16

Prosodic Typology

Sun-Ah Jun

16.1. INTRODUCTION

Studies on prosodic typology are in general rare probably because prosodic features are not easy to define and categorize, and also because prosodic features of languages have been described, if at all, with different assumptions and within different frameworks. Finding similarities and differences of prosodic features across languages would make sense only if these languages were described in the same framework in terms of the same prosodic categories. Comparisons of a prosodic system based on phonetic descriptions would have limitations because the similarities shown in the surface realization do not guarantee the same underlying distinctive prosodic features or structures. Since the advent of the phonological model of intonation in the 1980s, especially, the Autosegmental-Metrical (AM) model of intonational phonology, prosody has been described in terms of a prosodic structure and distinctive tonal categories (e.g. Pierrehumbert 1980; Gussenhoven 1984; Liberman and Pierrehumbert 1984; Beckman and Pierrehumbert 1986; Pierrehumbert and Hirschberg 1990; Ladd 1992, 1996). This has made it much easier to compare the prosody of languages categorically and closer to the goal of establishing prosodic typology (e.g. Ladd 2001)

In the AM model, prosody is described in two aspects: the prosodic structure of an utterance and the prominence relations within the structure (cf. Beckman 1996; Ladd 1996; Shattuck-Hufnagel and Turk 1996; Fougeron 1998; Jun 2003*a*). A prosodic structure is a hierarchical organization of prosodic units from the smallest prosodic unit (Mora or Syllable) to the largest (Intonation Phrase or Utterance). Within a phrase, some words are more prominent than others; and

I would like to thank Bob Ladd for his valuable comments on an earlier version. I would also like to thank those who have provided me with their papers and their expertise. They are in alphabetical order: Gösta Bruce, Dana Chahal, Mariapaola D'Imperio, Gorka Elordieta, Jennifer Fitzgerald, Janice Fon, Sonia Frota, Svetlana Godjevac, Matt Gordon, Larry Hyman, Kikuo Maekawa, Aniruddh Patel, Shu-hui Peng, Pilar Prieto, Stefanie Shattuck-Hufnagel, and Jennifer Venditti.

within a word, some syllables are more prominent than others. A prosodic structure and prominence relations are realized by suprasegmental features such as pitch, duration, and/or amplitude as well as segmental properties such as the realization of consonants and vowels. Furthermore, the prosodic property of an utterance is a combination of prosody at the word level and prosody at the phrase level. Postlexical prosody is constrained by the lexical prosody, and postlexical prosodic information contains information about the lexical prosody.

One of the most well-known properties of prosodic typology at the lexical level is word prosody: whether a lexical item has tone, stress, or lexical pitch accent (Trubetzkoy 1939; Beckman 1986; Ladd 1996; Fox 2000 and references therein). Under this typology, languages have been categorized as tone languages such as Mandarin and Hausa, as stress languages such as English and German, or as lexical pitch-accent languages such as Japanese and Basque. However, as shown in the previous chapters, a language can be specified with more than one such lexical feature or specified with none of them (see also Remijsen 2001). For example, a tone language can have stress, as in Mandarin, or not, as in Cantonese. Languages known to have stress can also have lexical pitch accent, as in Chickasaw, or not, as in English and other West Germanic languages. Languages known to have lexical pitch accent can have stress, as in Serbo-Croatian and Swedish, or not, as in Japanese. Finally, non-tonal and non-lexical-stress languages can have lexical pitch accent, as in Tokyo Japanese, or not, as in Seoul Korean. In sum, whether a language has stress or not is independent of whether a language has tone or lexical pitch accent specification.

Furthermore, all such lexical features interact with postlexical prosody, especially intonation, and intonational features are not directly predictable from the lexical prosodic features. That is, a pitch event in a syllable can provide postlexical information such as postlexical pitch accent (or sentence stress), a phrasal tone, or a boundary tone. Postlexical pitch accent can be from lexical pitch accent (e.g. Japanese) or from stress, either lexical (e.g. English and most other West Germanic languages) or postlexical (e.g. French). A phrasal tone, which marks a prosodic unit such as a Prosodic Word or an Accentual Phrase, can be found in languages with lexical pitch accent (e.g. Japanese, Serbo-Croatian), stress (e.g. Chickasaw, Farsi) or with no lexical specification (e.g. Korean). The postlexical pitch event can also differ in its function. It can mark prominence (e.g. pitch accent in English and other Germanic languages), demarcate a boundary of a prosodic unit (e.g. the accentual phrase in Japanese and Korean, the intonation phrase boundary tone in most languages), or both (e.g. demarcating pitch accent in French; Jun and Fougeron 1995, 2000, 2002). Therefore, in order to study prosodic typology, we need to examine postlexical prosodic features as well as the lexical prosodic features.

Another well-known property of prosodic typology is the rhythm or timing unit of a language. Languages have been categorized as mora-timed as in Japanese, syllable-timed as in Spanish, or stress-timed as in English (e.g. Pike 1945; Bloch 1950; Abercrombie 1967; Lehiste 1970). Though studies have shown that isochrony between syllables or between stress intervals does not have acoustic correlates and the division between syllable-timed and stress-timed is not straightforward (e.g. Lehiste 1977; Cutler 1980; Nakatani *et al.* 1981; Beckman 1982; Roach 1982; Dauer 1983, 1987; Cooper and Eady 1986), recent studies on the variation in vowel duration show that syllable-timed languages show less contrast between two consecutive vowel durations than stress-timed languages (Ramus *et al.* 1999; Low *et al.* 2000; Grabe and Low 2002; Ramus 2002).[1]

Even though our perception of the prosody of a language is influenced by the rhythm unit smaller than a word (Mora for mora-timed, Syllable for syllable-timed, or Foot for stress-timed), it is also influenced by the prosodic grouping at or above the Word such as a Prosodic Word, an Accentual Phrase, or an Intonational Phrase. A prosodic unit above the Word also has a rhythmic nature. Each unit often begins or ends with a prominent syllable or word, either due to pitch accent, a boundary tone, or both. A small prosodic unit such as an Accentual Phrase often has only one pitch accent, and is either followed by unaccented words (e.g. Japanese) or preceded by unaccented words (e.g. French); and similarly, an Intonation Phrase in stress accent languages often ends with a nuclear pitch accent, the most prominent pitch accent in the phrase. Furthermore, the edge of a prosodic unit above the Word tends to occur at a regular interval, and each prosodic unit avoids being too short (like stress clash) or too long (like stress lapse) unless it is influenced by other factors such as focus. For example, in Korean the size of the Accentual Phrase is three to five syllables on average, and an Accentual Phrase shorter than three syllables or longer than seven syllables is very rare (Korean Telecom 1996; Kim *et al.* 1997; Jun and Fougeron 2000).

As we cannot predict the postlexical prosody of pitch (e.g. intonational pitch accent, phrase accent) based on the lexical prosody (i.e., tone, stress, lexical pitch accent) of a language, the prosodic units above the Word are not predictable from the timing unit of the language. A language can have an Accentual Phrase, a small prosodic unit above the Word, whether it is mora-timed (e.g. Japanese), syllable-timed (e.g. French), or stress-timed (e.g. Chickasaw).

[1] However, this method does not tell what rhythmic unit a language has, e.g. syllable, foot. It only suggests a rhythmic unit of a language indirectly through the variability of syllable duration. A language with little contrast in vowel duration would be perceived as syllable-timed, but it is not clear if a language with more contrast in vowel duration is always perceived as stress-timed. More languages need to be examined to see if there is a strong correlation between the rhythmic unit and the degree of contrast in vowel duration.

As seen in previous chapters, prosodic units above the Word are often defined by intonation, but the analyses of tone languages such as Cantonese and Mandarin suggest that these languages do not have a tonal event consistently marking a prosodic unit. In Mandarin, an utterance is not always marked by a boundary tone, and there is no tonal event marking a prosodic unit within an utterance. But, the inventory of break indices suggests that speakers of this language perceive at least one level of prosodic grouping between the Word (syllable) and the Utterance (e.g. minor group, major group). This indicates that whether a prosodic grouping larger than the Word is marked by intonation or not is influenced by the functional load of pitch in the language (cf. Jun 1998). In sum, to study prosodic typology, prosodic units defined by the degree of juncture, i.e., break index in ToBI (Tones and Break Indices), should be included along with prosodic units defined by intonation.

In this chapter, I will attempt to provide prosodic typology by comparing the prosodic features of languages described in this book and several other languages which have been described in the AM model of Intonational phonology. Prosodic features which are common to these languages as well as those specific to a certain language will be discussed. The prosodic features of eleven languages whose ToBI transcription systems are described in this book (see Chapter 2 for detailed descriptions of ToBI) will be discussed in more detail. Finally, I will discuss the flexibility and the extension of the ToBI system: how the ToBI system has been extended to incorporate different prosodic systems while maintaining its integrity.

