequency** or **domain** of pitch accents/AP/word tones*
Languages where every word receives a pitch accent or AP/word boundary tone are more macro-rhythmic than those with less or more frequent pitch accents or AP/word boundaries per word. Similarly, edge-prominence languages where the edge of every word is marked by a tone or where every word forms one AP are more macro-rhythmic than those where the edge of a word is not always marked by a tone or where an AP tends to include more than one word. The "frequency" criterion is, therefore, equal to "the domain" of head/edge tones.

In sum, macro-rhythm is defined as phrase-medial tonal rhythm whose unit is equal to or slightly larger than a word, and the tones forming a tonal unit can be pitch accents, lexical tones, or boundary tones. Strong macro-rhythm is achieved if every word forms one tonal unit and the tonal units have similar Fo shape.

17.3 Macro-rhythmicity in each prominence type

Based on the criteria in (5), we can predict that Egyptian Arabic would be more macro-rhythmic than English given that the most common prenuclear pitch accent is L+H* in the former (Chahal & Hellmuth, this volume) but H* in the latter (Dainora 2001, 2006). On the other hand, Egyptian Arabic might have a similar degree of macro-rhythm as Bangladeshi Bengali (Khan 2008, this volume) because both languages mark each word by a rising pitch. However, since Egyptian Arabic is a head-prominence language while Bengali is a head/edge-prominence language, their way of marking prominence is not the same. Thus, in this section, macro-rhythm of various languages is classified within each type of prominence (see Tables 17.1–17.3). Though the degree of macro-rhythmicity is presumably gradient, languages are divided roughly in three macro-rhythm groups: Strong, Medium, and Weak.

17.3.1 *Head-prominence languages*

Head-prominence languages are those in which phrase-level prominence is marked by the phrase *head*, which is either derived from the head of a word, i.e. a designated syllable or mora of a word, or determined at a phrasal level. Therefore, languages that have a pitch accent at a phrasal level, regardless of whether the pitch accent location is determined lexically (e.g. lexical pitch accent languages such as Japanese and Swedish) or postlexically (e.g. lexical stress accent languages such as English and Arabic or postlexical stress-accent languages such as French) are head-prominence languages.

Tone languages would also belong to head-prominence languages because phrasal prominence in a tone language is derived from the tonal specification of a specific syllable(s) or mora(s) in the lexicon. Thus, in this model, a head is not necessarily a metrical one, and not necessarily observing the property of culminativity (Hyman 2006). A pitch accent head at a phrasal level can be metrical (English) or tonal (Japanese) and both types observe culminativity by being associated with one syllable or mora in a word, but a tone head in a free (or "unlimited") tone language is associated with all syllables/morae in a word. That is, a tonal head is "distributed" over a word. The domain of a head is, therefore, equal to or smaller than a word. Phrasal prominence in a tone language is realized by manipulating pitch range, amplitude, duration, and phrasing while not changing the identity of the head tones. A similar method of phrasal prominence is used by some lexical pitch accent languages (e.g. Japanese, Somali; however, phrasal prominence tone is introduced postlexically in Swedish and Papiamentu (Remijsen et al. 2008, this volume)).

This subsection includes languages whose prominence is marked only by the head, but not by the edge. Head-prominence languages that also cue prominence by marking the edge of a prosodic unit (e.g. Japanese, French) will be described in the next subsection, head/edge-prominence languages.

TABLE 17.1 Macro-rhythm of head-prominence languages defined by the inventory of pitch accents/lexical tones and the domain (i.e. frequency per content word) of head tones. The most common tone(s) in the tone inventory are italicized and underlined

