Prosodic Theory and Practice

Edited by: Jonathan Barnes, Stefanie Shattuck-Hufnagel

Citation:

Prosodic Theory and Practice
Edited by: Jonathan Barnes, Stefanie Shattuck-Hufnagel
DOI: 10.7551/mitpress/10413.001.0001
ISBN (electronic): 9780262543194
Publisher: The MIT Press
Published: 2022

OA Funding Provided By:

OA Funding from MIT Press Direct to Open

The MIT Press

4 The ToBI Transcription System: Conventions, Strengths, and Challenges

Sun-Ah Jun

4.1 Introduction

Decisions regarding what information a prosodic transcription should include vary depending on the goal of transcription. For example, the goal may be to (i) represent the distinctive features of the prosodic system of a language, (ii) learn how prosodic categories are phonetically realized and how variable the categories are, (iii) train machines to achieve better synthesis and recognition of human speech, (iv) compare prosody across dialects or across languages, or (v) develop a corpus to explore various linguistic phenomena related to prosody. Thus, prosodic transcriptions may include annotations of prosodic properties that are meaningful at the level of phonology, phonetics, or even lower (i.e., the implementation of phonetic rules for a specific language), or they can be nondistinctive but still categorical in nature, capturing prosodic properties not specific to a certain language.

ToBI, which stands for tones and break indices, is the most well-known annotation system for prosody at the level of phonology. It aims to transcribe the phonological properties of intonation (the "tones" part) and the perceived degree of juncture between words (the "break indices" part), which together represent the prominence patterns and prosodic structure of an utterance. Thus, the tone labels used in the ToBI system of a specific language are meant to represent underlying tonal targets that are both meaningful to native speakers of the language and systematic and consistent across native speakers of the language variety.

The ToBI system was originally designed for transcribing the intonation and prosodic structure of English utterances (Silverman et al. 1992; Beckman and Hirschberg 1994; Beckman and Ayers Elam 1997; Beckman, Hirschberg, and Shattuck-Hufnagel 2005), and so it began as a transcription system for the prosody of English, more specifically for mainstream American English. However, after the development of the ToBI systems for other dialects of English, such as GlaToBI for Glasgow English (Mayo, Aylett, and Ladd 1997), as well as a few typologically different languages—for example, GToBI for German (Grice and Benzmüller 1995), K-ToBI for Korean (Beckman and Jun 1996; Jun 2000), and J_ToBI for Japanese (Venditti 1997), *ToBI* has come to refer to a general framework for prosodic transcription systems based on phonological properties.

The original ToBI was therefore renamed MAE_ToBI (Mainstream American English ToBI, although sometimes it is still simply called "English ToBI") when several ToBI systems of typologically various languages (e.g., Bininj Gun-wok, Cantonese, Chickasaw, German, Greek, Japanese, Korean, Mandarin, Serbo-Croatian) were presented at a satellite meeting of the Fourteenth International Congress of Phonetic Sciences in San Francisco in 1999.[1] Since then, many ToBI systems have been developed, covering languages whose prosodic systems had been rarely studied or studied in different frameworks, such as Spanish (Beckman et al. 2002; Face and Prieto 2007), Bangladesh Bengali (Khan 2008, 2014), Mongolian (Indjieva 2009), Catalan (Prieto et al. 2009; Prieto 2014), European Portuguese (Frota 2014; Frota et al. 2015), and French (Delais-Roussarie et al. 2015), among others.

This chapter introduces what the ToBI conventions are and how the ToBI transcription system works and discusses the strengths and weaknesses of the ToBI system. It also describes some of the recent developments in the prosodic transcription system that attempt to address the ToBI system's known limitations. Before introducing how the ToBI system works, section 4.2 offers a brief description of the theoretical background and framework on which ToBI is based. Section 4.3 presents the ToBI conventions and how the ToBI system has been applied to typologically various languages, showing the workings of the phonological theory it has adopted. Section 4.4 presents the strengths of the ToBI system; section 4.5 discusses the problems and challenges ToBI users face, as well as recent developments that have been made in response to such challenges; and section 4.6 concludes the chapter.

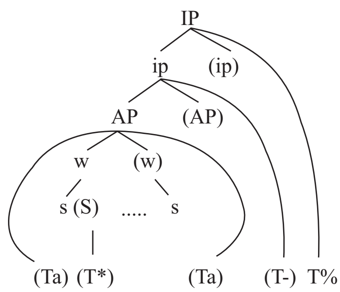
4.2 The Theoretical Background of the ToBI System

The ToBI system includes two main types of annotation: tonal labels and break index labels. The tonal labels are based on the autosegmental-metrical (AM) model of intonational phonology. (See Bruce 1977; Liberman 1975; Pierrehumbert 1980; Liberman and Pierrehumbert 1984; Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988; Ladd 1983, 2008a; and Pierrehumbert, commentary to chapter 11, this volume. See also Arvanti, chapter 1, this volume, for a detailed description of the AM model of intonational phonology.) In this model, an intonational contour is analyzed as a linear sequence of high (H) and low (L) tones. As an autosegment, the intonational tones can be associated with a specific syllable or mora in a word, that is, head, with the edge of a specific prosodic unit, or both, reflecting the metrical and/or prosodic structure of the utterance (Beckman 1996). When the tone is associated with an intonationally prominent syllable (either because it is metrically or rhythmically strong or because it is lexically marked), it is called a *pitch accent* (marked with a star [*], e.g., H*, following Goldsmith 1976 and Pierrehumbert 1980). When the tone is associated with the edge of a prosodic unit, it is called a *boundary tone* (e.g., H%, following Liberman 1975 and Pierrehumbert 1980). In this way, these intonational tones achieve two major functions: (i) marking the prominence relationship among the syllables (or moras) within a word and among the words within a phrase and (ii) marking prosodic grouping and hierarchical prosodic structure of the utterance.[2]

The AM model of intonational phonology, illustrated well in the models proposed by Beckman and Pierrehumbert (1986) and Pierrehumbert and Beckman (1988), assumes three prosodic units defined by intonation: an *intonational phrase* (IP), which is the largest prosodic unit; an *intermediate phrase* (ip), a prosodic unit smaller than an IP; and an *accentual phrase* (AP), which is slightly larger than a word (w) and smaller than an ip. These prosodic units form a strictly layered hierarchical prosodic structure (Selkirk 1986; Nespor and Vogel 1986; Pierrehumbert and Beckman 1988). A higher-level prosodic unit is assumed to be exhaustively parsed into one or more lower-level prosodic units. Therefore, any word or syllable should be part of each prosodic unit, and because an IP dominates an ip, which dominates an AP, an IP-final syllable is assumed to also be an ip-final syllable, as well as an AP-final syllable. A tree diagram of intonationally defined prosodic structure is shown in (1).[3] Each prosodic unit can be marked by a boundary tone (T% for an IP boundary tone, T- for an ip boundary tone, Ta for an

AP boundary tone; where T = H or L or its combinations), which is typically realized on the last syllable of a prosodic unit, regardless of whether the syllable is stressed. Among the boundary tones, only an IP boundary tone is obligatory. Lowercase *s* refers to a syllable, and *S* is a metrically strong/lexically marked syllable (the latter of which can be associated with a pitch accent, T*).

(1) Intonationally defined prosodic structure.



The AM model of intonational phonology allows only two tonal levels, H and L, and as in the analysis of African tone languages (Goldsmith 1976; Leben 1976), a rising or falling contour tone is represented by a combination of these level tones. This differs from the British intonation models (e.g., Palmer 1922; Halliday 1967; Crystal 1969) and from Bolinger's (1951, 1958) model where a contour tone (configuration) was represented as a single tonal unit. That is, a pitch accent in the AM model can be monotonal or bitonal, depending on the shape of the fundamental frequency (F0) contour around an accented syllable. A monotonal pitch accent represents a level tone, such as H* and L*, and a bitonal pitch accent represents a rising tone such as L + H* or a falling tone such as H* + L. By decomposing a contour tone into an L and an H level tone, the AM model can also distinguish contour tones where the F0 peak or valley is aligned differently relative to an accented syllable. That is, when a bitonal pitch accent is associated with an accented syllable of a prominent word, the starred tone of the pitch accent is realized on the accented syllable of the word. The unstarred tone, which either precedes (called a *leading tone*) or follows (called a *trailing tone*) the starred tone, is typically realized on the syllable immediately preceding or following the star-toned syllable, respectively. For example, L + H* is a rising tone with an F0 peak during the accented syllable and an F0 valley on the immediately preceding syllable, while L* + H is a rising tone with an F0 valley during the accented syllable and an F0 peak on the immediately following syllable.

As a phonological model of intonation adapted mostly from Pierrehumbert (1980), the AM model of intonational phonology posits the simplest possible abstract underlying representation. It accounts for the surface variation of underlying tones by rules that map the phonological representation (abstract level tone target sequences) to the phonetic representation of intonation (the F0 contour). Thus, even though the AM model assumes only two tonal levels in the underlying representation, various tonal levels realized in the surface F0 contour are explained by the rules that affect pitch range, such as upstep or downstep relative to the immediately preceding tone, the strength of prominence (e.g., a high nuclear pitch accent is often higher than a high prenuclear pitch accent in English), the lexical status of a tone (e.g., a lexical H tone is higher than a phrasal H tone in Japanese), or the function of tone (e.g., an L% boundary tone is typically lower than a L* pitch accent).

In addition to the rules adjusting the pitch range for a tonal category's realization, the model also adopts the target-interpolation rule to explain the surface F0 contour over syllables that have no underlying tonal targets. As a phonological model, only distinctive pitch targets are associated with a certain syllable, and syllables without a tonal target receive their surface F0 values through direct interpolation between two adjacent tonal targets, that is, the tones before and after the toneless syllable(s). In general, interpolation is done locally between two tonal targets, but it can be sensitive to the tone's function. For example, in English, interpolation does not occur between a prominence-marking pitch accent and a boundary tone in the same way that interpolation occurs between two pitch accents.