The organization of this chapter is as follows. Section 16.2 presents a summary of the prosodic systems of eleven languages based on the ToBI models described in this book, and Section 16.3 proposes a model of prosodic typology based on aspects of prominence and rhythmic/prosodic structure. Section 16.4 presents the flexibility of and other issues regarding the ToBI transcription system, and Section 16.5 concludes the chapter.

16.2. SUMMARY OF ELEVEN ToBI SYSTEMS

Table 16.1 shows a summary of the eleven ToBI systems described in the previous chapters, focusing on the tones and break indices. It includes the ToBI systems of Mainstream American English (MAE_ToBI), standard German (GToBI), Athens Greek (GRToBI), Neapolitan Italian (IToBI),[2]

[2] The chapter on Italian ToBI (Chapter 13) describes the tonal patterns and prosodic structure of four dialects (three Southern dialects, Neapolitan, Bari, and Palermo, and a central one, Florentine). Neapolitan is chosen arbitrarily. It is claimed that the other dialects use the same tonal shapes (falling, rising, high, and low) but that the tonal inventories are not identical due to different tone-text alignments.

TABLE 16.1 Summary of the eleven ToBI systems introduced in the previous chapters

Lang.	Types of tiers—extra only	Types of break indices (BI)	Types of tones on the tones tier	Prosodic units
English		0, 1, 2, 3, 4	L*, H*, L+H*, L*+H, H+!H*	
			L -, H-	ip
			L%, H%	IP
			! (for H pitch accent), <, >	
German		3, 4,	L*, H*, L+H*, L*+H, H+!H*, H+L*	
		2r (rhythm mismatch), 2t (tone mismatch)	L-, H-, !H-	ip
			%, L%, H%, ^H%	IP
			!, ^ (both for H pitch accent), <, >	
Greek	Prosodic word (=phonetic transcription)	0, 1, 2 (ip), 3 (IP), s (sandhi), m (mismatch)	L*, H*, L+H*, L*+H, H*+L	
			L-, H-, !H-	ip
			L%, H%, !H%	IP
			! (for *), <, >, w (for L* undershoot)	
Italian (Neapolitan)		0, 1, 2, 3, 4	L*, H*, L+H*	
			L*n, L+H*n, L*+Hn, H+L*n	ip
			H(*)L-	ip
			L%	IP
			! (for *), n (for nuclear pitch accent)	
Serbo-Croatian	Glosses	0, 1, 2 (IP), m (mismatch)	L*+H, H*+L	Wd
			%L, %H	Wd
			Ø-, LH-, L%, H%, HL%	IP
			>, # (pitch range rising)	
Japanese	Finality	0, 1, 2 (AP), 3 (IP), finality, m (mismatch)	H*+L	AP
			H-, L%, %L	AP
			H%, LH%, HL%	IP
			<, >, w (for L% or %L undershoot)	

TABLE 16.1 (*Continued*)

Lang.	Types of tiers—extra only	Types of break indices (BI)	Types of tones on the tones tier	Prosodic units
Korean	Phonol tone,	0, 1, 2 (AP),	L, H, +H, L+, Ha, La, LHa	AP
	Phonetic tone	3 (IP),	L%, H%, LH%, HL%, LHL%,	IP
		m (mismatch)	HLH%, LHLH%, HLHL%, LHLHL% <, >	
Mandarin	Romanzi, Syll,[a] Stress, Sandhi, Code	0, 1, 2 (minor grp), 3 (major grp), 4 (breath grp), 5 (prosodic grp)[b]	L%, H%, %reset, %q-raise, %e-prom, %compress	Breath group
Cantonese	Syllable, Foot[c]	0, 1, 2 (IP)	lexical tones (55, 33, 22, 335, 223, 221, 553)	Wd
			L%, H%, H:%, HL%, %, -%, %fi (frame initial boundary)	IP
Chickasaw	Phonetic transcription	0,	H^{λ}	Wd
		1 or 2 (AP),	LHHL, HL, LL, LHH	AP
		3 (IP),	H*, !H*	IP
		m (mismatch)	L%, H%, HL% <	IP
BGW	Gloss (Syll)	0, 1, 2,	H*, H*<, ^H*, L+H*	
		3 (IP),	Lp	PhP
		4 (Utt)	L%, H%, LH%, %L, %H	IP
			! (for * tones)	

[a] Unlike Cantonese ToBI, four lexical tones (55, 51, 32, 124) are labelled in the Romanzi (Romanization) tier in Mandarin.

[b] The authors note that a large percentage of inter-transcriber disagreements involved confusion between BI 4 and BI 5. They suggest that these two labels may be collapsed later.

[c] In Cantonese, the Syllable tier tags an alphabetic transliteration for every syllable. It is similar to Mandarin ToBI's Romanzi tier (Mandarin ToBI's Syllable tier tags phonological syllables). The Foot tier tags syllable fusion and the domain of emphasis. It does not function as a prosodic unit and differs from the foot in stress languages. Cantonese ToBI recommends three more tiers for some sites: Phones, Sociolinguistic variables, and Code.

Serbo-Croatian (SCToBI), Tokyo Japanese (J_ToBI), Seoul Korean (K-ToBI), Mandarin (M_ToBI), Cantonese (CToBI), Chickasaw, a Western Muskogean American Indian language (Ch-ToBI), and Bininj Gun-wok, an indigenous Australian language, also known as Mayali (BGW ToBI).

To compare the prosodic systems of these languages, their ToBI systems are compared in four categories: types of tiers, types of break indices, types of tones on the Tones tier, and types of prosodic units defined by intonation (not by juncture which is represented by break indices).

For the Types of Tiers, only the tiers other than the four original tiers (Words, Tones, Break Indices, Miscellaneous) proposed in American English ToBI are provided. This is because all the ToBI systems described in this book have these four tiers and only some have additional tiers, reflecting the different prosodic systems of the languages.

For the types of break indices, the numerical number is given when the meaning of the index is the same as or similar to that of the English break index (i.e., 0 for a weakened word boundary, 1 for a phrase-medial word boundary, 3 for a minor phrase boundary such as an Intermediate Phrase, 4 for a major phrase boundary such as an Intonation Phrase, and 2 for mismatch). When the meaning of any break index is different from this default meaning, the meaning (i.e., the name of a prosodic unit marked by the break index) is given in parentheses right after the break index (e.g. 2(IP) in Serbo-Croatian means the break index 2 marks the end of an Intonation Phrase). Diacritics for the break indices such as 'p' (disfluent pause), '–' (ambiguity between two indices), and '?' (uncertainty) are not included because most ToBI systems include these labels and these labels have the same meaning. However, diacritics specific to a certain ToBI system, such as 's' in Greek, are included with the interpretation in parentheses (e.g. s (sandhi)).

For the types of tones on the Tones tier, tones in each row have the same function and belong to the same prosodic unit. That is, they are either lexical tones (e.g. Chinese tone and Chickasaw lexical pitch accent), the head of a prosodic unit such as pitch accent (marked by *), or the boundary tone marking the edge of a prosodic unit such as an Accentual Phrase (AP), an Intermediate Phrase (ip), or an Intonation Phrase (IP). For Mandarin and Cantonese, tones for marking pitch range information (e.g. %reset, %q-raise) are grouped together with the boundary tone of the highest prosodic group. The name of the prosodic unit corresponding to each tone type is given in the last column, Prosodic Units, on the same row. The prosodic unit corresponding to the postlexical pitch accent is not given because the domain of the postlexical pitch accent is not known (cf. see Beckman and Edwards 1990 for a discussion of English pitch accent domain). Diacritics were added to the tone type if they are used as a part of a distinctive tone type as in English H+!H* or German ^H%. However, if diacritics were used with more than one tone type, the symbol is given in the last row within the tones column for

each language (e.g. '! (for H pitch accent)' or '! (for *)', meaning downstep can happen to all H pitch accent tones).

The summary table shows that all languages have prosodic units above the word. Except for the two tone languages (Mandarin and Cantonese), all have two prosodic units marked by intonation. Though these languages differ in the number of prosodic units above the Word, ranging from one (Serbo-Croatian and Cantonese) to four (Mandarin), most of them have two prosodic units above the Word: the largest phrase, often called Intonation Phrase, and a smaller phrase, called by various names such as Intermediate Phrase, Accentual Phrase, or Phonological Phrase. The prosodic units higher than the Word are often identified by intonation as well as by the degree of juncture, as implied by the fact that a break index and a boundary (or phrasal) tone mark the same prosodic unit. However, this mapping is not perfect in natural speech as indicated by the use of a break index (BI 2 or a diacritic 'm') for mismatch cases in most ToBI analyses. Currently, the two ToBI analyses of tone languages do not include any mismatch labels probably because tonally defined prosodic units smaller than an IP or Breath Group do not exist or have not been identified yet.