Head-prom. Languages	*/tone inventory	Domain of head
	Strong macro-rhythm	
Arabic, Egyptian	*L+H**	≈ Cwd
Catalan	H*, L*, L+H*, *L+<H**, L*+H, H+L*	≥ Cwd
Greek	H*, L*, L+H*, *L*+H*, H*+L	≈ Cwd
Italian, Neapolitan	*H**, L*, *L+H**, L*+H, H+L*	≈ Cwd
Papiamentu	*lexical HL*, focal LH	≥ Cwd
Portuguese, Brazilian[a]	H*, L*, *L+H**, L*+H, H+L*	≈ Cwd
Samoan[b]	*LH**	≈ Cwd
Spanish, Castilian[c]	H*, L*, *L+<H**, L*+H, H+L*	≈ Cwd
Swedish[3]	*HL*, H*L*, focal H	≥ Cwd
Tone languages with a single tone melody		
	Medium macro-rhythm	
Arabic, Lebanese	*H**, L*, L+H*, H+!H*	≥ Cwd
Dutch	*H**, L*, L*H, *H*L*, H*LH	≥ Cwd
German	*H**[4], L*, L+H*, L*+H, H+L*, H+!H*	≥ Cwd
English[d]	*H**, L*, L+H*, L*+H, H+!H*	≥ Cwd
Jamaican Creole English	*H**, L*, L+H*, L*+H, *H+L**	≥ Cwd
Tone languages with multiple tone melodies		
	Weak macro-rhythm	
Cantonese	6 tones (levels and contours)	≤ Cwd
Mandarin	4 tones (L, H, LH, HL)	≤ Cwd
Portuguese, European	*H*/(H)*, L*, L*+H, H+L*, H*+L	> Cwd
Wolof[e], Kuot[f]	no pitch accent[5]	
Free tone languages		

[a] Libman (2008), de Moraes (2007); also see Vigário & Frota (2003) for some northern varieties of European Portuguese; [b] Yu (2009), Orfitelli & Yu (2009), N.B. LH* is my label. In their model, it is represented as LH; [c] Estebas-Vilaplana & Prieto (2010); [d] Dainora (2001, 2006); [e] Rialland & Robert (2001); [f] Lindström & Remijsen (2005).

[3] This data is based on Bruce (1977, 2005). Different tonal representations have been proposed by Riad (1998, 2006) (H* for accent 2, L or no tone for accent 1, LH for prominence tone), and by Myrberg (2010) (same as Bruce for word accent types but (H)LH for prominence tone). Myrberg posits an AP (Accent Phrase) as the domain of HL-word accent, but this AP is not marked by an edge tone. So, considering all these models, Swedish would belong to a head-prominence language.

[4] The rising slope of German H* is steeper than that of English, thus close to L+H* phonetically (Grice, pc). So, it is possible that German has higher degree of macro-rhythm than English if quantified based on Fo contour.

[5] Wolof (Rialland & Robert 2001) and Kuot (Lindström & Remijsen 2005) have been claimed to have lexical stress but no intonational pitch accent. Since these languages have stress, they would belong to head-prominence languages, but because what is intonationally marked in these languages is a large phrase and not a word-or AP-size unit, these languages would have weak degree of macro-rhythm.

Table 17.1 provides the information corresponding to the three criteria in (5) for various languages analyzed in the AM model of intonational phonology, in three groups of macro-rhythm: strong, medium, weak. The first column lists languages belonging to each macro-rhythm group. The second column lists the inventory of pitch accents in declaratives[6] or lexical tones in each language, with the most common one(s) underlined and italicized. Since a downstepped High tone (e.g. !H*) counts as a High tone in macro-rhythm, !H* is not included in the pitch accent type inventory. The third column lists the domain of pitch accent or lexical tones, i.e. frequency of head tones per word. "≈ Cwd" means the domain is approximately one content word (= Cwd), i.e. the head tone occurs roughly once every content word, "≥ Cwd" or "≤ Cwd" means the domain is slightly larger or smaller than a content word, respectively, and "> Cwd" means the domain is clearly larger than one content word. (In Tables 17.1, 17.2, and 17.3, the information in the 2nd and 3rd column for each language is given based on the descriptions provided in the earlier volume of *Prosodic Typology* (Jun 2005b) and the current volume. For those languages not included in these volumes or those that have better references, the references are provided under the table.)

Head-prominence languages belonging to the strong macro-rhythm group have a rising or falling pitch accent as the most common tone in a phrase. Among these languages, Egyptian Arabic and Samoan would be more macro-rhythmic than others because they have only one type of pitch accent while others have multiple types of pitch accent.