Distinctive intonational tone categories, such as pitch accents and boundary tones, proposed for a specific variety of a language will form the tonal inventory for that variety's ToBI system. But a complete ToBI annotation system should also include labels for the break index (BI) of the language. The BI in the ToBI system is in general marked by Arabic numerals, representing the perceived degree of juncture between any two adjacent words. Typically, the higher the numeral, the larger the perceived juncture between the words. Therefore, the distribution of BIs reflects prosodic grouping of words and thus the utterance's hierarchical prosodic structure.

The convention of using numbers to represent the degree of juncture is adopted from work by Price and her colleagues (Price et al. 1991; Wightman et al. 1992), who examined the mapping between syntactic and prosodic structure, and the role of prosody (especially boundary marking and prominence) in the resolution of various types of syntactic ambiguity. As Beckman et al. (2005) described in detail, the ideas that Price and her colleagues investigated have a long tradition, going back to the work known as *instrumental phonetics* (see Ladd [1996] 2008a and Price et al. 1991 for a review), where researchers examined the phonetic correlates of syntactic structure (e.g., Lieberman 1967; Lehiste 1973; Klatt 1975; Cooper and Paccia-Cooper 1980; Scott 1982), and work on *prosodic phonology* (e.g., Gee and Grosjean 1983; Selkirk 1986; Nespor and Vogel 1986), where researchers tried to predict prosodic structure from syntactic structure and explained the domain of postlexical phonological rules based on a sentence's syntactic structure.

In sum, the ToBI system is a transcription system for annotating the phonological properties of prosody, especially the prominence and prosodic structure of an utterance, analyzed in the AM model of intonational phonology for a specific language variety. The following section describes the conventions of the ToBI system based on data from English and other typologically various languages, illustrating how the phonological approach of AM theory is manifested in various ToBI systems.

4.3 ToBI Conventions and Various ToBI Systems

The goal of developing the original ToBI system was to provide a common vocabulary for tagging phonological properties of prosody and intonation to a community of users working at different sites. This common vocabulary would then facilitate the ability to share and interpret each other's data and to build a large prosodically transcribed database. In their chapter describing the original ToBI system, Beckman et al. (2005, 12–14) state that a viable ToBI system must conform to the following four principles. First, the conventions should be "accurate," conforming to an established body of research in the intonational phonology of the language variety, and, when possible, the conventions should also be informed by work on dialectology, pragmatics, and discourse analysis for the language. Second, the conventions should be "efficient," by not transcribing information that can be extracted from the signal or derived from online resources automatically (e.g., location of lexical stress). Also, prosodic phenomena should not be labeled if they cannot be labeled consistently due to a lack of understanding of the phenomena by the research community. Third, the conventions should be "easy" enough to learn by consulting a manual (which includes sound files illustrating ToBI transcriptions), and "reliable" enough to be applied consistently among transcribers. For this, intertranscriber consistency should be evaluated and the conventions should be reviewed regularly and upgraded, if needed, by agreed-on groups. Finally, a ToBI transcription should never replace a permanent record of the speech signal with a symbolic record. Instead, it should integrate symbolic labels with continuous measures derived directly from the signal. This is because a ToBI system is not simply a transcription system, but also "a tool for observing the signal and communicating one's observations to the larger community in a common language" (Beckman et al. 2005, 14).

A ToBI transcription therefore requires an audio recording of an utterance and a record of the F0 contour, aligned with symbolic labels written on four parallel quasi-independent tiers: words, tones, BIs, and miscellaneous (misc). In the original ToBI system, UNIX-based xwaves software was used to display the waveform, pitch track, and four tiers (placed above the pitch track), and a label on each tier was placed right-justified relative to a certain time point. That is, the end of each label is aligned with the end of the acoustic signal corresponding to each word (for the words and BIs tiers) or with the time point of a tonal event (for the tones tier) or a nonspeech event or interval (for the misc tier). Figure 4.1 provides an example, showing the English utterance "Okay, they have a couple flights," transcribed in MAE_ToBI, using xwaves. Later ToBI systems have used different software, such as PitchWorks (Scicon R&D; http://www.sciconrd.com/index.aspx), WaveSurfer (Department of Speech, Music, and Hearing, KTH; http://www.speech.kth.se/wavesurfer/), or Praat (http://www.fon.hum.uva.nl/praat/). Praat has been used widely in recent years by ToBI users, and unlike the format used in xwaves, has two different types of labeling tiers: it places a label either as an event that has some duration, in an *interval tier*, or as an event occurring at a specific time point, in a *point tier*. In this format, an interval tier is appropriate for the words tier, which marks the beginning and the end of each word determined by examining the waveform and spectrogram, and a point tier is appropriate for the other three tiers. Figure 4.2 shows an example of ToBI transcription of the same utterance as in Figure 4.1, but displayed in Praat. The subsections that follow describe the conventions on each tier based on data from English and other languages.

4.3.1 The Words Tier

The words tier is where an orthographic transcription of each word in the utterance appears, aligned with the acoustic representation of the word (in the waveform or the spectrogram, or both). What counts as a word is language-specific, and how to transcribe it can be decided by the research community for a specific ToBI system. For example, even though the morphosyntactic function of a case marker or a postposition is similar between Japanese and Korean, these morphemes are treated differently prosodically in each language. They are entered as an independent word in J_ToBI (aligned with BI 1), but as part of a word in K-ToBI. This is because case markers and postpositions are sometimes produced separately from a stem in Japanese, but not in Korean.

The orthography of a word can be written in language-specific characters if the font is available, or in a romanized text of the word, as in J_ToBI or K-ToBI. In Pan-Mandarin

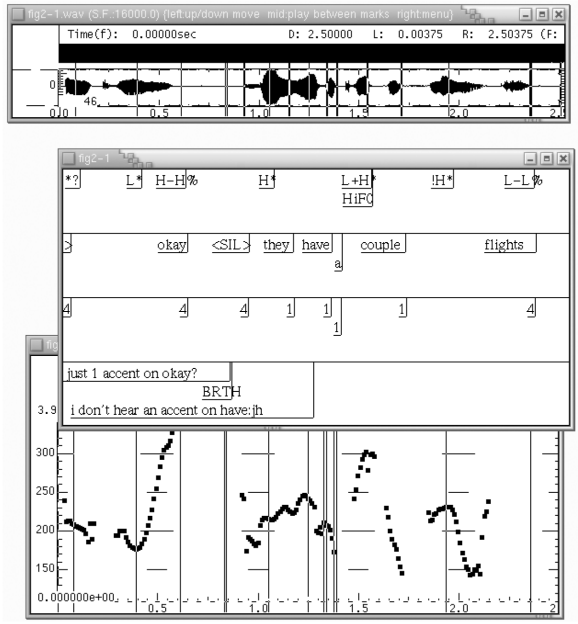


Figure 4.1
An example English utterance, "Okay. They have a couple flights," transcribed in MAE_ToBI in xwaves. The tiers are shown in the order of words, tones, BIs, and miscellaneous. *Source*: Courtesy of Beckman, Hirschberg, and Shattuck-Hufnagel (2005, figure 2.1); reproduced by permission of Oxford University Press. https://global.oup.com/academic/product/prosodic-typology-9780199208746?cc=us&lang=en&. BRTH, breath.

ToBI (Peng et al. 2005), syllable-by-syllable transcriptions in Chinese characters are given on the words tier, but transcription of syllables in modified Pinyin romanization is given on the romanization tier, called the *Romazi tier*. Similarly, the words tier in Cantonese C-ToBI (Wong, Chan, and Beckman 2005) provides a syllable-by-syllable transcription in Chinese characters, but because not every labeling platform has Chinese characters, the C-ToBI developers suggest making the words tier optional and a syllables tier obligatory, where an alphabetic transliteration is provided for every syllable in the words tier.

Finally, when the boundary of a word is not easy to locate (for example, due to cliticization of adjacent function words), the orthography of a prosodic word rather than a morphosyntactic word (e.g., *gonna* rather than *going to*) can be written on the words tier.

4.3.2 The Tones Tier

The tones tier is for transcription of distinctive tonal events and other tone-related tags. Distinctive tonal events can include (i) phrasal tones, marking the edge of a prosodic unit (e.g., Ha for an accentual phrase boundary tone in Bengali [Khan 2008, 2014] and Korean [Jun 2000, 2005a] or L% for a low boundary tone of an IP in many languages); (ii) postlexical pitch accents, marking the prominence of a word (e.g., L*, L+H* in English, German [Grice, Baumann, and Benzmueller 2005], Greek [Arvaniti and Baltazani 2005], or Bininj Gun-wok [Bishop and Fletcher 2005]); or (iii) lexically specified tonal events (e.g., lexical pitch accent H*+L in Japanese [Venditti 1997, 2005] and

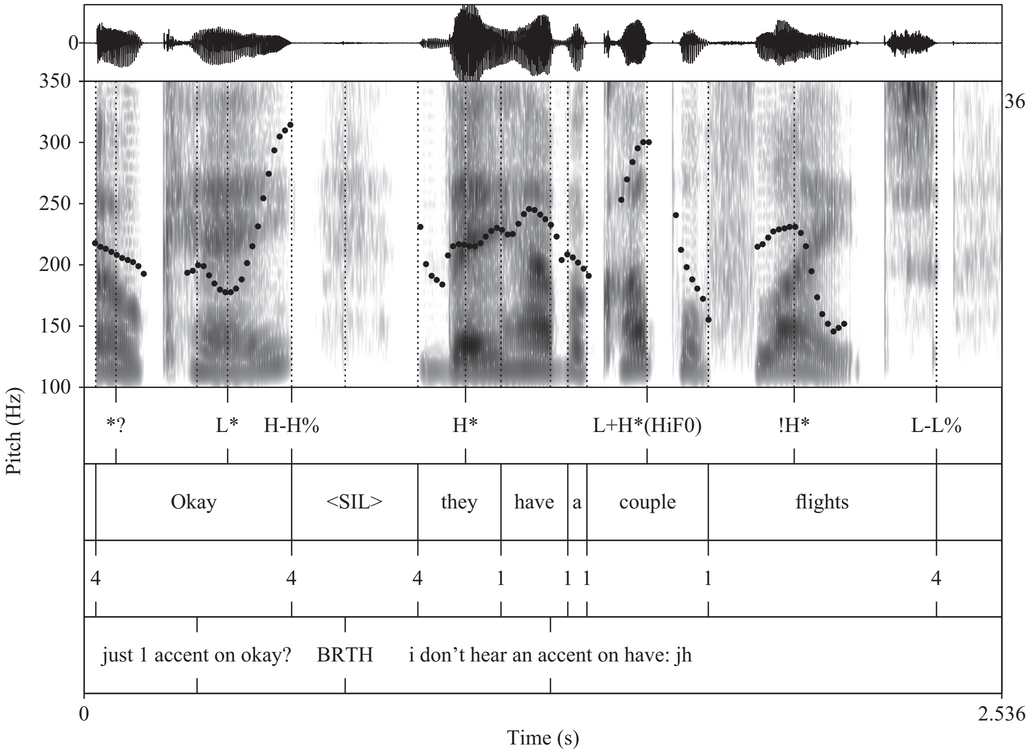


Figure 4.2
An example of MAE_ToBI transcription of the same utterance shown in figure 4.1 but transcribed in Praat.