The ToBI analyses of languages also show that languages differ in the types of tones (lexical tone, pitch accent, or boundary tone) and tonal inventories they have. The tonal types can tell us whether a language is a tone language, an accent language, or whether it has no lexical specification of prosody. The current version of ToBI systems proposed for contour tone languages (Mandarin, Ch. 9, and Cantonese, Ch. 10) shows that they do not have pitch accent (a * tone) and that they have smaller 'intonational' tone inventories compared to other languages.[3] They only have a few boundary tones at the edge of the largest prosodic unit and have some labels marking pitch range specifications. The language with no lexical specification of prosody (Korean) does not have a * tone, either, but has the largest variety of boundary tones. The types of tones, however, cannot distinguish stress-accent languages from lexical pitch-accent languages. This is because the AM model does not specify whether pitch accent is a lexical property or a postlexical property. A major difference between these languages, based on the summary table, is in the number of tonal inventories. Stress-accent languages have multiple types of pitch accent while lexical pitch-accent languages have only one or two types of pitch accent. Among the languages which include pitch accent, H* is the

[3] It could be possible to propose a postlexical pitch accent in Mandarin when a syllable receives sentence level stress. A ToBI system of African tone languages would be needed to get a better picture of the prosody of tone languages in general as well as the prosodic properties specific to Chinese-like contour tone languages.

most common and L* is the least common. For all these languages, high or rising tones (a sequence of L and H) are more common than falling tones, and single boundary tones (i.e., L%, H%) are more common than multiple boundary tones (e.g. LH%, LHL%).

For tiers, three types of tiers have been added in addition to the original four tiers. A first type is to represent language-specific prosodic features, a second type is to help labellers who are not native speakers of the language, and a third type is to label information needed to investigate the interaction between prosody and sub-areas of linguistics such as the role of prosody in syntax-phonology mapping, discourse structure, and dialect differences. The first type includes the syllable-related tiers such as the Syllable and the Stress tiers in a tone language (where each syllable carries a tone, and the degree of stress affects the realization of the underlying tone) and the Phonetic Tone tier in Korean (where some of the phrasal tones show variations in tonal categories, categorically affecting the tonal shape of the phrase, but have no distinctive function). The second type includes the English Gloss tier in some of the non-English languages (e.g. Serbo-Croatian, BGW). The third type includes the Phonetic Transcription tier where allophones are specified to study the domain of a segment sandhi (e.g. Greek, Chickasaw), the Code tier for a dialect specification (e.g. Mandarin) to study code switching and dialect typology, and a Finality tier for marking discourse finality (e.g. Japanese) to study the relationship between prosody and discourse structure.

Finally, the summary table shows that the Tones tier uses diacritics to cover both distinctive tonal targets and the phonetic realizations of the tonal targets. The diacritics for the local pitch range adjustments (e.g. !, ^, #) are distinctive in some languages (e.g. H+!H* in English and German, ^H* in BGW) but not in others (e.g. !H* in Greek and Chickasaw). Similarly, the diacritics for timing in the tone-text alignment (e.g. '<' for a peak delay and '>' for an early peak) are used to mark a distinctive tonal category in Bininj Gun-wok (H*<), but are used to mark a phonetic event in all others. This suggests that the function of diacritics in the ToBI system is specific to its variety. This issue will be further discussed in Section 16.4.

Though not represented in the summary table, languages also differ in their tone-text association and their alignment. For example, the H- phrase accent in Greek interrogatives is associated with a stressed syllable, if available, but not in English. The L*+H pitch accent in Greek is realized as a rising tone with the fo minimum before the stressed syllable and the fo maximum after the stressed syllable (Arvaniti *et al.* 1998), but the same pitch accent is realized in English as a rising tone with the fo minimum *during* the stressed syllable and fo maximum after the stressed syllable (Pierrehumbert and Steel

1989; Hirschberg and Ward 1992; Beckman and Ayers-Elam 1997). Similarly, the realization of the same prosodic unit also differs across languages. For example, a falling tone (HL%) can mark the end of an Intonation Phrase in Japanese and Korean, but not in English, German, or Greek. The end of an Intonation Phrase is significantly lengthened in English and most other languages, but it is not always lengthened in Japanese. A High tone can mark the end of an Accentual Phrase in Korean, but not in Japanese.

Furthermore, languages, and even the dialects within the same language, can differ in their mapping between a tone and the meaning. Though there are cases where a tone-meaning mapping is fairly consistant across languages or dialects (e.g. a low boundary tone for a statement, a high boundary tone for a yes-no question), the mapping is specific to a language and a dialect. A basic meaning of a certain tone type (e.g. boundary tones, focus tone) is given in each ToBI system, but given that the relationship between a tone and the meaning is not one-to-one, but many-to-many even in a single dialect, it would be premature to include tone-meaning relations in prosodic typology. In order to find out tone-meaning mappings across languages, further research in meaning, including pragmatic and sociolinguistic meaning, should be carried out first.

16.3. PROSODIC TYPOLOGY

In this section, a model of prosodic typology is proposed based on the prosodies of various languages which have been analysed in the framework of the AM model of intonational phonology. Comparisons across different ToBI models and AM models of intonation show similarities and differences in prosody across languages. As mentioned earlier, the types of tones and the tonal inventories proposed in these models distinguish languages in terms of their prosodic structure and, to some degree, the properties of their lexical prosody (whether a language has stress, lexical pitch accent, tone, or none of these). However, the AM model does not directly specify information about lexical prosody or timing units. For example, it does not distinguish languages which have lexical pitch accent only (e.g. Japanese) from languages which have both lexical pitch-accent and stress (e.g. Serbo-Croatian). It also does not distinguish languages which have similar prosodic structures and tonal types but differ in their timing (e.g. stress-timed English vs. syllable-timed Italian). This suggests that prosodic typology would not be complete unless we include prosodic events below the word as well as those above the word (see Hirst and Di Cristo 1998 for a similar view). When combining

lexical and postlexical prosodic events, there seem to be two main aspects of variation which define prosodic typology. These are the prominence and the rhythmic/prosodic pattern of an utterance.

Prominence, as a category in prosodic typology, concerns both how the prominence of a lexical item in the language is realized prosodically (lexical prosody) and how the prominence relation among the words is realized postlexically. Although there are no agreed upon criteria for categories of lexical prosody (see Beckman 1986; van der Hulst 1999; Fox 2000 for various proposals on types of lexical prosody and their definitions), three categories of lexical prosody—tone, stress-accent, and lexical pitch-accent—are chosen in this chapter based on the phonetic realization of prosodic features, their function, and the relation to intonation (following Beckman 1986; Ladd 1996; Cruttenden 1997). A language is categorized to have a 'tone' feature if the language has 'prescribed pitches for syllables or sequences of pitches for morphemes or words' (Cruttenden 1997: 8–9), i.e., pitches having paradigmatic contrast. A language is categorized to have a 'stress-accent' feature if a certain syllable in a word is more prominent than other syllables by duration and/or amplitude, showing syntagmatic contrast. The syllable has no lexical specification of pitch but can be realized with a certain pitch pattern determined by intonation. Finally, a language is categorized to have a 'lexical pitch-accent' feature if a certain, not every, syllable of a word has lexical specification of pitch, showing syntagmatic contrast, but does not exhibit phonetic 'stress' in the sense of Beckman (1986). Therefore, unlike the stress-accented syllable, the pitch pattern of a lexical pitch-accented syllable, if realized, is fairly independent of intonation.

Data suggest that there are two ways of prominence realization at a postlexical level: culminatively by marking the head of a prosodic unit and demarcatively by marking the edge of a prosodic unit (Hyman 1978; Beckman 1986; Beckman and Edwards 1990; Ladd 1996; Venditti *et al.* 1996). Prominence is realized culminatively at a postlexical level when a syllable or word becomes prominent through a local manipulation of suprasegmental features such as pitch, duration, and/or amplitude. This is represented as the postlexical pitch accent (marked as *) in the AM model. The realization of postlexical pitch accent depends on what suprasegmental features the language employs for the realization of the lexical prosody. If the postlexical pitch accent is from the lexical pitch-accent as in Japanese, it does not change the duration or amplitude of the syllable (Beckman 1986; Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988). On the other hand, if the postlexical pitch accent is from the stress-accent as in English, it does change both the duration and amplitude of the syllable. Acoustic correlates of

postlexical pitch accent are language specific, and this difference will not be captured in a model of prosodic typology where languages are compared based on phonological categories.