Head-prominence languages belonging to the medium macro-rhythm group include languages such as English, German, Jamaican Creole English, Lebanese Arabic, and Dutch. These languages are less macro-rhythmic than Spanish-type languages in that they generally have multiple pitch accent types and the most common phrase-medial pitch accent in declaratives is a level tone (e.g. H*), and the domain of the head is slightly larger than one content word. In American English, the most common pitch accent in declaratives is H* (Dainora 2001, 2006), thus similar to Neapolitan Italian, but in English verbs and items with old information often do not receive a pitch accent while in Italian every content word receives a pitch accent even when delivering old information (Ladd 1996/2008).

Table 17.1 also shows that different varieties of the same language can have different degrees of macro-rhythm. Egyptian Arabic and Brazilian Portuguese are expected to have stronger macro-rhythm than Lebanese Arabic and European Portuguese, respectively, because the former group has a single rising pitch accent (L+H*) on almost every content word while the latter group has multiple types of pitch accent with the domain of pitch accent being larger than a word, i.e. not every

[6] The degree of macro-rhythm might change if other sentence types are included.

content word carries a pitch accent (see Chahal & Hellmuth, this volume, for Arabic; Frota, this volume, for Portuguese). Table 17.1 puts European Portuguese under the Weak macro-rhythm group and Lebanese Arabic under the Medium macro-rhythm group because the domain of pitch accent is clearly larger than a Word in European Portuguese but only slightly larger than a Word in Lebanese Arabic.

The macro-rhythm of tone languages would also range from Strong to Weak. Free tone or limited tone languages such as Cantonese or Yoruba would have Weak macro-rhythm since each syllable/word can carry any of the tone types, thus not likely to have any regular alternation of High and Low within a phrase. Among these, contour tone languages like Cantonese would be less macro-rhythmic than level tone languages like Yoruba because the H/L alternation can occur *within* a syllable more often in the former than the latter. On the other hand, tone languages with one tonal melody would have a stronger degree of macro-rhythm. If a tone language has a single tonal melody as in Somali (Hyman 1981) and Safwa (Odden 1988, 1990), it would be similar to lexical pitch accent languages, having a strong degree of macro-rhythm. If a tone language has multiple tone melodies such as Kimatuumbi (Odden 1988), Isthmus Zapotec (van der Hulst & Smith 1988), and Skou (Donohue 2003, cited in Hyman 2006), it would be similar to English-type languages, having a medium degree of macro-rhythm.

17.3.2 *Head/edge-prominence languages*

Head/edge-prominence languages are languages where the prominence at a phrase level is marked by both the *head* and the *edge* of a phrase. That is, they are head-prominence languages but also have a word/phrasal tone marking the edge of a word boundary such as an AP. This prominence type includes languages that have a postlexical pitch accent as well as an AP-like phrasal or boundary tone (e.g. French, Bengali, Farsi, Georgian, Tamil, Kiche) or a lexical pitch accent as well as a word/AP boundary tone (e.g. Lekeitio Basque, Serbo-Croatian, Tokyo Japanese).

Typically, these languages have only a few fixed types of pitch accents, and if having multiple types, they often have one dominant one, rising or falling tone. In addition, the location of stress/pitch accent is often the edge (initial or final syllable) of a word. For example, French (Jun & Fougeron 2002) and Kiche (Nielsen 2005) have pitch accents on word-final syllables, while Bengali (Hayes & Lahiri 1991; Khan 2008, this volume), Georgian (Vicenik & Jun, this volume), and Tamil (Keane, this volume) have pitch accents on word-initial syllables. Furthermore, the boundary tones marking the edge of a word/AP are often the opposite tonal type of the pitch accent (e.g. a L* pitch accent with a H boundary tone or a L boundary tone with a H* pitch accent), thus showing a rising or falling tonal pattern over a word/AP. Therefore, in general the head/edge-prominence languages are expected to have stronger degree of macro-rhythm than the purely head-prominence languages.