L* + H in Serbo-Croatian [Godjevac 2005]), lexical tones in Cantonese [Wong, Chan, and Beckman 2005], or H^{λ} morphological pitch accent in Chickasaw [Gordon 2005]).

Because distinctive tonal categories and their realizations are language- or dialect-specific, the same tonal label is not necessarily realized in the same way across languages, and the same F0 contour can be transcribed with different tonal labels across languages. For example, as mentioned earlier for English intonation, a rising pitch accent L* + H is used to label an F0 contour with an F0 valley on an accented syllable and an F0 peak on the immediately following syllable. However, in Greek, the same tonal label is used for an F0 contour with an F0 valley at the very beginning or slightly before the onset of the accented syllable and an F0 peak at the beginning of the postaccentual vowel (Arvaniti, Ladd, and Mennen 1998, 2000; Arvaniti and Baltazani 2005). This is possible because each of these languages has another, contrasting rising pitch accent, such as L + H*, whose F0 peak is aligned later than that of L* + H. Thus, what justifies a tonal label in a language is its status as contrastive or noncontrastive, not the exact details of its phonetic realization.

Furthermore, the same F0 contour with the same tone-text alignment can be transcribed differently across languages depending on the lexical prosody of the target language. For example, a phrase-medial rising F0 contour over a two syllable word can be categorized as either L* + H or L + H* if the first or the second syllable, respectively, is the metrical head of the word (as in English, German, and Spanish ToBIs). If both tones are not associated with a head syllable, the contour can be categorized as a LH boundary tone or a LH tonal melody (as in Mongolian and Korean ToBIs) or an L boundary tone followed by an H phrasal one (as in J_ToBI). Or, if the first syllable is the head syllable of a word while the following syllable is the last syllable of a word, the sequence can be categorized as an L* pitch accent followed by an H boundary tone (as in Bangladesh Bengali ToBI).

In addition to the difference in tone-text alignment, tone labels reflect differences in the tonal scale or pitch range. While the AM model assumes only two tonal levels, H and L, it is also known that an H tone target is sometimes lower compared to the preceding H target, more so than would be expected given simple declination over the utterance. Furthermore, such reduction in pitch range may apply iteratively to the subsequent H tone targets. H tones realized in such a reduced pitch range are called "downstepped H tones." Because the context of downstep is similar to the downdrift phenomenon in African tone languages (Leben 1973), Beckman and Pierrehumbert (1986) proposed downstep for their AM model of English intonation to be triggered by the preceding bitonal pitch accent within the same intermediate phrase. However, the original ToBI conventions for English, influenced by Ladd (1983), adopted a theory-neutral downstep diacritic, !, before H to annotate a downstepped H tone. With this change, the MAE_ToBI revised the tonal label H+L* in Beckman and Pierrehumbert (1986) to H+!H* to reflect an F0 fall from a high to mid target rather than from a high to low target, and revised the label H*+L to H* followed by !H*. This therefore increased the tone inventory for English by adding a downstepped version of all pitch accents that have an H tone component (i.e., !H*, L+!H*, L*+!H, !H+!H*) and also a downstepped version of the H phrase accent (i.e., !H-).[4] This introduction of a downstep diacritic made the tonal category in the ToBI system more faithful to the surface F0 contour than that proposed in the AM model.[5]

The tones tier also includes tones marking the prosodic structure of an utterance. As mentioned in section 4.1, a language can have at least three prosodic units marked by intonation above the word level: IP, ip, and AP.[6] The tones marking the edge of a prosodic unit, called *boundary tones*, are accompanied by a diacritic, indicating that the prosodic unit that the tone is affiliated with. For example, *a* is used to mark an AP boundary tone (e.g., Ha, La), - for an ip boundary tone (e.g., H-, L-), and % for an IP boundary tone (e.g., H%, L%). Unlike the pitch accent, which is located within the accented syllable, specifically aligned with the F0 minimum for L*-type tones (e.g., L*, L*+H) and with the F0 maximum for H*-type tones (e.g., H*, L+H*), the boundary tone is typically located on the tones tier, aligned with the end of the relevant prosodic unit. When an ip is IP-final, though, both ip-boundary and IP-boundary tones can be realized sequentially at the edge of an IP, as in English (e.g., L-H%). In some languages, however, the boundary tone marking a lower-level unit (e.g., an ip) can be overridden by that marking a higher-level unit (e.g., an IP), as in Bangladesh Bengali or Korean. Thus, how multiple boundary tones are realized when they are hosted by the same syllable is language-specific.

The pitch accents and boundary tones mentioned so far are all postlexical tones, but the tones forming an intonation contour can also originate from a lexical tone, such as a lexical pitch accent (in Japanese and Swedish), a morphological pitch accent (in Chickasaw), or a lexical tone (in Mandarin and Cantonese). In the ToBI system of most languages, both lexical and phrasal tones are labeled on the same tier, a tones tier, because the lexical status of the tone is recognizable from the diacritics on the tone symbol. For example, in J_ToBI (1997, 2005), there is only one lexical pitch accent and it is marked with a star diacritic (i.e., H* + L). All other tones are phrasal tones: an AP phrasal tone marked with a diacritic - (i.e., H-) or a boundary tone marked with a diacritic % (i.e., an L% AP-boundary tone and a few IP-boundary tones [e.g., H%, LH%]). In Chickasaw ToBI (Gordon 2005), there is one morphological pitch accent, and it is marked with a lambda (i.e., H^{λ}). All other tones are phrasal tones: a postlexical pitch accent marked with a star (e.g., H*), AP phrasal tones without any diacritics (e.g., L, H), and an IP boundary tone marked with % (e.g., H%). Similarly, both lexical tones and IP-boundary tones are labeled on the tones tier in Cantonese ToBI (Wong, Chan, and Beckman 2005), but lexical tones are represented using Chao's system of tone numbers (e.g., 35, 53) and IP tones are marked with % (e.g., L%, HL%). However, lexical tones and postlexical tones can also be labeled on two separate tiers. In the Pan-Mandarin ToBI (Peng et al. 2005), the lexical tone is labeled on the Romazi tier while postlexical tonal events such as IP-boundary tones (i.e., L%, H%) and pitch-range–related labels (e.g., %reset, %e-prom) are labeled on the tones tier.

Finally, as shown in the original ToBI system, the tones tier can also include tone-related tags that are not intended to mark distinctive tonal events. Such tags are instead used to mark surface F0 patterns related to the phonetic realization of a tonal category or to discourse information. The most common example is to use a diacritic to mark atypical or allotonic realizations of an underlying tonal category, such as the use of < to mark a delayed F0 peak and > for an early F0 peak when the F0 peak is realized later or earlier than expected, respectively, from the typical tone-text alignment of the underlying tone. Another example is the use of *w* to mark a weak realization of an underlying tonal target, as in the undershoot of an L target in Greek prenuclear pitch accents (i.e., wL* + H) or in Japanese AP boundary tones (wL% or %wL).[7] Tone-related tags can also include a label that can be used to retrieve F0 values to study discourse structure (e.g., HiF0 in English on the highest F0 point among pitch accents within an ip), or a label to capture how pitch range is influenced by disfluency (e.g., %r at the onset of pitch-range reset after disfluency). Because a ToBI system is a tool for observing the signal and creating a communal corpus, adding these tags on surface F0 information related to the distinctive tonal categories allows researchers to learn how an underlying tone is phonetically realized in what context and how intonation interacts with other components of grammar and reflects various communicative pressures.

This idea is further extended in K-ToBI, version 3 (Jun 2000, 2005a), where the tones tier is split into two tiers, a phonological tone tier and a phonetic tone tier, to capture the variability and distribution of the surface tonal categories of the AP. Unlike English intonation, where all tonal categories are distinctive and each type of pitch accent contributes compositionally to the pragmatic interpretation of an utterance (Pierrehumbert and Hirschberg 1990), most of the tonal targets in Korean intonation are not distinctive. Distinctive tones are those marking the end of prosodic units larger than a word, and tones that are not AP-final are not distinctive. For example, an AP-initial tone can be L or H depending on the laryngeal feature of the AP-initial segment,[8] that is, predictable and thus not distinctive. AP-internal tones can be H on the second syllable or L on the penultimate syllable, but the presence of these tones varies by factors not fully known, though some of the tendencies can be explained by the AP length and by speech style. Even the AP-final tone, though usually H, can be L for factors that are also not well understood. Because the tones in ToBI systems are intended to be distinctive, nondistinctive yet categorical F0 targets such as these should not be included in the tones tier. For this reason, Jun (2000) chose to place these on a separate tier called the phonetic tone tier and put only distinctive tones on the phonological tone tier for

K-ToBI. It was hoped that labeling these surface tonal categories in a large corpus will help researchers to determine what factors and conditions trigger such variation, and thus improve the current model of Korean intonational phonology.

In sum, the tones tier is used to transcribe distinctive tonal categories of a specific language variety, regardless of their function (marking prominence or prosodic structure) and lexical status (lexical or phrasal), as well as to transcribe surface tonal events that are meaningful and interesting to the research community of the target language.