Prominence can also be realized demarcatively at a postlexical level when the prominent word comes at a certain location in a prosodic unit (e.g. the beginning or the end), and a phrasal tone, mostly the same tone type, marks the edge of the prosodic unit. Examples of demarcative prominence are the Prosodic Word boundary tone in Serbo-Croatian (%L), the Phonological Phrase boundary tone in Bininj Gun-wok (Lp), and the Accentual Phrase boundary tone in Japanese (L%) and in Korean (Ha) (see Table 16.2 for more data). That is, the function of postlexical pitch accent in English and other West Germanic languages is preformed by prosodic phrasing in 'edge' prominence languages. For example, the function of postlexical pitch accent in English and other Germanic languages (such as marking focus or disambiguating an ambiguous string) is performed by placing words in the same or different prosodic units, i.e., prosodic phrasing, in Japanese and Korean (Ladd, 1996; Venditti *et al.* 1996; Venditti 2000). Venditti *et al.* showed that the function of pitch accent and deaccenting in English is delivered by phrasing and dephrasing in Korean and Japanese. In English, contrastive focus is realized by L+H* pitch accent followed by deaccenting, while, in Japanese and Korean, contrastive focus is realized by inserting a prosodic boundary before or after the focused word and dephrasing postfocused items (see Jun 1996, 2003*a*; Jun and Lee 1998; Ueyama and Jun 1998). Ladd (1996: 196–7) agrees with this view and notes that sentence accentuation can be seen as one manifestation of prosodic structure, and deaccenting or dephrasing are just different surface symptoms of the same deep structural effects.

Another category affecting prosodic typology is the rhythmic/prosodic pattern of an utterance. The rhythmic pattern refers to the timing unit which is smaller than the word, and the prosodic pattern refers to the prosodic unit above the word. These patterns are combined together as a group since a timing unit can be represented as a prosodic unit, and a prosodic unit has a rhythmic nature, as mentioned in Section 16.1. In stress-timed languages, rhythm is perceived through sequences of prominent and non-prominent syllables, and the prosodic unit including the prominent and non-prominent syllables is a Foot. Rhythm is also perceived when a sequence of syllables or words shows a repeating tone pattern (e.g. LLH - LLH - LLH; LHL - LHL - LHL). The grouping of syllables/words marked by a tonal pattern is a postlexical prosodic unit such as an Accentual Phrase or a Phonological Phrase. Rhythmicity will increase if the tonal pattern is the same. The rhythmic property of a postlexical prosodic unit can also be derived from the lengthening of a syllable at the end of the unit. A similar

view is indicated by Cruttenden (1997: 7–21). He calls a Foot a 'rhythm group' and describes the boundary of an intonation-group as a break in the rhythm.

The rhythmic patterns derived by the postlexical prosodic unit, which I call a 'macro' rhythmic unit, would vary more than those derived by foot, syllable, or mora, which I call a 'micro' rhythmic unit. As mentioned in Section 16.1, the rhythmic constraints on the micro rhythmic patterns also seem to apply to the macro rhythmic patterns, though less strictly. That is, prosodic units at the same level tend to have a similar size and the boundary of each prosodic unit tends to occur at a regular interval. Evidence can be found in the size constraint on a prosodic unit. Studies on prosodic phonology of various languages (e.g. Nespor and Vogel 1986; Delais-Roussarie 1995; Jun 1996, 2003*b*; Selkirk 2000) have shown that a phonological phrase is constrained by the size of the unit (e.g. constraints on the minimum and maximum size of prosodic constituents).

The constraint on the size of a prosodic unit also applies to a larger prosodic unit. Even though the rhythmic nature of the unit would be weakened as the size of the postlexical prosodic unit increases, from Prosodic Word to Intonation Phrase, we can still feel a broad sense of rhythm from the distribution of Intonation Phrases. For example, when a sentence is short, it is produced in one Intonation Phrase, but when the sentence gets longer, it tends to have two Intonation Phrases with the boundary coming in the middle of the sentence (e.g. Gee and Grosjean 1983, Ferreira 1993). Similarly, in many languages, when a relative clause (RC) is short, speakers tend to produce no major prosodic break between the relative clause and the head noun, but when the RC is long and heavy, they tend to produce a major prosodic boundary between the head noun and the RC (e.g. Lovrić *et al.* 2000; Quinn *et al.* 2000; Jun and Koike 2003).

An informal observation on the size of Accentual Phrase (AP, the same level of the Phonological Phrase in Selkirk's hierarchy) and Intonation Phrase (IP) across four languages based on the short story, *The North Wind and the Sun*, shows that APs tend to have a similar size within each language and are similar in size even across languages. In Japanese, most APs have 4–5 syllables (Ueyama 1998), and in Korean and French, most APs have 3–4 syllables or 1.2 content words (Jun and Fougeron 2000). The size of an IP in each language varies more in general but becomes similar when short IPs consisting of a connective or an exclamation (e.g. *Alors*, 'Therefore', in French) are excluded. The average number of syllables within an IP ranges from 7–10 syllables in English, French, and Japanese to 12–15 syllables in Korean. The syllable count or the word count, however, could misrepresent the phenomenon because it would depend on the structure of a syllable and whether a language has

function words, and if so, how it treats the function words in forming a prosodic unit. Another, perhaps a better, way to measure the rhythmic pattern of an IP would be the duration of the phrase. It is reported that the average duration of an IP is 1.5 seconds in more than one language (data from two subsamples of the Verbmobile-database and MARSEC corpus for German and English, reported by Batliner 2001; also from an interview with a French novel writer, reported in Fagyal 1995). More data should be compared to confirm the rhythmic nature of a macro prosodic unit. When doing so, we should control the type of discourse and genre and/or the speech style because these factors would affect the duration of the prosodic units.

Languages seem to differ in how an utterance is rhythmically and prosodically organized. Based on the AM model of various languages, some languages have only one prosodic unit above the word (e.g. Serbo-Croatian) while others have three (e.g. Bininj Gun-wok, Farsi). Though the rhythmic patterns and prominence realizations at the lexical level are closely related to those at the postlexical level, the postlexical prosodic patterns are not fully predictable from the lexical properties. Table 16.2 shows the prosodic features of twenty-one languages specified in two categories—the Prominence and the Rhythmic/Prosodic Unit—with each category being divided into two levels, lexical and postlexical. Lexical prominence is further divided into three categories reflecting the three features of word prosody. This arrangement enables us to represent languages which have more than one way of marking lexical prominence (e.g. both stress and lexical pitch-accent or both tone and stress). 'LPA' under the lexical prominence column stands for 'lexical pitch-accent'. Postlexical prominence is divided into two categories reflecting the two ways of prominence realization: marking the head (culminative) and marking the edge (demarcative). Head marking is achieved by postlexical or intonational pitch accent, and edge marking is achieved by the phrasal tones marking the edge of a prosodic unit (similar to or slightly higher than the Word level). When the category of prominence for a given language is not agreed upon among the researchers, '(x)' is given.

Similarly, the category 'Rhythmic/Prosodic unit' is divided into two categories, lexical and postlexical. The lexical rhythmic unit is further divided into three categories reflecting the three types of timing unit: mora-timed, syllable-timed, and stress-timed (a 'Foot' is given as the rhythmic unit of stress-timing). Though this three-way division is not as categorical as the table implies, each language is, at the moment, categorized as belonging to only one timing unit except for Italian. For languages for which there is no published data about the timing unit, a category is chosen based on consultation with native speakers and researchers working on the language

TABLE 16.2. Prosodic typology based on prominence and the rhythmic/prosodic unit

Prosody	**Prominence**					**Rhythmic/prosodic unit**					
	Lexical			*Postlexical*		*Lexical*			*Postlexical*		
Language	tone	stress	LPA	head	edge	mora	syll	foot	AP	ip	IP
English		x		x				x		x	x
German		x		x				x		x[a]	x
Dutch		x		x				x			x+1
Greek		x		x				x		x	x
Italian		x		x			x	(x)[b]		x	x
Spanish		x		x			x			(x)[c]	x
Portuguese		x		x			x				x
Arabic		x		x				x		x	x
Farsi		x		x	x		x		x	x	x
BGW		x		x	x			x	(x)[d]		x+1
Swedish		x	x	x				x			x
Sb-Croat.		x	x	x	x			x			x
Chickasaw		x	x	x	x			x	x[e]		x
Japanese			x	x	x	x			x	(x)[f]	x
Basque			x	x	x		x		x	x	x
French				x	x		x		x	(x)[g]	x
Bengali				x[h]	x		x		x		x
Korean					x		x		x		x
Mandarin	x	x		(x)[i]			x			x	x
Cantonese	x						x				x
Kinande	x				x		x		x		x

Notes: Spanish data is from Nibert 1999; Prieto 1999; Sosa 1999; Beckman *et al.* 2002. Portuguese is limited to European Portuguese, and the data is from Frota 2000; Frota *et al.* 2002. Arabic is limited to Lebanese Arabic, and the data is from Chahal 2001. Farsi data is from Jun *et al.* 2003. Basque is limited to the Lekeito dialect, and the data is from Elordieta 1997, 1998; Jun and Elordieta 1997; Elordieta and Hualde 2001. French data is from Jun and Fougeron 1995, 2000, 2002. Bengali data is from Hayes and Lahiri 1991; Ladd 1996; Lahiri and Fitzpatrick-Cole 1999. Kinande data is from Hyman 1990.