Table 17.2 Macro-rhythm of head/edge-prominence languages defined by the inventory of pitch accents/lexical tones and AP/Word tones as well as the domain (or frequency) of head or edge tones. The most common tone(s) in the tone inventory are in italics and underlined. "AP" represents a prosodic unit equal to or slightly larger than a word

Head/Edge-prom. Languages	*/AP/word tone inventory	Domain of head/edge
	Strong macro-rhythm	
Basque, Lekeitio	*H*+L* / *AP-initial LH*	≥ Cwd
Basque, standard	*(L+H)** / optional AP-final rise	≥ Cwd
Bengali, Bangladeshi	H*, *L**, L*+H / *AP-final H* and L	≈ Cwd
Farsi[a]	*L*+H* / AP-final H	≈ Cwd
French[b]	AP-final L*, H*, *LH**, optional AP-initial LH	≥ Cwd
Georgian	*L**, L*+H / *AP-final H*, L, optional AP-med H+L	≥ Cwd
Japanese, Tokyo	*H*+L* if accented / *AP-initial LH*	≥ Cwd
Kiche, Cantel[c]	H*, *LH** / AP-final H	≈ Cwd
Serbo-Croatian	*L*+H, H*+L* / word-initial L%	≈ Cwd
Tamil	*L** / *AP-final H*	≈ Cwd
Tongan[d]	*LH* / AP-final H*	≈ Cwd
	Medium macro-rhythm	
Bininj Gun-wok	*H**, <H*, ˆH*, L+H* / PhP-final L	≥ Cwd
Chickasaw	*H**, H*λ* / optional AP-initial LH or final HL	>/<MorphWd[e]
Dalabon	*H**, L+H* / optional AP-final L or H	≈ Cwd

[a] Sadat-Tehrani (2008), Arbisi-Kelm (2007), Scarborough (2007); [b] Jun & Fougeron (2002); [c] Nielsen (2005); [d] Kuo & Vicenik (2012); [e] The domain of AP can be larger or smaller than a morphological word.

Table 17.2 shows macro-rhythm of head/edge-prominence languages in two groups: Strong and Medium. The format is the same as that of Table 17.1 except that the tone inventory includes the edge tones of AP or Word in addition to pitch accents and lexical tones.

Head/edge-prominence languages belonging to the Strong macro-rhythm group have a fairly fixed location of stress with most content words being realized as a rising tone, either by a combination of a level tone pitch accent and a boundary tone of the opposite tone type (e.g. L* and the AP-level high tone Ha in Bengali, Georgian, and Tamil) or a rising pitch accent followed by a H boundary tone (e.g. L*+H or LH* and Ha in Farsi and Kiche), or a rising pitch accent (LH*) simultaneously marking the head and the edge of an AP (e.g. French). Like Kiche, Farsi has a rising pitch accent followed by Ha but it would be slightly less macro-rhythmic than Bengali, Tamil, or Kiche because Farsi shows some variation in stress location depending on grammatical category (noun vs. verb) while Bengali and Tamil have fixed word-initial stress and Kiche has fixed word-final stress (in native and nativized words).

Though lexical pitch accent languages such as Tokyo Japanese, Serbo-Croatian, and Lekeitio Basque are included in the Strong macro-rhythm group, they would have a slightly weaker degree of macro-rhythm than the Bengali-type languages because, in addition to having a rising tone marking an AP/word boundary, these languages have more than one way to mark their heads (e.g. two types of lexical pitch accent in Serbo-Croatian, and accented vs. unaccented AP distinctions in Tokyo Japanese), and the domain of AP is often larger than a word (e.g. an AP includes multiple unaccented words in Lekeitio Basque). Standard Basque has no lexical pitch accent but has a rising postlexical pitch accent, (L+H)*, on most content words. Thus, it would be more macro-rhythmic than Lekeitio Basque. However, Standard Basque would still be less macro-rhythmic than the Bengali type because its stress location is variable and has an optional AP-final boundary tone.

Finally, the macro-rhythm of Bininj Gun-wok, Chickasaw, and Dalabon would be weaker than that of all the other head/edge-prominence languages mentioned in this section because, like English, these three languages do not have a fixed location of stress, and they have multiple types of pitch accent, with H* being the most common. Unlike English, however, they have an AP-like tonal unit, but the AP phrasal/boundary tone is either variable or optional (e.g. Chickasaw, Dalabon), and the size of AP is often larger than a word (e.g. Bininj Gun-wok) or variable (e.g. Chickasaw), thus not contributing much to the regularity of tonal rhythm in a phrase.