4.3.3 The Break Indices Tier

The BIs tier is for transcribing the perceived degree of juncture or the strength of association between each pair of words (and between the final word and the silence at the end of the utterance), capturing the hierarchical nature of an utterance's prosodic grouping. The original ToBI system (for English) proposed five indices in numerical numbers: four numbers for four different degrees of disjuncture and one number for a mismatch between the degree of juncture and the tonal cues that mark prosodic grouping. Among the four degrees of juncture, two corresponded to prosodic boundaries higher than the word level, that is, IP and ip, and two lower than the ip-level juncture, that is, ip-medial word boundary and a boundary weaker than a typical word boundary. (2) shows a brief description of five BIs in MAE_ToBI, represented in numeric numbers from 0 to 4. On the BIs tier, each index is labeled and aligned with the end of each word in the words tier.

(2) The BIs in MAE_ToBI

0 A juncture smaller than a typical phrase-medial word boundary. It corresponds to the juncture between a content word and a function word when the function word is cliticized. It can also be used for a weakened word juncture due to coalescence of adjacent segments across a word boundary.

1 A typical phrase-medial word boundary.

3 A juncture corresponding to an ip boundary.

4 A juncture corresponding to an IP boundary.

2 A mismatch between a tonal cue and the degree of juncture that mark a prosodic boundary. It includes two types of mismatch: (i) a juncture corresponding to ip or IP but with no tonal event defining ip or IP, and (ii) a juncture weaker than an ip level but with a clear tonal event for an ip or IP.

Because the BI is tightly linked to the prosodic structure of a language, the number of BI categories will differ across languages, reflecting the number of prosodic units a language has. However, the number of prosodic units defined by a tone does not necessarily match the number of BIs, because a prosodic unit is not necessarily marked by intonation. For example, Cantonese has only one prosodic unit defined by intonation, IP, but its ToBI system has three BI numbers: 2 for an IP boundary, 1 for a foot boundary, and 0 for a foot-internal syllable boundary. Furthermore, the Pan-Mandarin ToBI has six BI numbers, but only one BI (BI 4) corresponds to a prosodic unit defined by pitch reset.

As the list of BIs shows, an increasing value of BI refers to an increasing degree of juncture: $0 < 1 < 3 < 4$. For this reason, BI 2, which refers to a mismatch in the original ToBI, has sometimes been misinterpreted as a juncture larger than BI 1 and smaller than BI 3. Furthermore, the BI 2 in MAE_ToBI included two types of mismatch cases, as described in the BIs list in (2). To improve this situation, some of the later ToBI systems (e.g., Japanese, Korean, Greek, and Chickasaw) used BI 2 as the juncture larger than BI 1 and smaller than BI 3, and added a diacritic *m* after a BI number for mismatch cases. For example, 2m is used to label a BI 2–like juncture but without having tonal cues for the prosodic unit corresponding to BI 2 (which is an AP in Japanese, Korean, and Chickasaw, but an ip in Greek). Similarly, 3m is used to label a BI 3–like juncture but without having tonal cues for the prosodic unit corresponding to BI 3 (which is an IP in all four languages mentioned).

Because some of the BIs do correspond to a tonally defined prosodic unit, labeling both a BI on the BIs tier and a boundary tone on the tones tier would mark the same prosodic unit. For example, in MAE_ToBI, BI 3 and BI 4 are labeled on the BIs tier aligned with a phrase accent and an IP boundary tone, respectively, on the tones tier. This seems to violate the "efficiency" principle of the ToBI system, mentioned at the beginning of section 4.3. Namely, the system has to be efficient and should not transcribe predictable information. So, the researchers of a specific ToBI system could decide not to label BIs that are predictable from other tiers but label only those not predictable such as mismatch cases. But this would work if both tones and BI tiers are fully specified. In the cases where a research group wants to examine only the tonal patterns of a prosodic structure or only the degree of juncture or grouping between words, the tier of interest should be fully labeled regardless of whether some label is predictable from the other tier. That is, depending on the goal of the researchers, what to label on what tier can be different.[9]

Finally, a BI can also transcribe a disfluent juncture, a juncture produced disfluently or unnaturally for various reasons (e.g., a memory problem, a sudden change in a word choice, interruption by other speakers, stuttering), and the original ToBI proposed three types of labels for a disfluent juncture, as in (3).

(3) ToBI labels for disfluent juncture
1p for an abrupt cutoff of a word
2p for prolongation, but without a phrase accent
3p for prolongation, with an accompanying phrase accent

These two types of disfluency, cutoff and prolongation, have been adopted by most of the later ToBI systems, but these are not enough to transcribe all types of disfluencies found in speech.[10] To transcribe stuttered speech, Arbisi-Kelm (2006) increased the types of disfluency from two to five by adding "restart," "pause," and "filler." However, BIs describing disfluent junctures have not been used extensively across ToBI systems, probably because most ToBI systems have been proposed based on reading and scripted speech or semi-spontaneous speech and not much from free conversation speech.

4.3.4 The Miscellaneous Tier

The misc tier is for transcription of any comments or remarks about the utterance. It can be a comment about the utterance such as silence, audible breaths (an example is "BRTH" in figure 4.1), laughter, false start, hesitation, disfluencies, and other spontaneous speech effects. It can also be used to include labelers' remarks for themselves about a transcription, or questions to other labelers (e.g., "just 1 accent on okay?" in figure 4.1). If a label on the misc tier is for a localized event, it should be located at the time point of the event (e.g., a label "disfl" is written at the approximate time point where some disfluency is perceived), but if the event has an identifiable interval such as laughter does, the convention is to use paired labels to mark the starting point of the event with < and the ending point of the event with > (e.g., laughter< . . . laughter>).

4.3.5 ToBI Labels for Uncertainty and an Alt Tier

The original ToBI system provided a convention to transcribe a labeler's uncertainty about a BI or a tonal category. When a labeler is not sure about the degree of juncture between two adjacent BIs, the convention is to add a dash (-) after a BI of a larger juncture (e.g., 4- is used for a BI ambiguous between 3 and 4).[11] For a tonal category, *? is used when a labeler is not sure whether a syllable is pitch accented or not, and X*? is used when a labeler is sure about the presence of pitch accent on a syllable but not sure about the pitch accent type. The same rule is applied to other tonal categories. So, uncertainty about the presence and the type of a phrase accent is labeled as -?and X-?, respectively, and that of an IP boundary tone is labeled as %?and X%?, respectively.

However, this way of labeling uncertainty is often found to be unsatisfactory because it does not record relevant details of the ambiguity. That is, while it tags that the labeler is uncertain about the transcription, it does not specify what possible tonal labels were considered by the labeler, or whether one potential label should be preferred over another. To fix this problem, a fifth tier, Alt (alternative), was proposed at the Workshop on ToBI for Spontaneous Speech, held in Boston in 2004, and the use of an Alt tier has been tested by labeling a large speech sample in Brugos et al. (2008). Brugos and her colleagues proposed adding a question mark next to an ambiguous or uncertain tonal label (e.g., H*?) or a BI label (e.g., 3?) in the respective tier and providing an alternative tonal or BI label on the Alt tier (e.g., !H* or 2), with the alternative label aligned with the corresponding labels on the tones or BIs tiers. If a sequence of labels is ambiguous, instead of one label, they proposed adding a square bracket before and after the affected labels on the Alt tier (e.g., [2 L+H*]).

4.3.6 Other Tiers

As mentioned, the original ToBI system provided four core tiers (tones, words, BIs, misc), but later ToBI systems have added new tiers either to capture language-specific prosodic properties (e.g., Romazi, syllable, stress, and sandhi tiers in Pan_Mandarin ToBI; foot and syllable tiers in C_ToBI), to study the interface between prosody and other areas of linguistics (e.g., the phonetic transcription tier in GRToBI and Chickasaw ToBI, the prosodic word tier in GRToBI, the finality tier in J_ToBI), or to help labelers transcribe non-native or unfamiliar languages (e.g., the gloss tier in Chickasaw ToBI and Bininj Gun-wok ToBI; see Jun 2005b for examples of each tier used in different ToBI systems). Creation of these noncore tiers in various languages exemplifies the flexibility and extendibility of ToBI system as a tool for transcribing language-specific prosodic features as well as other linguistic properties that are useful to the research community or system users.

4.4 Strengths of the ToBI System

The strengths of the ToBI system come from the transcription of tones being phonological, especially based on the AM model of intonational phonology, and from the architecture and mechanism of the transcription system.

First, as a phonological transcription system based on the AM framework of intonational phonology, the ToBI system transcribes distinctive tonal categories of intonation that mark prominence and a hierarchical prosodic structure in a specific language variety. That is, the tones that are transcribed are not just a sequence of F0 turning points on the surface F0 contour, but are contrastive in the language by performing linguistic functions, either marking prominence or information structure, or marking prosodic structure, or both.

Though changes in pitch heavily contribute to the perception of prominence and the degree of juncture, perception of prominence and juncture is influenced by multiple sources of information such as acoustic differences in F0, duration and intensity, voice quality, and even linguistic structures (Cole, Mo, and Baek 2010; Cole, Mo, and Hasegawa-Johnson 2010; Bishop 2012). Because perception can be subjectively influenced by an individual's bias and linguistic background, ToBI requires observation of visual representations of acoustics, especially the F0 track and the waveform/spectrogram of the utterance. This holistic approach to transcription is necessary because ToBI transcription is phonological, capturing the information of metrical and prosodic structure as well as the tonal category. In this way, it is different from other transcription systems focusing on only F0 representations, such as INTSINT (International Transcription System for Intonation; Hirst 1998, and chapter 3, this volume; Hirst and Di Cristo 1998). In INTSINT, an F0 contour is labeled as a sequence of symbols referring to either an absolute tone level (e.g., *t* or ⇑ for the top of the speaker's pitch range) or a relative tone (e.g., *h* or ↑ for a higher tone than the preceding tonal segment), without considering the function of the tone or the prosodic structure of an utterance. Thus, the ToBI tone labels can be equated to broad phonemic transcription of the International Phonetic Alphabet (IPA), while INTSINT symbols correspond to a narrow phonetic transcription of the IPA. That is, the goal of ToBI is not simply to regenerate the surface F0 contours of any sentence in any language, but to capture tonal properties that are linguistically meaningful, especially for the function of marking prominence and prosodic structure.