[a] Féry (1993) and Grabe (1998) adopt Gussenhoven's (1984) model and do not propose an ip for German. However, as noted in Chapter 3, the differences are of a theoretical rather than a typological nature.
[b] Chapter 13 (Italian) explains that not all dialects are syllable-timed. Some Southern dialects tend more towards stress-timing.
[c] Sosa (1999) and Nibert (2000) propose an ip in Spanish but Beckman *et al.* (2002) do not.
[d] The smallest prosodic unit larger than a word in BGW is a Phonological Phrase (PhP), mostly containing one morphosyntactic word. The authors (Chapter 12) claim that not all dialects of BGW show this unit.
[e] The Chickasaw AP can be larger or smaller than a prosodic word.
[f] Beckman and Pierrehumbert (1986) and Pierrehumbert and Beckman (1988) proposed an ip for Japanese.
[g] Jun and Fougeron (2000) proposed an ip for French with a note that they need more data to confirm the unit.
[h] Lexical accent in Bengali is realized as pitch accent at the postlexical level. It does not show a stress feature as in English (Ladd 1996, 2001). However, Lahiri and Fitzpatrick-Cole (1999) found that focus clitics are lexically specified with high tone.
[i] Postlexical pitch accent in Mandarin was not proposed in Chapter 9, but a more recent proposal includes stress-driven pitch accent (Janis Fon, pers. comm. June 2003).

concerned. As can be seen in the table, the classification of the timing unit does not seem to affect the classification of other prosodic categories. Further research on the timing unit may change the category of the rhythmic unit and even the arrangement of this column.

Under the postlexical column of the Rhythmic/Prosodic unit, 'AP' is a cover term referring to the smallest rhythmic/prosodic unit above the word, whose boundary is marked suprasegmentally. This unit contains in general one morphosyntactic content word. 'ip' is also a cover term referring to a prosodic unit larger than a Word and smaller than an Intonation Phrase (e.g. the Intermediate Phrase of English, German, and Greek). This unit contains in general more than one content word and would be larger than an Accentual Phrase-like prosodic unit if a language has both (e.g. Japanese intonation proposed in Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988). '(x)' in the AP or ip column means that either the existence of the category is not fully confirmed (e.g. 'ip' in Spanish) or not found consistently across all dialects of the language (e.g. 'ip' in Bininj Gun-wok). Finally, 'IP' refers to the largest prosodic unit marked by intonation. 'x+1' under the IP column means that one more prosodic unit above the Intonation Phrase (e.g. the Utterance) has been proposed for these languages. This unit is in general not marked by intonation but by the degree of finality through phrase final lengthening and pause.

The languages are arranged in four groups following the 'traditional' characteristics of their word prosody: stress-accent languages, lexical pitch-accent languages, non-stress and non-lexical pitch-accent languages, and tone languages. In assigning the prosodic features to each language, the models described in the previous chapters are used as a reference, but if another phonological model of the same language assumes different prosodic features or structures, a superscript is given next to (x) in the table and a note is given below the table caption. For languages whose models are not described in the book, a reference is given below the table caption. It should be noted that, as in all studies on typology based on phonological categories, the categorization would be somewhat different if one adopted different models of prosody for a certain language. But, it is believed that the prosodic features of a language described in this table would not be dramatically different across different models if those models assume the same principles specified in the AM model of intonational phonology. However, the names of certain prosodic units and the number of prosodic units may differ across the models.

Table 16.2 shows that characterizing the prosodic properties of languages based on prominence and rhythmic/prosodic units allows us to observe generalizations of the relationship between the types of prominence and the types of rhythmic/prosodic unit, both at the lexical and the postlexical level. Generalizations found from the table are listed below.

Generalizations on Prosodic Prominence and the Rhythmic/Prosodic Unit.

(a) All languages have at least one prosodic unit above the word.
(b) In stress languages, the prominence of a word is always marked by postlexical pitch accent (i.e., marking the head of the word), but not often by marking the edge of the word.
(c) Most of the lexical pitch-accent languages mark the prominence of the word in two ways at the postlexical level: culminatively by marking the head of the word and demarcatively by marking the edge of the word.
(d) Languages that do not have any feature of lexical prosody mark the prominence of the word demarcatively at the postlexical level.
(e) Non-stress languages can have postlexical pitch accent.
(f) The number and type of rhythmic/prosodic units at the postlexical level are not predictable from the lexical rhythmic unit of the language, nor from the type of lexical prominence.
(g) There is no direct relationship between the type of lexical prominence and the type of lexical rhythmic unit. Also, there is no relationship between the type of postlexical prominence and the type of postlexical rhythmic/prosodic unit. However, the edge marking of the postlexical prominence is predictable from the AP category in the postlexical rhythmic/prosodic unit.
(h) As 'b–d' above imply, the type of postlexical prominence is partially predictable from the type of lexical prominence.

The prosodic categorization in Table 16.2 also captures the prosodic differences and similarities across languages. As described in Beckman and Pierrehumbert (1986), English, a stress-accent language, and Japanese, a pitch-accent language, are both specified as having postlexical pitch accent and having two postlexical prosodic units. But, the table shows that, in addition to the lexical prosody and lexical rhythm, these languages differ by the way the prominence is realized postlexically: English uses head marking but Japanese uses both head and edge marking. The Accentual Phrase in Japanese tonally marks a word boundary most of the time (i.e., about 68 per cent of the rising tones at the beginning of an AP mark the beginning of a word (Warner and Arai 2001)), and the boundary tone is always Low (L% or wL%); but the Intermediate Phrase in English does not—the Intermediate Phrase final boundary tone can be either High or Low and is not realized on a certain syllable of a word but realized over multiple syllables between the nuclear pitch accented syllable and the last syllable of the Intermediate Phrase, i.e., post-nuclear tail (see Chapter 2).[4] The categorization also captures the

[4] An Intermediate Phrase (ip) also seems to be larger than an AP-like type when compared in terms of the number of syllables and words. When examining 20 Intermediate Phrases from the BU FM

similarities and differences between French and Korean. In both languages, the prominence of the word is not marked at the lexical level but marked at the postlexical level by the rising tone at the edge of an Accentual Phrase (as mentioned earlier, the AP in these languages are about 3–5 syllables long on average and in general include one content word; Jun and Fougeron 2000; Schafer and Jun 2002). The difference between these languages lies in the fact that in French the AP final full syllable carries postlexical pitch accent while in Korean no syllable carries postlexical pitch accent. The table also shows that Japanese and Basque are similar in many prosodic features but their lexical rhythmic unit is different.

The current categorization, however, cannot capture the differences between stress languages that differ in the frequency and the type of postlexical pitch accent. In some stress languages like Spanish and Greek, pre-nuclear pitch accent occurs on almost all content words, and the type of pitch accent is basically the same (e.g. L*+H for Greek) except for a few cases; while in other stress languages like English and German, pitch accent does not occur on every content word and the type varies more (see Dainora 2001 for the frequency and the type of pitch accent in English). In the former case where pitch accent occurs at a regular interval (i.e., almost every content word) with a similar type of pitch accent, each of the accents would provide a cue for a word boundary, functioning similarly to the Word boundary tone in Serbo-Croatian or the Accentual Phrase boundary tone in Korean. Since the prosodic features of a language described in the AM model of intonation focus on the prosodic structure defined by distinctive pitch events and the degree of juncture, the perceptual equivalence of word segmentation, whether it is marked by the head tone or by the edge tone of the unit, is not captured in the model. This model also does not include the category 'quantity' or length (e.g. long vowels, long consonants, heavy syllable), unlike previous discussions of prosodic typology (e.g. Hirst and Di Cristo 1998; van der Hulst 1999; Fox 2000). This is because the length distinction, if it affects the tone type, is incorporated into the tone type (e.g. Serbo-Croatian's two types of rising tone and falling tone, Japanese L vs. wL). In ToBI, the distinction in quantity is given in the Word tier as part of the lexical information.

In addition, the current typology does not include how languages differ or are similar in terms of the melodic category in Ladd's (2001) intonational typology (i.e., the relation between tune and meaning/function and the

News corpus, produced by three radio news speakers, Shattuck-Hufnagel found that an Intermediate Phrase included two content words on average, ranging from one to four content words (Stefanie Shattuck-Hufnagel, pers. comm. June 2003). This agrees with Ueyama's (1998) finding. Based on a short story (*The North Wind and the Sun*) read by two speakers, she found 5–6 syllables per ip (cf. 3–4 syllables per AP).

realization of tunes across languages). As mentioned in Section 16.2, the relation between a tone and its meaning, or its function, is not fully studied and the currently available analyses do not provide enough data to compare across languages. Ladd's (2001) other two categories of intonational typology (accentual and prosodic) are also not captured by the current prosodic categories because Ladd's typology is concerned with the relation between prosody and morphosyntax or phonology, while the current typology is concerned with the shape (the tonal categories and the structure) of prosody itself and its realizations. As will be described in the next section, the AM model of intonation will be used as a tool to investigate the interaction between prosodic features and sub-areas of linguistics, thus helping to explore the accentual and prosodic typology in Ladd (2001).