17.3.3 *Edge-prominence languages*

Edge-prominence languages are languages that do not have any lexically specified head (stress, pitch accent, tone), nor any postlexically marked head, so the prominence at the word and phrasal level is only marked by the *edge* of a word and a phrase. That is, they only have AP-like phrasal/boundary tones (e.g. Seoul Korean, Halh and Oirat Mongolian, West Greenlandic, accentless dialects of Japanese). Therefore, languages belonging to this group are "head-less AP-languages." An AP has been typically defined as a tonally marked prosodic unit which often contains one word, regardless of whether the AP has a head or not (e.g. in Tokyo Japanese, AP can include one pitch accent (H*+L) or none, Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988). Since each word in edge-prominence languages tends to be marked by a rising tone either at the beginning or the end of a word or both, these languages generally have strong macro-rhythm, similar to the Spanish type in head-prominence languages or the Bengali type in head/edge-prominence languages.

Table 17.3 shows that all edge-prominence languages that have been analyzed so far show a rising, falling, or rising-falling tone pattern over a word, thus belonging to the Strong macro-rhythm group. The format of the table is the same as that of Tables 17.1 and 17.2 except that the tone inventory includes edge tones only.

Table 17.3 Macro-rhythm of edge-prominence languages. For each language, the inventory and the domain of AP/Word tones are provided. The most common tone(s) in the tone inventory are in italics and underlined. PW = prosodic word

Edge-prom. Languages	AP/word tone inventory	Domain of edge
	Strong macro-rhythm	
Greenlandic, West	*PW-final HLH, HL*, LH	≈ Cwd
Japanese, Kobayashi	*Word-final H (= LH word tone melody)*	≈ PW
Japanese, Koriyama	*LHL AP tone melody*	≥ PW
Japanese, Yamagata	*LHL AP tone melody*	≥ PW
Korean, Seoul	*AP-initial LH or HH*, *AP-final LH*, L, H	≥ PW[a]
Mongolian, Halh	*AP initial LH*	≥ Cwd
Mongolian, Oirat[b]	AP tone melodies H, L, *LH*, HL, HH	≥ Cwd

[a] Jun & Fougeron (2002), Kim (2004); [b] Indjieva (2009)

Though edge-prominence languages generally have strong macro-rhythm, we could expect some variable degrees of macro-rhythm among these languages depending on the number of AP phrase/boundary tone types and the domain of AP tones. Edge-prominence languages with a fixed word/AP boundary tone (e.g. Kobayashi Japanese, Igarashi, this volume; Halh Mongolian, Karlsson, this volume) or a fixed AP tone melody (e.g. accentless Japanese dialects, Igarashi, this volume; Kumamoto Japanese, Maekawa 1994) would be more macro-rhythmic than those with multiple tonal melodies of AP or Word (e.g. West Greenlandic, Arnhold, this volume; Oirat Mongolian, Indjieva 2009; Korean, Jun 2000, 2005a). Macro-rhythm in Korean might be even weaker than that in West Greenlandic or Oirat Mongolian because the AP initial tone in Korean varies between High and Low depending on the laryngeal feature of the AP-initial segment while the other languages do not have this added segmental interaction. However, since the Korean AP-final syllable is typically marked by the rising tone LHa, it is expected that the macro-rhythm of Korean is still stronger than that of English type.

17.4 A new, revised model of prosodic typology

So far, we have classified languages based on the type of prominence marking and the degree of macro-rhythm. Classifying languages based on the prominence type allows us to explain how languages that seemingly differ prosodically from the traditional view of word prosody share a common property (e.g. Bengali, French, and Japanese all mark a word by both the head and the edge of a word while English and Papiamentu mark a word by the head only), or how languages that have same word prosody differ prosodically (e.g. Spanish and Farsi, though both are stress

languages, differ by the way they mark phrasal prominence). In addition, as mentioned in the Introduction, classifying languages based on the degree of macro-rhythm can allow us to explain how languages have similar or different phrase-medial intonation patterns regardless of their prominence type and word prosody (e.g. In tonal rhythm, Spanish and Egyptian Arabic differ from English, though these three languages all mark word prominence by stress; Spanish and Egyptian Arabic are similar to Bengali though Bengali marks word prominence by an AP tone as well as by stress.).