Because ToBI is language-specific and the tonal categories of each ToBI system are contrastive in each language, defined based on the same theoretical assumptions and principles and using a common vocabulary, comparing the tonal types and the structure of tonal sequences across various ToBI systems allows us to identify intonational universals and study prosodic typology.

For example, as mentioned in sections 4.2 and 4.3.2, the function of ToBI tones, whether they mark prominence or a prosodic structure, can be identified by tonal symbols, especially diacritics. In the ToBI system of languages that have stress, a prominence marking tone is labeled with a diacritic * (e.g., H*, L*+H), called a (postlexical) pitch accent, while tones marking the boundary of a prosodic unit are labeled with an edge-marking diacritic such as %, -, or *a* (to mark a boundary tone of an IP, ip, and AP, respectively). But the * diacritic is also added to a lexical pitch accent, as shown in the Japanese and Serbo-Croatian ToBI systems. Unlike the postlexical pitch accent found in English, the lexical pitch accent in Japanese is not metrically strong, so an accented mora is not realized with higher amplitude or longer duration than an unaccented one (Beckman 1986). Thus, a starred tone marks a syllable or mora's special status as the "head" of a word, but that head is not necessarily associated with stress. If a language has no lexical prosody and has no word head, as in Korean, Mongolian, and West Greenlandic, no starred tone is used to categorize the F0 contour. In these headless languages, distinctive tonal categories are all phrasal tones or boundary tones marking a prosodic unit (e.g., an HLH tonal rhythm in West Greenlandic aligned with the right edge of a phonological word [Arnhold 2014] or a rising boundary tone on the left edge of an intermediate phrase [i.e., -LH] in Mongolian [Karlsson 2014]). In these headless languages, a word becomes prominent by being located at the edge of a larger prosodic unit. In Jun's (2014a) model of prosodic typology, these languages are categorized as edge-prominence languages, while languages like English, where prominence is marked by the head of a prosodic unit, are categorized as head-prominence languages (cf. Beckman 1996).[12] So, comparing diacritics used in the tonal labels of a ToBI system can provide data for typological analyses of prominence marking and prosodic structure.

Along the same line, by comparing tonal categories across ToBI systems, we can identify the most common versus rare tonal categories, or the most common prosodic units across languages. We can also learn which languages have a contrast on the rising slopes (e.g., H* versus L+H*) or on differences in the alignment of rises (e.g., L+H* versus L*+H) or falls (e.g., H+L*, H*+L), and so on. Languages can also be classified based on the number and type of pitch accents and boundary tones they have, and on how regular their tonal pattern is (for such comparisons, see Jun 2005b, 2014a).

Labeling distinctive categories of prosody by referring to their phonetic realizations in ToBI can also provide excellent data for studying the acquisition of prosody as a second or foreign language or the prosody of languages in contact. This is because one needs to know the distinctive tonal categories and prosodic structure, as well as their phonetic realizations, in both languages to study prosodic transfer or interference between them (e.g., McGory 1997; Ueyama and Jun 1998; Jun and Oh 2000; Dupoux et al. 2008; Atterer and Ladd 2004; Mennen 2004, 2007, 2015; Schepman, Lickley, and Ladd 2006; Trofimovich and Baker 2006; O'Brien and Gut 2010; Lee and Jun 2016). For example, both Japanese and Korean L2 (second language) speakers of English produced the nuclear pitch accent on the focused word in English declarative sentences (i.e., L+H*) better than that in English interrogative sentences (L*). This is because a focused word in Japanese and Korean is always produced with an H tone regardless of the sentence type (McGory 1997; Ueyama and Jun 1998). This suggests that difficulty in producing L2 prosody is due to the differences in the prosodic category between L2 and the speaker's native language (L1). Difficulty in producing L2 prosody can also come from the differences in phonetic realizations of the same prosodic category between L1 and L2. An example can be seen when bilingual speakers of German tend to produce a delayed onset of F0 rise when producing English prenuclear rising pitch accent, because the onset of F0 rise is aligned later in German than in English (Atterer and Ladd 2004). Similarly, both Greek and Dutch have a prenuclear rising pitch accent, but Dutch speakers of Greek often produce an earlier F0 peak when producing Greek prenuclear rising pitch accent, because the F0 peak is aligned later than the accented syllable in Greek than in Dutch (Mennen 1998, 2004). Transcribing the prosody of utterances produced by L2 speakers or bilinguals by referring to the ToBI systems of the two languages involved reveals how the prosodic properties of two languages, either categorical or phonetic, interact in the speech of L2 learners or bilinguals (Lee and Jun 2016).

An additional strength of the ToBI system can come from the flexibility of adding new tiers. Each tier in ToBI can be used to transcribe different prosodic and linguistic information, and the misc tier can even carry nonlinguistic information. As noted in section 4.3.6, ToBI systems for various languages have added new tiers to tag language-specific prosodic information as well as other linguistic information tailored to the particular interests of the research community of the ToBI variety. For example, C-ToBI added a foot tier to transcribe the fusion forms, that is, to mark the degree of lenition between two adjacent syllables, as well as emphasized syllables and phrases. The goal of transcribing syllable fusion was to explore the factors affecting syllable fusion, and the goal of transcribing emphasized syllables was to explore whether the domain of prominence has any relationship with the prosodic hierarchy.[13] The addition of a phonetic tone tier in K-ToBI also allowed researchers to explore the factors affecting the various tonal patterns observed in the AP. By transcribing the allotonic surface tonal categories in the phonetic tone tier for various utterances produced in different speech styles, Yoo and Jun (2016) recently found that, between the two AP-medial tones, the H tone is realized more frequently than the L tone and is more common in spontaneous speech than in formal speech. A new tier can therefore be added to a ToBI system of a specific language for an exploratory purpose, and what to label in the tier can be decided by the particular group of researchers working on the target language variety.

In addition to testing and improving existing models of intonation by transcribing language-specific information in a new tier, ToBI also represents a useful tool for studying the interfaces between prosody and various subareas of linguistics. Utilizing parallel quasi-independent multitiers aligned with each other, researchers can create and examine ToBI-labeled corpora (e.g, Ostendorf et al. 2001; Grice and Savino 2003) or add ToBI labels to an existing database such as the English Broadcast News Speech (Fiscus et al. 1998) and the Buckeye Corpus (Pitt et al. 2005). For example, the MAE_ToBI-labeled corpus of the Boston University Radio Speech Corpus (Ostendorf, Price, and Shattuck-Hufnagel 1995) was examined to study the realization of glottal stops and glottalization of word-initial vowels in different prosodic positions (e.g., Dilley, Shattuck-Hufnagel, and Ostendorf 1996; Garellek 2013). News speech corpus (Fiscus et al. 1998) and National Public Radio news excerpts have also been labeled in MAE_ToBI to examine the prosodic realization of reflexive pronouns (e.g., *himself, themselves*) relative to their syntactic and semantic structures (Ahn 2015). Similarly, Grice and Savino (2003) analyzed dialogues from the map tasks in Bari Italian by adopting the coding scheme used in the English Human Communication Research Center (HCRC) Map Task (Anderson et al. 1991) and transcribed the intonation of the utterances in the corpus following the ToBI notation in Grice et al. (2005). They discovered that in Bari Italian, polar questions asking about new information employ a rising pitch accent (L+H*), while questions about given information (i.e., confirmation-seeking questions) are expressed with a falling pitch accent (H*+L or H+L* depending on a contrast setting). Furthermore, Stirling et al. (2001) transcribed part of the speech corpora from the Australian National Database of Spoken Language (ANDOSL) map task corpus (Miller et al. 1994) in Australian English ToBI (AuE_ToBI, Fletcher and Harrington 1996) and investigated correlations between discourse structure and various prosodic properties.

Finally, researchers can also use tonal categories and BIs from ToBI transcriptions to test linguistic theories and hypotheses in experimental contexts. ToBI is a particularly useful tool when manipulating prosody for experimental stimuli and for evaluating and categorizing prosodic structures produced by experimental subjects. Perhaps the best illustration of this type of use is research on the role of prosody in sentence processing (e.g., Speer, Kjelgaard, and Dobroth 1996; Schafer et al. 1996; Kjelgaard and Speer 1999; Schafer et al. 2000; Schafer, Speer, and Warren 2005; Ito and Speer 2008; Snedeker and Casserly 2010; Jun 2010; Lee and Watson 2011; Speer and Foltz 2015; Jun and Bishop 2015). These studies explored how manipulations of prosodic structure influence listeners' interpretation and processing of syntactic structure, or, conversely, how their production of prosody changes as a result of manipulating syntactic structure or pragmatic context. Prosodic manipulation and evaluation in these studies was facilitated by the adoption of the tonal categories and/or BIs from the ToBI conventions of the target language. For example, to test Fodor's implicit prosody hypothesis (Fodor 1998, 2002),[14] the prosodic phrasing of sentences produced by subjects (e.g., "Someone shot the servant of the actress who was on the balcony") was assessed by labeling a BI and a boundary tone, if present, after each head noun ("the servant" and "the actress"; see Bergmann, Armstrong, and Maday 2008 and Jun 2010 for English data using MAE_ToBI; Jun and Koike 2008 for Japanese data using J_ToBI; and Jun and

Kim 2004 for Korean data using K-ToBI). Speer and Foltz (2015) also tested the implicit prosody hypothesis by manipulating the presence and type of pitch accent on target words in visual-to-auditory cross-modal priming experiments. Based on ToBI-labeling speech samples taken from paragraphs read aloud, they found that only speakers who reliably produced L+H* pitch accents on contrastively focused words in the produced speech sample (i.e., with explicit prosody) responded faster to an auditory probe word in L+H* following the silent reading of a correctively focused prime word. That is, priming of contrastive intonation from implicit to explicit prosody depended on the speech styles of individual speakers, which were established by way of a ToBI annotation of individual speakers' prosody.