In sum, the prosodic categorization given in Table 16.2 was chosen to capture the prosodic differences and similarities across languages described in the AM model of intonational phonology. These prosodic categories seem to be useful and significant in determining the prosodic typology of the languages described so far. In order to verify the proposed model of prosodic typology, more languages should be analysed in the AM model of intonational phonology.

16.4. ToBI TRANSCRIPTION: ITS FLEXIBILITY AND EXTENSIONS

As shown in Section 16.2, ToBI systems differ across languages, reflecting the different prosodic systems of languages. The conventions and principles assumed in the ToBI system (see Chapter 2), which are based on American English intonation, have been applied to various languages whose prosodies differ substantially. This suggests that the ToBI system is flexible without losing the integrity of the system. It is flexible to include other stress languages because the tonal inventories on the Tones tier and the break indices on the Break Index tier can change, reflecting the tonal patterns and the prosodic structure of the target language (e.g. H+L* is found only in Greek; ^H% is found only in German; the break index 3 is a juncture for an Intonation Phrase level in Bininj Gun-wok, but is a juncture for an Intermediate Phrase in German and Greek). It is also flexible to include lexical pitch-accent languages because the lexical tonal event is represented on the Tones tier in the same way as the postlexical pitch accent (e.g. H*+L). It is flexible enough to include a language which has lexical pitch-accent as well as stress-accent (e.g. Chickasaw) because it allows us to add a diacritic on a tone to distinguish

lexical pitch accent from postlexical pitch accent and to label them on the same tone tier (i.e., H^{λ} for lexical pitch-accent and H^{*} for postlexical pitch accent). This is an extension of the usage of the diacritic to mark the affiliation of the boundary tone to a different prosodic unit (e.g. H- for an Intermediate Phrase and H% for an Intonation Phrase). It is also flexible enough to include tone languages, with or without stress, because it allows labellers to create new tiers (e.g. Syllable, Stress) to tag information unique to the prosody of the language. Finally, it is also flexible enough to include a language which has no stress and no lexical pitch-accent but only has a phrasal tone and a boundary tone. This is because it allows us to have the Tones tier with no starred tone (*T) if no tone is associated with a syllable, either metrically strong or marked in the lexicon, and it allows us to expand the Tones tier to distinguish the underlying, phonological tone from the surface, phonetic tone.

One of the principles assumed in the ToBI system is that we should transcribe *distinctive* prosodic information. As mentioned in Chapters 1 and 2, the prosodic model assumed in ToBI is a phonological model, not a phonetic one. The question remains what the distinctive properties of prosody are. An apparent answer would be that a prosodic feature is distinctive if modifying that feature affects the semantic and pragmatic meaning of an utterance. But, a prosodic feature can also change the sociolinguistic meaning such as dialect variation or speech style while keeping the semantic and pragmatic meaning the same. To a research community working on the sociolinguistic meaning of prosody, the prosodic features changing the dialect identity or speech style must be 'distinctive' and should be transcribed. This suggests that there are different levels of distinctiveness, and which level of distinctiveness one should include in the transcription would depend on the interests of the research group. Similarly, distinctiveness in meaning can be realized in more than one way (cf. systemic differences in Ladd's (1996) typology). It could be done by a distinct tonal shape such as High or Rising (LH), or by pitch range manipulations such as pitch range expansion and reduction related to focus. In other words, pitch range manipulation can be phonologized to mark certain semantic or pragmatic meanings (cf. Ladd 1996; Jun and Lee 1998). Mandarin ToBI tags pitch range information on the Tones tier (e.g. %q-raise, %compress), and this is an extension of the conventional ToBI transcription because the tonal inventories on the Tones tier have always been a tone symbol (H or L and its combinations), possibly combined with some diacritics (e.g. *, -, %, !, ^, >).

Related to this is the issue of distinctive tone levels in ToBI. In the AM model of intonation, intonation contours are represented by only two tone

levels, High and Low. Tonal levels other than H and L were explained by phonological rules and phonetic realization rules such as downstep and upstep. However, in the ToBI models, downstepped High and upstepped High tones are explicitly marked by diacritics (! and ^) and they help to avoid making a transcription too abstract. This made it possible to represent four tone levels which are paradigmatically contrastive (e.g. H+!H* vs. H+L*, H% vs. ^H% in German; H% vs. !H% in Greek).

Another extension of the Tones tier transcription can be found in Mandarin and Cantonese ToBI. As contour tone languages, Mandarin and Cantonese have multiple levels of pitch change within one syllable. So, instead of using H and L, they use numerical numbers (e.g. 221, 55, 32) for the transcription of lexical tones. Cantonese tags the numerical tone on the Tones tier, but in Mandarin ToBI, the lexical tones are labelled on the Romanization tier and 'true' intonational tones are labelled on the Tones tier. The authors of Mandarin ToBI maintain that they did not include the lexical tone on the Tones tier because the realization of the underlying tone varies due to the degree of stress and tone sandhi, and the distinction between the tonal reduction and tone sandhi was not easy to resolve. For these two languages, it might be possible to add the lexical tone information on the Words tier or Romanization tier and mark the surface realizations (postlexical level) of the lexical tones on the Tones tier together with the 'true' intonational tones.

Another principle of the ToBI system is that the transcription conventions should be efficient. This means that we should not transcribe predictable or redundant information, either predictable from the intonation system (e.g. English nuclear pitch accent is not transcribed separately from pitch accent since it is predictable from the location of pitch accent within an Intermediate Phrase, while in Italian nuclear pitch accent is not predictable from the location of the pitch accent, and Italian ToBI labellers tag the nuclear pitch accent with a diacritic 'n' after the tone, e.g. L+H*n) or extractable from an on-line dictionary (e.g. the location of lexical stress in English) or from other tiers (e.g. the break index 4 from L% or H% on the Tones tier or vice versa).[5] This suggests that we should not label the Accentual Phrase initial tones in Korean (predictable from the segment types in the Words tier) nor the Japanese Accentual Phrase initial L or wL tone (predictable from the syllable structure from the Words tier). Furthermore, we should not label Japanese pitch accent but only label the *location* of pitch accent at the postlexical level because there is only one type of pitch accent (i.e., H*+L)

[5] In the currently available ToBI models, German ToBI is the only one which does not label the unambiguous break indices 3 and 4 to avoid the problem of redundancy.

and this will be predictable from the intonation system of Japanese. Similarly, if a language has a default pre-nuclear pitch accent occurring in almost all content words (e.g. Greek, Spanish), we may not need to label the pitch accent but only label non-default pitch accent.

However, not transcribing a certain value or a label due to redundancy or predictability would not necessarily increase the efficiency of the transcription. Skipping a certain label would require labellers to check other tiers or would require them to spend some time deciding if the labels were predictable or not, and this could confuse the labellers and increase transcription time. Furthermore, since all ToBI systems are 'an ongoing research program rather than a set of "rules" cast in stone' (Chapter 2), researchers often want to mark all prosodic information, whether it be redundant or not, to study the relation between prosodic features on different tiers and their phonetic realizations. Labelling information on all tiers despite the redundancy would also be useful when one is interested in examining only one type of prosodic information and not the others. For example, researchers who are interested in the degree of juncture between words relative to the morphosyntactic information of the utterance would need to look at the Break Index tier and the Words tier (or probably add a Syntax tier as described in the next paragraph), but not necessarily the Tones tier.

The ToBI systems described in this book show that a new tier can be added to describe the prosodic system specific to a language or a research community. Some tiers, like the Stress tier in Mandarin ToBI, are directly related to the realization of the tone, while others provide information about the speech (e.g. names of the dialect in the Codes tier) or information about the segment (the Phonetic Transcription tier in Chickasaw or the Prosodic Word tier in Greek). The Prosodic Word tier in Greek was added because the researchers were interested in studying the domain of segment sandhi, related to the prosodic structure. This tier will provide data for phonology-syntax mapping or prosodic phonology. All this suggests that a ToBI system is a tool with which to do research into all aspects of prosody and the interface between prosody and sub-areas of grammar. Researchers can add a Discourse tier (similar to the Japanese Finality tier) to tag a discourse structure and examine its relation to the pitch range and the degree of juncture, or add a Morphosyntax tier to study syntax-prosody relations. In addition, a tier can be split, as in the Tones tier in Korean ToBI being split into the Phonological tone tier and the Phonetic tone tier. This was done to study the distinctiveness of the phrasal tone, the realization of the underlying tone, and the tone-text alignment. Similarly, a Tone tier could also split into a Lexical Tone tier and a Postlexical Tone tier for a tone language or a language with a lexical

specification of pitch for a part of the lexicon (e.g. Chickasaw, Zapotec). This would show how lexical tones are modified by or interact with an intonational tone. Finally, researchers interested in speech synthesis or recognition could add an Implementation tier, providing an interpretation of the tonal categories, which would be useful as input into the synthesizer. Adding a new tier would not decrease the integrity of the system as long as the system is used to study different aspects of prosody or the role of prosody in linguistics through the transcription of tones and break indices.