Section 17.3 showed that adding the parameter of macro-rhythm to the types of prominence marking captured differences between languages that belong to the same prominence type and similarities among languages that belong to different prominence types. However, since both the prominence type and macro-rhythm are parameters of phrasal prosody, we can further distinguish languages in each group based on their word prosody, i.e. whether a word is marked in the lexicon by stress, pitch, both of these, or none. For example, European Portuguese, Cantonese, and Mandarin belong to the head-prominence with Weak macro-rhythm group, but these languages differ by whether they have stress (European Portuguese), tone (Cantonese), or both (Mandarin). Therefore, combining all three parameters—prominence type, macro-rhythm, word prosody—would make a better model of prosodic typology by capturing both lexical and phrasal properties of prosody across languages.

Table 17.4 presents a revised model of prosodic typology, showing how languages introduced in Tables 17.1–17.3 are classified using the three parameters. For each of the three types of prominence marking (head, head/edge, edge), languages are grouped by four types of word prosody (stress, tone/lexical pitch accent, both of these, and none) and three degrees of macro-rhythm (strong, medium, weak). For word prosody, tone and lexical pitch accent categories are combined into one because the distinction is not always clear-cut (Hyman 2006), and their way of marking phrasal prominence in intonation is similar (see section 17.3.1). The three columns representing the three degrees of macro-rhythm are separated by a dotted line to show that, though the degree of macro-rhythm is represented in three levels, the boundary between the levels is not categorical but gradual. Shaded cells indicate that it is not likely to find a language that satisfies the combination of the three parameters. Languages in parentheses, written in a smaller font and italics, mean that they are possible candidates for the given combination of parameters but no supporting intonational models are available as of yet. That is, for the combination of head-prominence & tone/lexical pitch accent & strong macro-rhythm parameters, an example language would be a lexical pitch accent language with no AP or a tone language with one type of tonal melody such as Somali and Safwa. For the combination of head-prominence and tone/lexical pitch accent and medium macro-rhythm parameters, an example language would be a tone language with multiple tonal melodies such as Skou and Kimatuumbi.

Table 17.4 Prosodic typology based on three parameters (Prominence Type, Word Prosody, and Macro-rhythm) and example languages for each combination. Shaded cells denote combinations of parameters that are expected to have no example languages

Prom. type	Word prosody	Macro-rhythm		
		Strong	Medium	Weak
Head	Stress	Brazilian Portuguese, Castilian Spanish, Catalan, Egyptian Arabic, Greek, Italian, Samoan	Dutch, English, German, Jamaican Creole English, Lebanese Arabic	European Portuguese, Wolof, Kuot
	Tone/lexical pitch accent	(*Kihehe, Safwa*[a], *Somali*[b])	(*Isthmus Zapotec*[c], *Kimatuumbi*[a], *Skou*[d])	Cantonese
	Both	Papiamentu, Swedish		Mandarin
	None			
Head/ Edge	Stress	Bengali, Georgian, Kiche, Tamil, Tongan	Bininj Gun-wok, Dalabon	
	Tone/lexical pitch accent	Japanese, Leketio Basque		
	Both	Serbo-Croatian	Chickasaw[7]	
	None	French, Standard Basque		
Edge	None	Accentless dialects of Japanese, Halh Mongolian, Oirat Mongolian, Seoul Korean, West Greenlandic		

[a] Odden (1988, 1990); [b] Hyman 1981; [c] van der Hulst & Smith (1988); [d] Donohue (2003, cited in Hyman 2006).

Among the languages analyzed in the AM model of intonational phonology, those that are known to have no word-level prosody ("None" in Word prosody column) are either edge-prominence or head/edge-prominence languages. At the moment, no example is available for the category of head-prominence & no word prosody. It is expected that a language with postlexical pitch accent (but no lexical pitch accent or lexical stress) and without AP-like tones would belong to this category. It is further expected that this type of language would have a fairly regular occurrence of postlexical pitch accent that marks word prominence, thus having strong or medium, but not weak, degree of macro-rhythm (for reasons, see next section). More research is needed to confirm this expectation.

[7] Chickasaw has a morpholexical pitch accent as well as a phonological postlexical pitch accent (Gordon 2005).