In sum, the ToBI system, based on the AM model of intonational phonology, is an important tool for advancing knowledge on the phonetics and phonology of prosody. It is particularly useful for illuminating the relation between prosody and other linguistic phenomena in a way not possible before the system's creation. As mentioned in Beckman, Hirschberg, and Shattuck-Hufnagel (2005, 46), ToBI is "an ongoing research program rather than a set of rules cast in stone for all time." In addition to facilitating linguistic research, it is a valuable tool for creating communal corpora that can be used to address a variety of research questions. It can also be used by engineers and speech scientists to enhance the performance of systems for recognizing and synthesizing intonation. With ToBI-labeled data, we can investigate how prosodic categories are phonetically realized, and how prosody delivers phonological, syntactic, semantic, and discourse-related information. We can also explore how languages differ in their marking of such complex information prosodically.

4.5 Challenges and Recent Developments

This section describes challenges that ToBI users face as the result of some of the system's weaknesses. While some of these challenges stem from properties of the AM theory that ToBI adopts, others arise from the nature of transcribing real speech using a discrete phonological transcription system. This section also describes some of the recent developments the ToBI community has devised to address ToBI's weaknesses.

One of the challenges ToBI users encounter is that the prominence patterns and prosodic structure captured by ToBI transcriptions are limited to the size of an utterance corresponding to an IP. This is a consequence of the fact that, in the AM model of intonational phonology, the IP is the highest prosodic unit defined by a tone.[15] Therefore, prosodic events that occur across an IP boundary or over a sequence of IPs cannot be captured by the current ToBI system. This makes it difficult to use ToBI to capture how prosody is used to mark complex discourse structure, or possible syntagmatic relations between tones across IPs (e.g., Ladd 1990, [1996] 2008a; Calhoun 2006, 2012; Bishop 2013).[16] One of the main prosodic cues known to mark discourse structure is pitch range (Brown, Currie, and Kenworthy 1980; Hirschberg and Pierrehumbert 1986). While ToBI includes pitch reset among the acoustic cues defining a prosodic unit, no labels exist to cover pitch range changes across ips and IPs. Mandarin ToBI (Peng et al. 2005) includes a limited number of labels for pitch range, but they refer to effects within one sentence or phrase (i.e., one breath group): that is, %reset to mark pitch reset over a declarative phrase, %q-raise to mark raised pitch range (as in echo questions), %e-prom for the local expansion of pitch range within a phrase due to emphatic prominence, and %compressed for the local compression of pitch range after %e-prom. The MAE_ToBI system does include the label HiF0, intended to mark the highest F0 point for a pitch accent within an ip. But this was simply intended as a tag for accessing the F0 value for use later as a measure of pitch range to study the relationship between pitch range and discourse structure. Instead of using a tones tier, a new tier could be added for pitch-range information. J_ToBI, for example, added an extra tier to mark finality. Though the finality tier for J_ToBI only marks a juncture larger than an IP (delivering "the sense that the speaker is 'done' in her turn in discourse planning"; Venditti 1997, 2005), this type of tier could be used to label acoustic cues to the relationship between IPs and larger discourse segments.

Because most of the existing ToBI systems were developed on the basis of sentence-length laboratory speech with relatively simple discourse structure (e.g., reading of a story, news reports, interviews, short dialogues, speech from map tasks), ToBI transcriptions have not been widely used by researchers working on the discourse structure of free conversational speech. However, as noted in Beckman, Hirschberg, and Shattuck-Hufnagel (2005, 12), transcription conventions for ToBI are ideally "based on a large and long-established body of research on dialectology, pragmatics and discourse analysis for the target language," in addition to study of its intonational phonology. The inclusion of a tier that allows for annotating relationships between ips and IPs (or between larger discourse segments) would thus be a straightforward way to address ToBI's current weakness in this area.

In a related vein, the current ToBI system is also not optimized for labeling global, and gradual, prosodic events. This stems from basic properties of the AM model of intonation, which represents the intonation contour as a linear sequence of low and high tonal targets. If a researcher is more interested in the domain of global prosodic events (e.g., pitch range reset at the beginning of prosodic unit X or Y) or relative differences between domains (e.g., whether the pitch range of unit X is larger or smaller than that of unit Y), this could be addressed by creating a label to mark the onset or magnitude of the prosodic feature at the edges of units. However, modifications that require a researcher to address more gradual changes over domains present a much greater challenge to the ToBI system (e.g., annotating changes in the slope of declination, which marks sentence types in Danish intonation; Thorsen 1980, 1983).

Given the variability and gradience inherent in acoustic signals, the next challenge ToBI users face is not due to the system's reliance on the AM model specifically but to its basic status as phonological. First, prosodic categories are not always realized in the same way. As Cole and Shattuck-Hufnagel (2016), Cangemi and Grice (2016), and Arvaniti (2016) have emphasized, acoustic cues to a prosodic category are not always realized consistently. Instead, factors such as time pressures and segmental/prosodic context often render cues reduced or ambiguous, and prosodic categories can also be realized differently depending on the individual speaker's strategy to encoding prosodic contrasts (e.g., Mo, Cole, and Lee 2008; Cole and Shattuck-Hufnagel 2016). Moreover, criteria for defining prosodic categories, such as phrase-final lengthening for an IP or ip boundary, are gradient rather than categorical. Categorizing a prosodic event based on gradient and variable acoustic cues is not straightforward in the absence of other information supporting the category, and the final decision can be further influenced by the individual labeler's weighting of the acoustic cues.

Second, and related to this, recall from section 4.4 how ToBI labels are decided based on the labeler's perception of prominence and degree of juncture from an audio file, while at the same time observing the visual display of a pitch track together with a spectrogram or waveform. This means that the labeling is not fully based on objective criteria and can be influenced by the labeler's perception of acoustic cues, interpretation of linguistic context, and level of knowledge and experience in ToBI labeling. It is therefore recommended that ToBI labeling be done by multiple trained labelers to reduce the subjectivity of the annotations. However, because ToBI labeling is labor intensive and requires training in both acoustics and AM theory, it is not always easy to find multiple ToBI labelers to participate in the annotation of experimental data or databases.

The fact that ToBI systems are tools for phonological transcription means they are well suited to studying prosodic typology. One of the strengths of a phonological transcription system is that it enables comparison of how different languages (and language varieties) tonally mark prominence and prosodic structure. However, this also comes with a challenge. Because ToBI's tone labels are language-specific, reflecting contrasts in each language, the same surface F0 contour is not necessarily assigned the same labels across languages (just as the same Voice Onset Time [VOT] can be categorized as /p/ in Spanish but as /b/ in English). Additionally, as pointed out by Ladd (2008b) and Hualde and Prieto (2016), it is also possible that the same surface F0 contours, representing the same contrast in each language, can be labeled differently depending on what analysis and tonal categories are chosen by the developers of the ToBI system.

A good example illustrating this point can be found in the comparison of MAE_ToBI and GToBI (Grice et al. 2005). Both American English and Standard German have four levels of sentence-final pitch, representing the same basic meaning/function: low F0 at the end of a neutral statement, mid F0 at the end of a calling contour, high flat F0 for incompleteness, and super-high F0 at the end of a yes-no question. However, these four types of contours are transcribed differently: as L-L%, !H-L%, H-L%, and H-H% in MAE_ToBI, but as L-%, !H-%, H-%, and H-^H% in GToBI. This is because English ToBI developers adopted an upstep rule (i.e., an IP boundary tone is realized higher after a (!)H- phrase accent), but GToBI developers did not. Instead, GToBI's creators added two symbols intended to make their system for German more phonetically transparent. One symbol, an upstep diacritic ^, is used to directly mark a super-high tone, ^H; another, a toneless IP boundary symbol %, is used when the IP-final pitch level is the same as the ip-final pitch (i.e., H-% means a high plateau and L-% a low plateau). Thus the difference between these two ToBI systems was a matter of choosing how abstract versus how phonetically transparent the tonal representation should be. Because the tones in the original ToBI system, by introducing a downstep diacritic, were claimed to be more faithful to the surface F0 patterns than those proposed in the original AM model (see section 4.3.2), various ToBI systems have added tones and symbols to better capture surface F0 contours, but there are still no widely agreed-on rules or conventions regarding what symbols and labels to use and on what occasions to use them. (See the "portability" problem of ToBI notation argued in Hualde and Prieto 2016.) As more ToBI systems are developed in the future, there will be new symbols and labels created, making comparison across ToBI systems more difficult and the study of prosodic typology thus more challenging. Students and researchers will also likely find it more difficult to determine what some labels and symbols in different ToBI systems mean[17] and what level of abstraction they represent.[18]

In addition to the difficulty in interpreting symbols and levels of abstraction in different ToBI systems generally, there is also the challenge of deciding what level of abstraction is appropriate when a surface tone does not match an underlying tone. Should it be transcribed with an allophonic tone label or with an underlying tone label? As a system designed to transcribe phonological categories, clearly the underlying tone should be used. However, if the goal is simply to tag one's observations about the speech signal, one might want to use an allophonic tone label. Cases like this have been a known source of confusion and disagreement among labelers (e.g., Escudero et al. 2012; Armstrong 2017; Hualde and Prieto 2016), and the chosen solutions vary across ToBI systems. For example, when an underlying L% boundary tone is truncated after an L+H* on a phrase-final syllable (and thus realized as an approximately mid-level tone), should the boundary tone be labeled as L%, !H%, or something else? Prieto and Ortega-Liebaria (2009) and Hualde and Prieto (2016) used !H% in Catalan and Peninsular Spanish ToBIs, but Grice et al. (2005) used (L%)—L% in parentheses—for southern Italian varieties, as a notation for a partially realized underlying tone.