The flexibility of the system, however, could be a weakness of the system because there is no restriction on the number or types of tiers or on the types of labels or diacritics that could be added. In fact, we have seen that different ToBI systems use diacritics differently. Some use them as phonological markers and some as phonetic markers (e.g. !, <). But, this is probably unavoidable because a less developed ToBI system will include more symbols until the phonological status of the symbol is resolved by the examination of more data. The limits of the extension may also be resolved by developing more ToBI systems and using them in various types of research.

16.5. CONCLUSION

In this chapter, I have given a summary of the eleven ToBI systems described in this book. The types of tiers and break indices, types of the prosodic units, and the tonal inventories on the Tones tier reflect the prosodic structure and the intonation pattern of each language. Based on the thirteen languages described in this book and eight other languages described in the same theoretical framework, i.e., the AM model of intonational phonology, I have proposed a model of prosodic typology. The prosodic similarities and differences across the twenty-one languages are well captured by two prosodic categories, prominence and rhythmic/prosodic unit, with each category being further divided into lexical and postlexical levels.

I have also shown how flexible the ToBI system is in incorporating many languages which vary morphosyntactically and prosodically. I have discussed how the system has been extended by some of the ToBI models and have suggested possible additional extensions. To fully understand the universal and language specific characteristics of a prosodic system, the prosodies of more languages need to be described, and to verify the model of prosodic typology proposed in this chapter, more languages should be analysed in the framework of intonational phonology.

Finally, it is suggested that a ToBI system would be an excellent tool with which we could do research on prosodic interfaces including syntax-phonology mapping, semantic/pragmatic meaning, discourse structure, sentence processing, sociolinguistics, and speech technology. It is hoped that the next edition of this book will include a revised and improved version of each ToBI system described here and will include new ToBI systems of more languages.

REFERENCES

Abercrombie, D. (1967), *Elements of General Phonetics* (Edinburgh: Edinburgh University Press).

Arvaniti, A., Ladd, D. R., and Mennen, I. (1998), 'Stability of Tonal Alignment: the Case of Greek Prenuclear Accents', *Journal of Phonetics*, 26: 3–25.

Batliner, A. (2001), 'A Reply to "avg. length of intonational phrase"' in Prosody Discussion List (PROSODY@LIST.MSU.EDU), archive address: http://listserv.linguistlist.org/cgi-bin/wa?A2 = indo101b&L = prosody&D = 1&F = &S = &P = 88. January 2001, week2.

Beckman, M. E. (1982), 'Segment Duration and the "Mora" in Japanese', *Phonetica*, 39: 113–35.

—— (1986), *Stress and Non-Stress Accent* (Dordrecht: Foris).

—— (1996), 'The Parsing of Prosody', *Language and Cognitive Processes*, 11: 17–67.

——, and Ayers-Elam, G. (1997), 'Guidelines for ToBI labelling. Version 3.0, March 1997', ms (Ohio State University).

——, Diaz-Campos, E., McGory, J., and Morgan, T. (2002), 'Intonation across Spanish in the Tones and Break Indices Framework', *Probus*, 14: 9–36.

——, and Edwards, J. (1990), 'Lengthenings and Shortenings and the Nature of Prosodic Constituency', in J. Kingston and M. Beckman (eds.), *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech* (Cambridge: Cambridge University Press), 152–78.

——, and Pierrehumbert, J. (1986), 'Intonational Structure in Japanese and English', *Phonology Yearbook*, 3: 255–309.

Bloch, B. (1950), 'Studies in Colloquial Japanese IV: Phonemics', *Language*, 26: 86–125.

Chahal, D. (2001), 'Modeling the Intonation of Lebanese Arabic Using the Autosegmental-Metrical Framework: A Comparison with English', Ph.D. dissertation (University of Melbourne).

Cooper, W. E., and Eady, S. J. (1986), 'Metrical Phonology in Speech Production', *Journal of Memory and Language*, 25: 369–84.

Cruttenden, A. (1997), *Intonation* (Cambridge: Cambridge University Press, 2nd edn.).

Cutler, A. (1980), 'Syllable Omission Errors and Isochrony', in H. W. Dechert and M. Raupach (eds.), *Temporal Variables in Speech. Studies in Honor of Frieda Goldman-Eisler* (The Hague: Mouton), 183–90.

Dainora, A. (2001), 'An Empirically Based Probabilistic Model of Intonation in American English', Ph.D. Dissertation (University of Chicago).

Dauer, R. (1983), 'Stress Timing and Syllable Timing Reanalyzed', *Journal of Phonetics*, 11: 51–62.

—— (1987), 'Phonetic and Phonological Components of Language Rhythm', in the *Proceedings of XIth ICPhS* (Tallinn, Estonia, USSR), 5: 447–50.

Delais-Roussarie, E. (1995), *Pour une approche parallèle de la structure prosodique*, Thèse de Doctorat (Univ. Toulouse le Mirail).

Elordieta, G. (1997), 'Accent, Tone and Intonation in Lekeitio Basque', in F. Martínez-Gil and A. Morales-Front (eds.), *Issues in the Phonology and Morphology of the Major Iberian Languages* (Washington, DC: Georgetown University Press), 3–78.

—— (1998), 'Intonation in a Pitch-accent Dialect of Basque', *International Journal of Basque Linguistics and Philology*, 32: 511–69.

——, and Hualde, J. I. (2001), 'The Role of Duration as a Correlate of Accent in Lekeitio Basque', in *Scandinavia*, D. P., Lindberg, B., and Benner, H. (eds.), *Proceedings of Eurospeech 2001* (Aalberg University, Denmark, Center for Personkommunikation).

Fagyal, Zs. (1995), 'Phonostylistics of Read and Spontaneous French Speech: Situation and Speaker-dependent Temporal Variation', Ph.D. dissertation (University of Paris III Sorbonne Nouvelle).

Ferreira, F. (1993), 'The Creation of Prosody during Sentence Production', *Psychological Review*, 100: 233–53.

Féry, C. (1993), *German Intonational Patterns* (Tübingen: Niemeyer).

Fougeron, C. (1998), 'Variations articulatoires en début de constituants prosodiques de différents niveaux en français', Thèse de doctorat (Paris III, France).

Fox, A. (2000), *Prosodic Features and Prosodic Structure: The Phonology of Suprasegmentals* (Oxford: Oxford University Press).

Frota, S. (2000), *Prosody and Focus in European Portuguese* (New York: Garland).

——, Vigário, M., and Martins, F. (2002), 'Language Discrimination and Rhythmic Classes: Evidence from Portuguese', *Speech Prosody* 2002 (Aix-en Provence: Laboratoire Parole et Langage, France).

Gee, P., and Grosjean, F. (1983), 'Performance Structures: A Psycholinguistic and Linguistic Appraisal', *Cognitive Psychology*, 15: 411–58.

Grabe, E. (1998), *Comparative Intonational Phonology: English and German* (MPI Series in Psycholinguistics 7) (Wageningen: Ponsen and Looijen).

——, and Low, E. L. (2002), 'Durational Variability in Speech and the Rhythm Class Hypothesis', in C. Gussenhoven and N. Warner (eds.), *Laboratory Phonology 7* (Berlin and New York: Mouton de Gruyter), 515–46.

Gussenhoven, C. (1984), *On the Grammar and Semantics of Sentence Accents* (Dordrecht: Foris).

Hayes, B., and Lahiri, A. (1991), 'Bengali Intonational Phonology', *Natural Language and Linguistic Theory*, 9: 47–96.

Hirschberg, J., and Ward, G. (1992), 'The Influence of Pitch Range, Duration, Amplitude, and Spectral Features on the Interpretation of L*+H L H%', *Journal of Phonetics*, 20: 241–51.

Hirst, D., and Di Cristo, A. (1998), 'A Survey of Intonation Systems', in D. Hirst and A. Di Cristo (eds.), *Intonation Systems: A Survey of Twenty Languages* (Cambridge: Cambridge University Press), 1–44.

Hyman, L. (1978), 'Word Demarcation', in J. Greenberg (ed.), *Universals of Human Language, Vol. 2: Phonology* (Stanford: Stanford University Press), 443–70.

—— (1990), 'Boundary Tonology and the Prosodic Hierarchy', in S. Inkelas and D. Zec (eds.), *The Phonology-Syntax Connection* (Chicago: University of Chicago Press), 109–26.

Jun, S.-A. (1996), *The Phonetics and Phonology of Korean Prosody: Intonational Phonology and Prosodic Structure* (New York, NY: Garland).