17.5 Discussion and conclusion

17.5.1 *The typology model and macro-rhythm*

In this chapter, I have proposed a new model of prosodic typology by classifying languages using three parameters: type of prominence marking, word prosody, and macro-rhythm. Revised from the version proposed in Jun (2005b), macro-rhythm represents the regularity of the High/Low alternation relative to word boundaries in the intonation contour of a phrase. This parameter is significant in distinguishing languages that belong to the same type of word prosody and prominence marking but differ by the global pattern of phrase-medial intonation contour, while capturing similar intonation patterns across languages that do not belong to the same type of word prosody and prominence marking.

The typology model in Table 17.4 suggests that a language can have strong macro-rhythm regardless of its type of prominence marking. That is, an intonation pattern of a language can be rhythmic whether the intonation contour is composed of pitch accents (lexical or postlexical), lexical tones, or phrasal tones. However, in general, the degree of macro-rhythm is stronger in edge-prominence languages and weaker in head-prominence languages. Edge-prominence languages generally have strong macro-rhythm probably because, by having no word prosody, they regularly mark the edge of a word-size unit either by rising or falling pitch. Head-prominence languages, especially those that have lexical stress, have a wide range of macro-rhythm types because stressed syllables can have multiple pitch accent types and the domain of pitch accent can vary from one syllable to several words. On the other hand, languages that have lexical pitch accent are expected to have strong macro-rhythm because they typically have only one or two types of pitch accent: either rising or falling.

Then, what would be the function of macro-rhythm? Why do languages have a tonal rhythm and why is the melody in one language more rhythmic than that of another? Intonation models of various languages suggest that languages often mark content words prosodically, either on their head, their edge, or both. That is, intonation enhances word prominence, and the tonal rhythm of a phrase must be a result of marking a sequence of words in a phrase. Languages also mark groupings of words by changing the tonal pattern. Therefore, intonation also cues the organization of speech while facilitating word segmentation and sentence processing (e.g. Cutler & Norris 1988, 2002; Kielgaard & Speer 1999; McQueen, Otake, & Cutler 2001; Cutler & Otake 2002; Schafer & Jun 2002; Salverda, Dahan, & McQueen 2003; Welby 2007; Dilley & McAuley 2008; Ito & Speer 2008; Kim & Cho 2009; Warner et al. 2010).

One of the functions of intonation is to mark word prominence, but if a language has other ways to mark word prominence, such as amplitude and duration, it is

expected that tones would play a weaker role in marking word prominence. In other words, the strength of a language's macro-rhythm seems to be inversely correlated with the strength of phonetic realization of stress. Languages with phonetically strong stress (with strong amplitude and long duration) would have weaker macro-rhythm than those with phonetically weak stress (e.g. English vs. Bengali).

We can project this relationship to the differences in the realization of stress between head-prominence and head/edge-prominence languages. In head-prominence languages, stress is often realized fully through longer duration and stronger amplitude, but in head/edge-prominence languages, the acoustic realization of stress is generally weak and the existence of stress is often controversial among researchers. This weak stress is probably compromised by the presence of edge tones. That is, edge tones will help to mark the prominence of word when the head marking (stress) is weak, which is common in head/edge-prominence languages. That is, either a language with weak stress may need an edge tone to boost word prominence, or a language with an edge tone consistently marking word prominence may not need strong stress.

Another common phenomenon in head/edge-prominence languages is that the head tone is often located at the edge of a word, thus also functioning to mark the edge of a word (e.g. French). This pattern is also seen in those head-prominence languages where the location of stress is fairly fixed. In this case, there is no edge tone, but the head tone functions like an edge tone by indirectly marking the word edge or marking the vicinity of a word edge, as in Spanish and Samoan (Yu 2009). It seems that this type of language generally has stronger macro-rhythm than languages where the location of stress is more variable.

This leads us to predict the relationship between macro-rhythm and micro-rhythm. It is known that syllable- and mora-timed languages have less variability in vocalic duration than do stress-timed languages (e.g. Ramus et al. 1999; Grabe & Low 2002; Arvaniti 2009). Therefore, the duration cue would not help much for word prominence or word segmentation in syllable-timed languages, suggesting that these languages may employ more regular pitch variation to cue word prominence, i.e. stronger macro-rhythm. Alternatively, syllable-timed languages may have fewer tonal contrasts (thus being more macro-rhythmic), because the lack of variation in syllable duration would limit the number of tonal contrasts allowed in the syllable. Stress-timed languages may be generally less macro-rhythmic because word prominence and word segmentation would be marked by non-tonal cues such as amplitude and/or duration.