As a tagging tool, the ToBI system can include labels that represent surface tonal patterns, and this makes its level of tonal representation clearly less abstract than that proposed in the AM model of intonational phonology. However, the common practice of using ToBI-style labeling (employing primarily the words and tones tiers) when trying to build a model of intonational phonology of a language seems to have blurred the distinction between the ToBI system and the AM model of intonational phonology. This trend can be observed in publications such as Prieto and Roseano (2010), Jun (2014b), and Frota and Prieto (2015), where most models of intonational phonology and most ToBI transcription systems look quite similar to each other. Consequently, the BI part of ToBI has also become less emphasized in the ToBI system. In fact, it is not uncommon to hear researchers refer to the “ToBI model” of a particular language, which would seem to equate the transcription system with the intonational phonology model of the language.[19]

Many of the issues and challenges described herein motivated a recent proposal to develop an international prosodic alphabet (IPrA) within the AM framework (Hualde and Prieto 2016 and Jun and Fletcher 2014; see also presentations by Jun, Prieto, and Hualde at the satellite workshop of the International Congress of Phonetic Sciences 2015 in Glasgow[20]). The current proposal calls for ToBI systems to include two levels of tonal transcription—one distinctive, thus phonological, and the other nondistinctive but categorical, thus phonetic—and to create a set of discrete IPrA tonal labels and diacritics (the details of which are to be agreed on by the international community of ToBI users). The goal is that the symbols in IPrA can be used for both phonological and categorical phonetic transcriptions, like the symbols in the IPA, so that they can be used in both phonemic and phonetic representations. The labels for *phonological* transcription would only mark phonological contrast in the language, and the *categorical phonetic* labels would be for tones that are categorical in nature but not distinctive, or whose distinctiveness is not yet known. These categorical phonetic labels can (i) represent allophonic realizations of an underlying tonal category, (ii) be used as temporary labels before their phonological status is established (e.g., at the beginning stage of building a model of intonational phonology), or (iii) represent hybrid or exceptional tonal categories that are not part of the intonational model of any specific language (e.g., L2 speech, bilinguals’ speech, or speech produced by speakers of multiple languages in contact). The phonological tone labels can be used in the tones tier, and categorical phonetic tones can be labeled in a new, separate tier (e.g., an IPrA tier) or can be used in the tones tier together with the phonological tones, but written in square brackets (e.g., [L*] for an allophonic L*). It is hoped that separating categorical phonetic tones from phonological tones in ToBI will facilitate the development of more abstract phonological analyses of intonation and better represent the correspondence between underlying tonal categories and surface patterns in a systematic way. This will help us to clarify the levels of transcription that different ToBI systems are using and make more transparent comparisons across languages possible. Clarifying the levels of transcription will also increase interlabeler agreement rates and facilitate labeling large corpora and further improve the strengths of the ToBI system mentioned in section 4.4.

In addition to the IPrA proposal, some of the weaknesses of the ToBI system motivated the recent development of a few other transcription systems such as rapid prosody transcription (RPT) and rhythm and pitch (RaP) transcription. The RPT system was developed by Cole and her colleagues (Mo, Cole, and Lee 2008; Cole, Mo, and Baek 2010; and Cole, Mahrt, and Hualde 2014). Unlike ToBI transcription, which is labor-intensive and requires training and knowledge of the theoretical background, the RPT transcription is, as the name implies, a coarse-grained transcription done by untrained naive native speakers of the target language.[21] Because this type of annotation can be done as a group, it is possible to generate transcriptions from many labelers in a very short time. In the RPT task, untrained native speakers (such as college students) are given instructions defining “prominent” words and “prosodic juncture” in running speech. They are then asked to listen to a sample of running speech and identify prominent words and prosodic boundaries on a printed or electronically presented transcript of the speech sample. Identification of these prosodic events is done in real time (i.e., the “rapid” part of RPT), with the subjects typically listening to the entire speech sample four times, twice each for prominence and boundary transcription (they are not allowed to replay any specific parts of the sample). The length of speech samples used can vary from ten to sixty seconds, which generally includes one or more syntactic clauses and one or more prosodic phrases. In Cole and colleagues’ studies, the task was used to generate a prominence score (p-score) and a boundary score (b-score) for each word (i.e., the proportion of labelers who marked each word as prominent or as preceding a boundary, respectively). The magnitude of p-scores and b-scores for individual words and their distribution thus provides a measure of agreement among labelers and an estimation of the variability in the perception of prosodic events in the speech material (Mo, Cole, and Lee 2008; Cole and Shattuck-Hufnagel 2016).

Bishop and his colleagues (Bishop and Kuo 2016; Bishop, Kuo, and Kim 2020) compared the responses from an RPT task with the labels provided by expert ToBI annotators. Their results showed that RPT subjects identified just under 50% of the nuclear pitch accents (NPAs) and just under 50% of the IP boundaries that ToBI labelers assigned to the speech materials. Thus a little more than half of the time, the naive labelers failed to identify NPA-level prominence and IP-level boundaries. This is not surprising given that the RPT labelers were not trained to attend to prominence and phrasing at a detailed level, and the transcription was done in real time without examining pitch tracks or being able to hear parts of the sentence as many times as desired. However, from the words identified as prominent or having a prosodic boundary by the RPT labelers, pitch-accented words were identified as prominent more often (i.e., had higher p-scores) than unaccented words, and words at an IP boundary were identified as a boundary more often (i.e., had higher b-scores) than those at an ip boundary. This suggests that the p- and b-scores from a large set of speech materials annotated by a large group of RPT labelers can provide a good measure of the perceptual salience of prosodic categories. RPT may also provide useful information about the acoustic correlates of various prosodic categories, as well as about how the perception of prominence and boundaries is influenced by grammatical information. As pointed out in Cole and Shattuck-Hufnagel (2016), the patterning of the p- and b-scores from RPT may shed light on the mechanisms involved in prosodic processing in speech production and perception.

Next, the RaP system was developed by Dilley and her colleagues (Dilley 2005; Dilley and Brown 2005; Dilley et al. 2006; Breen et al. 2012) to address transcription difficulties associated with ToBI's vulnerability to the variability and gradience of tonal categories and to reflect the importance of rhythmic or metrical prominence distinct from pitch accent. In ToBI, tonal labels such as H* or L-H% include both the identity of the tonal target (e.g., H or L or !H) and information regarding its function in terms of both metrical prominence (i.e., pitch accents are explicitly marked by *) and prosodic structure (i.e., boundary tones are marked by % and -). The RaP system, in contrast, emphasizes the independence of pitch information from rhythmic/metrical and phrasal structure. Pitch information is transcribed in terms of three tonal targets (H, L, E) whose value is relative to a preceding tone (higher, lower, or equal to the preceding tonal labels, respectively). Tones in RaP, labeled in the pitch tier, are therefore represented phonetically. Metrical prominence (in three levels: strong, weak, none)[22] and prosodic structure (in two levels: IP and ip) are transcribed in the rhythm tier. Although RaP was proposed as "a method of labeling the rhythm and relative pitch of spoken English" (Dilley and Brown 2005, 2), the concept and principles of the system can be applied to other languages, and the system can be used by researchers interested in examining structural information separately from tonal information.

4.6 Conclusion

This chapter introduced the ToBI system for transcribing phonological aspects of prosody. It discussed both the strengths and challenges associated with the system, as well as some of the recent developments the research community has made in response to such challenges. As a phonological annotation system based on the AM model of intonational phonology, each language's ToBI system is unique, reflecting the prosodic properties of the language in question. The ToBI framework allows a common set of conventions to be used across languages, despite the fact that how prominence and phrasal structure are phonologically marked and phonetically realized is language-specific. In the ToBI system, tones and BIs are annotated based on the labeler's perceived prominence of words and degree of juncture between words, supplemented by the visual display of the utterance's acoustics, especially F0 track and waveforms (or spectrogram). Furthermore, as a tool to annotate and communicate observations made about the speech signal with other users of ToBI, the ToBI system allows communal corpora to be created for testing hypotheses about various linguistic and prosodic phenomena. The output of such research also serves as feedback to the ToBI transcription system itself and informs models of intonational phonology.

It was shown that most of the strengths as well as most of the weaknesses associated with ToBI stem from its being a phonological transcription system based on the AM model. The ToBI system's emphasis on transcribing distinctive phonological categories of tones and metrical/prosodic structure makes it the most valuable tool currently available for studying the grammar of prosody and intonation. ToBI users can pursue questions about prosody's relation to a number of other linguistic subfields, and it is especially well-suited to the study of prosodic typology and universals. However, as a phonological system, the labels proposed in various ToBI systems do not necessarily have the same phonetic values or represent the same level of abstraction from system to system, and so an effort should be made when comparing multiple ToBI systems to study prosodic typology. Efforts to address this particular issue are currently underway, such as the recent proposal to develop an IPrA, allowing for the categorical phonetic representation of tones in each ToBI system.

ToBI is a valuable tool for conducting research on prosody, and it is a powerful and flexible transcription system. With approximately fifty ToBI systems now established from typologically different languages, tremendous progress has been made since the original ToBI system was published in 1994, in terms of both prosody's relation to other linguistic subfields and our understanding of prosodic typology. Although ToBI transcription is labor-intensive, and becoming a confident labeler requires significant training, the benefits seem clear and justified. Ongoing and future tasks for the ToBI research community include improving on the system's current limitations, as well as continuing to expand the number (and diversity) of languages described within the ToBI framework. Taken together, the ToBI framework represents a powerful and flexible tool, advancing our understanding of prosodic systems and prosodic typology.

Notes

1. Most of the ToBI systems presented at the meeting, including the MAE_ToBI, were later published in Jun (2005c).

2. As will be shown in sections 4.3.2 and 4.4, not all languages mark intonational prominence by way of pitch accent. If a language has no stress or lexically marked syllables—and thus has no designated terminal element for words (e.g., Korean, Mongolian, or West Greenlandic), prominence is not marked by a pitch accent but by positioning the prominent word at the edge of a prosodic unit. Thus how prosodic prominence is marked varies across languages. See Jun (2014a) for a typology of prominence marking.

3. This prosodic structure is similar to, but slightly different from, the prosodic hierarchy proposed by prosodic phonologists in the 1980s, which was defined indirectly based on the syntactic structure of a sentence (e.g., Nespor and Vogel 1986; Selkirk 1986; Hayes 1989). In that approach, prosodic units commonly assumed to be larger than a word include, from the largest to the smallest, an utterance (Utt) > an intonational phrase (IPh) > a phonological phrase (PPh) > clitic group. Both an Utt and an IPh correspond to an IP in intonational phonology, but a PPh corresponds to either an ip or an AP in intonational phonology (see Jun 1998 for comparisons between these two approaches).

4. The H*+L pitch accent proposed in Pierrehumbert (1980) and Beckman and Pierrehumbert (1986) is not a falling tone as implied by the trailing L tone. As mentioned in Beckman, Hirschberg, and Shattuck-Hufnagel (2005), it is for a simple H* pitch accent, followed by a downstepped H*. The trailing L was simply included to provide the necessary bitonality required, as shown in the research on African tone languages (e.g., Leben 1973; Goldsmith 1976), to trigger downstep of a following H tone. That is, a sequence of H*+L pitch accents was used to refer to a series of stepping-down pitch accents, examined in Liberman and Pierrehumbert (1984).