—— (1998), 'The Accentual Phrase in the Korean Prosodic Hierarchy', *Phonology* 15/2: 189–226.

—— (2003*a*), 'Prosodic Phrasing and Attachment Preferences', *Journal of Psycholinguistic Research*, 32(2): 219–49.

—— (2003*b*), 'The Effect of Phrase Length and Speech Rate on Prosodic Phrasing', in the *Proceedings of the 15th International Congress of Phonetic Sciences* (Barcelona, Spain), 483–6.

——, and Elordieta, G. (1997), 'Intonational Structure of Lekeitio Basque', in A. Botinis, G. Kouroupetroglou, and G. Carayannis (eds.), *Intonation: Theory, Models, and Applications* (Proceedings of an ESCA Workshop, 18–20 Sept. 1997, Athens, Greece), 193–6.

——, and Fougeron, C. (1995), 'The Accentual Phrase and the Prosodic Structure of French', in the *Proceedings of XIIIth International Congress of Phonetic Sciences* (Stockholm, Sweden), 2: 722–5.

——, —— (2000), 'A Phonological Model of French Intonation', in A. Botinis (ed.), *Intonation: Analysis, Modeling and Technology* (Dordrecht, Netherlands: Kluwer Academic), 209–42.

——, —— (2002), 'Realizations of the Accentual Phrase in French Intonation', *Probus*, 14: 147–72, J. Hualde (ed.), a special issue on *Intonation in the Romance Languages*.

——, and Koike, C. (2003), 'Prosody and Attachment Preference in Japanese: Production and Perception', in the *Proceedings of the International Workshop on Prosodic Interfaces* (Nantes, France), 27–9 March 2003.

——, and Lee, H.-J. (1998), 'Phonetic and Phonological Markers of Contrastive Focus in Korean', in *Proceedings of the 5th International Conference on Spoken Language Processing* (Sydney, Australia), 4: 1295–8.

——, Scarborough, R., Arbisi-Kelm, T., Esposito, C., and Barjam, P. (2003), 'Intonational Phonology of Farsi', ms (UCLA).

Kim, J.-J., Lee, S.-H., Ko, H.-J., Lee, Y.-J., Kim, S.-H., and Lee, J.-C. (1997), 'An Analysis of Some Prosodic Aspects of Korean Utterances using K-ToBI Labelling System', *Proceedings of Conference on Sentence Processing* (Seoul, Korea), 87–92.

Korea Telecom Research and Development Group Report (1996), *A Study of Korean Prosody and Discourse for the Development of Speech Synthesis/Recognition System* [In Korean].

Ladd, D. R. (1992), 'An Introduction to Intonational Phonology', in G. Docherty and D. R. Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (Cambridge: Cambridge University Press), 321–4.

—— (1996), *Intonational Phonology* (Cambridge: Cambridge University Press).

—— (2001), 'Intonation', in M. Haspelmath, E. König, H. E. Wiegand, and H. Steger (eds.), *Language Typology and Language Universals: An International Handbook* (Berlin: Mouton de Gruyter), Vol. 2: 1380–90.

Lahiri, A., and Fitzpatrick-Cole, J. (1999), 'Emphatic Clitics and Focus Intonation in Bengali', in R. Kager and W. Zonneveld (eds.), *Phrasal Phonology* (Nijmegen: Nijmegen University Press), 119–44.

Lehiste, I. (1970), *Suprasegmentals* (Cambridge, MA: MIT Press).

—— (1977), 'Isochrony Reconsidered', *Journal of Phonetics*, 5: 253–63.

Liberman, M., and Pierrehumbert, J. (1984), 'Intonational Invariance under Changes in Pitch Range and Length', in M. Aronoff and R. Oehrle (eds.), *Language Sound Structure* (Cambridge, MA: MIT Press), 157–233.

Lovrić, N., Bradley, D., and Fodor, J. D. (2000), 'RC Attachment in Croatian With and Without Preposition', poster presented at the AMLaP Conference (Leiden).

Low, E. L., Grabe, E., and Nolan, F. (2000), 'Quantitative Characterizations of Speech Rhythm: Syllable-timing in Singapore English', *Language and Speech*, 43/4: 377–401.

Nakatani, L. H., O'Connor, K. D., and Aston, C. H. (1981), 'Prosodic Aspects of American English Speech Rhythm', *Phonetica*, 38: 84–106.

Nespor, M., and Vogel, I. (1986), *Prosodic Phonology* (Dordrecht: Foris).

Nibert, H. J. (2000), 'Phonetic and Phonological Evidence for Intermediate Phrasing in Spanish Intonation', Ph.D. dissertation (University of Illinois, Urbana-Champaign).

Pierrehumbert, J. (1980), 'The Phonology and Phonetics of English Intonation', Ph.D. dissertation (Massachusetts Institute of Technology).

——, and Beckman, M. E. (1988), *Japanese Tone Structure* (Cambridge, MA: MIT Press).

——, and Hirschberg, J. (1990), 'The Meaning of Intonation Contours in the Interpretation of Discourse', in P. R. Cohen, J. Morgan, and M. E. Pollack (eds.), *Intentions in Communication* (Cambridge, MA: MIT Press), 271–311.

——, and Steele, S. (1989), 'Categories of Tonal Alignment in English', *Phonetica*, 46: 181–96.

Pike, K. L. (1945), 'The Intonation of American English', in D. Bolinger (ed.), *Intonation* (Harmondsworth: Penguin), 53–83 [1972].

Prieto, V. P. (1999), 'Review of Sosa: La entonación del espoñol', in *Linguistics*, 39: 1191–9.

Quinn, D., Abdelghany, H., and Fodor, J. D. (2000), 'More Evidence of Implicit Prosody in Reading: French and Arabic Relative Clauses', poster presented at the 13th Annual CUNY Conference on Human Sentence Processing (La Jolla, CA), 30 March–1 April.

Ramus, F. (2002), 'Acoustic Correlates of Linguistic Rhythm: Perspectives', *Proceedings of Speech Prosody 2002* (Aix-en Provence: Laboratoire Parole et Langage, France), 115–20.

——, Nespor, M., and Mehler, J. (1999), 'Correlates of Linguistic Rhythm in the Speech Signal', *Cognition*, 73: 265–92.

Remijsen, B. (2001), 'Word-prosodic Systems of Raja Ampat Languages', Ph.D. dissertation (Leiden University) LOT Dissertation Series vol. 49.

Roach, P. (1982), 'On the Distinction between "Stress-timed" and "Syllable-timed" Languages', in D. Crystal (ed.), *Linguistic Controversies: Essays in Linguistic Theory and Practice in Honour of F. R. Palmer* (London: Arnold), 73–9.

Schafer, A., and Jun, S.-A. (2002), 'Effects of Accentual Phrasing on Adjective Interpretation in Korean', in M. Nakayama (ed.), *East Asian Language Processing* (Stanford: CSLI), 223–55.

Selkirk, E. (2000), 'The Interaction of Constraints on Prosodic Phrasing', in G. Bruce and M. Horne (eds.), *Prosody, Theory and Experiment: Studies presented to Gösta Bruce* (*Text, Speech, and Information Technology, Vol. 14*) (Dordrecht: Kluwer Academic).

Shattuck-Hufnagel, S., and Turk, A. (1996), 'A Prosody Tutorial for Investigators of Auditory Sentence Processing', *Journal of Psycholinguistic Research*, 25/2: 193–247.

Sosa, J. M. (1999), *La entonación del español: Su estructura fónica, variabilidad y dialectología* (Madrid: Cátedra).

Trubetzkoy, N. (1939), *GrundzUge der Phonologie. Traveaux du Cercle Linguistique de Prague 7* [Repr. (1968) Gottingen: Vandenhoek and Ruprecht].

Ueyama, M. (1998), 'Speech Rate Effects on Phrasing in English and Japanese', ms (UCLA).

——, and Jun, S.-A. (1998), 'Focus Realization in Japanese English and Korean English Intonation', in *Japanese and Korean Linguistics*, Vol. 7, CSLI (Cambridge University Press), 629–45.

van der Hulst, H. (1999), 'Word Accent', in van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe* (Berlin and New York: Mouton de Gruyter), 3–115.

Venditti, J. (2000), 'Discourse Structure and Attentional Salience Effects on Japanese Intonation', Ph.D. dissertation (Ohio State University).

Venditti, J., Jun, S.-A., and Beckman, M. E. (1996), 'Prosodic Cues to Syntactic and Other Linguistic Structures in Japanese, Korean, and English', in J. Morgan and K. Demuth (eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Mahwah, NJ: Lawrence Earlbaum Associates), 287–311.

Warner, N., and Arai, T. (2001), 'Accentual Phrase Rises as a Cue to Word Boundaries', a talk presented at the Annual meeting of the Linguistic Society of America (Washington, DC), 4–7 January 2001.