17.5.2 *Quantifying macro-rhythm*

The model of prosodic typology proposed in this chapter is based on the prosody of languages described in the AM model of intonational phonology. That is, the

typology was developed based on phonological models of intonation. Thus, the parameter of macro-rhythm was divided in three groups based on categorical data such as the number of head/edge tones or the type and the domain of head/edge tones. However, as mentioned earlier, the degree of macro-rhythm is not categorical but gradual, and the degree of macro-rhythm could be quantified by measuring various aspects of Fo contours. The following is a preliminary suggestion for this direction.

Since macro-rhythm is defined as phrase-medial tonal rhythm whose unit is equal to or slightly larger than a word, we can quantify the strength of macro-rhythm by calculating the degree of variation in the interval and shape of tonal units in a phrase. The interval of tonal unit can be represented by the interval of Fo peaks and valleys and the shape of tonal unit can be represented by the slope of Fo rise and fall. Therefore, a variation index of macro-rhythm (Mac*R*_Var) would be a sum of standard deviations (SD) in each of these measurements, shown in (6). An Fo contour with a higher variation index can be interpreted as having a weaker degree of macro-rhythm.

(6) Quantifying macro-rhythm (Mac*R*)
Mac*R*-variation index (Mac*R*_Var): the higher the value of Mac*R*_Var, the weaker the macro-rhythm
Mac*R*_Var = SDp+SDv+SDr+SDf
SDp: SD of peak-to-peak interval, SDv: SD of valley-to-valley interval,
SDr: SD of rising slope, and SDf: SD of falling slope

Another way to measure the degree of macro-rhythm would be the frequency of High/Low alternations in a phrase, Mac*R*_Freq. Since the tonal unit is expected to be equal to or slightly larger than a Word, Mac*R*_Freq would be an average number of H/L alternations per word within a phrase. Macro-rhythm would be strong if the Mac*R*_Freq is close to 1. Here, a word should be a Prosodic Word (PW), not a morphosyntactic word, and a phrase should be an Intonational Phrase (IP). An IP would be a better choice than an Intermediate Phrase (ip) because not every language has an ip. Since Mac*R*_Freq measures the frequency of a H/L alternation, this information can be derived from the peak-to-peak duration data from the Mac*R*_Var index.[8]

In quantifying macro-rhythm, I have considered the intonation patterns of declarative sentences described in the literature, mostly based on read data.

[8] Macro-rhythm proposed in the current chapter does not consider tone-text alignment, e.g. not distinguishing late rise vs. early rise (L*+H vs. L+H*). What matters is whether high and low tones are alternating at a regular interval. It's possible that a different peak alignment could be an important parameter in distinguishing intonation across languages. A parameter of tone-text alignment can be added later when it's proven important.

Therefore, it should be noted that the degree of macro-rhythm could change if we add other sentence types or include speech materials produced in different styles of speech. Further research is needed to quantify macro-rhythm across various data.

17.5.3 *Conclusion*

In conclusion, this chapter presents a revised model of prosodic typology from Jun (2005b). The new model includes three parameters—prominence type, word prosody, and macro-rhythm—with revised definitions of "macro-rhythm" and "head" from Jun (2005b). The parameter of macro-rhythm captures the regularity of phrase-medial intonation patterns across languages. This rhythm is different from traditional notions of speech rhythm, which is perceived from a sequence of syllables or feet through the amplitude and duration of each rhythm unit. Adding the dimension of phrase-level prosody and intonation provides a new direction in prosodic typology. I have shown how languages can be classified into different degrees of macro-rhythm and discussed how word prosody, prominence types, micro-rhythm, and realization of stress are all related to macro-rhythm by means of a simple function, i.e. marking prominence of a word, thus facilitating word segmentation, sentence processing, and language acquisition.

Finally, the model of prosodic typology proposed in the current chapter is only preliminary, based on a few dozen languages. In order to advance prosodic typology, more research is needed to establish intonational phonology of various languages and to quantify their macro-rhythm.