5. Allowing tonal labels to be more faithful to the surface F0 contour became the most prominent difference between the tonal inventory of the AM model of intonational phonology and that of the ToBI system. As will be discussed in section 4.5, this difference was a source of confusion to many researchers and significantly affected subsequent models of the intonational phonology of other languages.

6. A prosodic word can also be marked by intonation as in Serbo-Croatian (Godjevac 2005).

7. Because the undershoot allophones of an L boundary tone are predictable from the words tier in J_ToBI, both wL% and %wL symbols are removed in X-JToBI (Maekawa et al. 2002), a revised version of J_ToBI.

8. If the AP-initial segment is either an aspirated or a tense consonant, the tone is H. Otherwise, the tone is L.

9. The GToBI had the same number of prosodic units as in MAE_ToBI, but the GToBI developers proposed not to label a BI lower than 3 and proposed to label two types of mismatch cases, labeled 2r and 2t. BI 2r refers to rhythmic break with tonal continuity, and BI 2t refers to tonal break with rhythmic continuity.

10. In the original ToBI, filler words (such as *uh* and *um*, also known as filled pauses) were not treated as a type of disfluency. But this and other types of disfluency could be added in other ToBI systems.

11. When a labeler is sure about the degree of juncture at a certain word boundary but the tonal realization does not match the corresponding degree of juncture, a mismatch label (2 or #m, discussed in section 4.3.3) should be used.

12. Languages can cue a word's prominence by marking both the head and the edge of the word, and this type of language is categorized as a head-/edge-prominence language (examples are Georgian, Bengali, French, and Japanese). The tonal inventory of these languages includes a starred tone as well as a boundary tone of a word or AP (see Jun 2014a).

13. In C-ToBI, the domain of the foot is reflected in the BI tier by aligning the foot-medial weakened syllable with a BI 0, while the foot-final syllable is aligned with a BI 1. However, the domain of prominence is not linked with the BI tier, reflecting the unknown relationship between the domain of prominence and the prosodic hierarchy. The developers of C-ToBI state that the architecture of independent parallel tiers "highlights one of the principal advantages of the looser structure of the original ToBI framework" (Wong, Chan, and Beckman 2005, 293).

14. The hypothesis was proposed to explain cross-linguistic differences in attachment preferences. Fodor claimed that the difference is due to implicit default prosody assigned to the syntactic structure of each sentence. She predicted that, in the example sentence where the attachment of a relative clause (RC) is ambiguous between the first noun ("the servant") or the second ("the actress"), languages that prefer high attachment (i.e., RC attachment to the first noun) would have a larger juncture after the second noun than after the first noun, and the opposite is true for low attachment preference languages. She also assumed implicit prosody generated in silent reading is equal to explicit prosody generated in reading aloud.

15. The AM model does not include a prosodic unit above the IP, such as the utterance-level unit proposed by some prosodic phonologists (e.g., Nespor and Vogel 1986). The primary reason for this is the absence of distinctive tones or other categorical prosodic events that mark any unit above the IP.

16. In the AM model of English intonation adopted in the MAE_ToBI, an ip is the domain of nuclear pitch accent and the domain of pitch reset. Thus, capturing prominence relation across ips is limited.

17. Examples where ToBI systems assign different symbols for the same phenomenon are easily found among the systems currently available. For example, undershoot has been marked with either *w* (in Japanese and Greek ToBIs) or *u* (in Bangladesh Bengali ToBI), and an upstep has been marked with ^ (in German and Bininj Gun-wok ToBI) or with ¡ (in Spanish and Catalan ToBIs).

18. For example, downstep can be phonological or phonetic across languages. Its status can also vary within a language depending on tone types (e.g., in GRToBI, downstepped H is contrastive for a phrase accent (!H-) or a boundary tone (!H%) but not for a pitch accent (e.g., !H* or !H*+L).

19. In this case, a "ToBI model" would seem to refer to an intonational phonology model that is more transparent and surface-rich than the traditional AM model of intonational phonology.

20. For more information about the workshop, visit http://linguistics.ucla.edu/ipra_workshop/.

21. The RPT methods have been used for Hindi, Russian, French, and Spanish in addition to English (for some cautionary notes on applying the RPT across languages, however, see Gussenhoven 2015). It has also been used with fluent bilinguals and learners of English as L2 (see Cole and Shattuck-Hufnagel 2016).

22. In English ToBI, metrical prominence is also represented at three levels (Beckman & Edwards 1994): NPA, pitch accent, no accent. The NPA is the last pitch accent within an ip. Because it is predictable from the location of pitch accent and an ip boundary tone, NPA is not labeled with a separate symbol in English ToBI. But in languages where NPA is not predictable as in Italian, an NPA tone is marked by a diacritic *n* (e.g., L* + Hn).

References

Ahn, B. 2015. "Giving Reflexivity a Voice: Twin Reflexives in English." PhD diss., UCLA.

Anderson, A., M. Bader, E. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, et al. 1991. "The HCRC Map Task Corpus." *Language and Speech* 34:351–366.

Arbisi-Kelm, T. 2006. "An Intonational Analysis of Disfluency Patterns in the Speech of Stutterers." PhD diss., UCLA.

Armstrong, M. E. 2017. "Accounting for Intonational Form and Function in Puerto Rican Spanish Polar Questions." *Probus* 29 (1): 1–40. https://doi.org/10.1515/probus-2014-0016.

Arnhold, A. 2014. "Prosodic Structure and Focus Realization in West Greenlandic." In Jun 2014b, 216–251.

Arvaniti, A. 2016. "Analytical Decisions in Intonation Research and the Role of Representations: Lessons from Romani." *Laboratory Phonology* 7 (1): 1–43.

Arvaniti, A., and M. Baltazani. 2005. "Intonational Analysis and Prosodic Annotation of Greek Spoken Corpora." In Jun 2005c, 84–117.

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

Arvaniti, A., D. R. Ladd, and I. Mennen. 2000. "What Is a Starred Tone? Evidence from Greek." In *Papers in Laboratory Phonology V*, edited by M. Broe and J. Pierrehumbert, 119–131. Cambridge: Cambridge University Press.

Atterer, M., and D. R. Ladd. 2004. "On the Phonetics and Phonology of "Segmental Anchoring" of F0: Evidence from German." *Journal of Phonetics* 32: 177–197.

Beckman, M. E. 1986. *Stress and Non-Stress Accent*. Dordrecht, the Netherlands: Foris.

Beckman, M. E. 1996. "The Parsing of Prosody." *Language and Cognitive Processes* 11:17–67.

Beckman, M. E., and G. M. Ayers Elam. 1997. "Guidelines for ToBI Labeling." Unpublished manuscript. Ohio State University. https://www.ling.ohio-state.edu/research/phonetics/E_ToBI/.

Beckman, M. E., M. Díaz-Campos, J. T. McGory, and T. A. Morgan, 2002. "Intonation across Spanish, in the Tones and Break Indices Framework." *Probus* 14:9–36.

Beckman, M. E., and J. Edwards. 1994. "Articulatory Evidence for Differentiating Stress Categories." In *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, edited by P. A. Keating, 7–33. Cambridge: Cambridge University Press.

Beckman, M. E., and J. Hirschberg. 1994. "The ToBI Annotation Conventions." Unpublished manuscript. Ohio State University. https://web.archive.org/web/20100622195855/http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/ToBI/ToBI.6.html.

Beckman, M. E., J. Hirschberg, and S. Shattuck-Hufnagel. 2005. "The Original ToBI System and the Evolution of the ToBI Framework." In Jun 2005c, 9–54.

Beckman, M. E., and S.-A. Jun. 1996. "K-ToBI (Korean ToBI) Labelling Convention." Version 2. Unpublished manuscript. Ohio State University and UCLA. https://linguistics.ucla.edu/people/jun/ktobi/k-tobi-V2.html.

Beckman, M. E., and J. Pierrehumbert. 1986. "Intonational Structure in Japanese and English." *Phonology Yearbook* 3:255–309.

Bergmann, A., M. Armstrong, and K. Maday, K. 2008. "Relative Clause Attachment in English and Spanish: A Production Study." In *Proceedings of the Fourth International Conference on Speech Prosody*. https://www.isca-speech.org/archive/sp2008/papers/sp08_505.pdf.

Bishop, J. 2012. "Information Structural Expectations in the Perception of Prosodic Prominence." In *Prosody and Meaning (Interface Explorations)*, edited by G. Elordieta and P. Prieto, 239–269. Berlin: Mouton de Gruyter.

Bishop, J. 2013. "Prenuclear Prominence: Phonetics, Phonology, and Information Structure." PhD diss., UCLA.

Bishop, J., and J. Fletcher. 2005. "Intonation in Six Dialects of Bininj Gun-wok." In Jun 2005c, 331–361.

Bishop, J., and G. Kuo. 2016. "Do 'Autistic-Like' Personality Traits Predict Prosody Perception?" Paper presented at LabPhon 15 (Workshop on Personality in Speech), Cornell University, Ithaca, New York, July 17.

Bishop, J., G. Kuo, and B. Kim. 2020. "Phonology, Phonetics, and Signal-Extrinsic Factors in the Perception of Prosodic Prominence: Evidence from Rapid Prosody Transcription." *Journal of Phonetics* 82. https://doi.org/10.1016/j.wocn.2020.100977.

Bolinger, D. L. 1951. "Intonation: Levels versus Configurations." *Word* 7:199–210.

Bolinger, D. L. 1958. "A Theory of Pitch Accent in English." *Word* 14:109–149.

Breen, M., L. C. Dilley, J. Kraemer, and E. Gibson. 2012. "Inter-Transcriber Reliability for Two Systems of Prosodic Annotation: ToBI (Tones and Break Indices) and RaP (Rhythm and Pitch)." *Corpus Linguistics and Linguistic Theory* 8:277–312.

Brown, G., K. Currie, and J. Kenworthy 1980. *Questions of Intonation*. London: Croom-Helm.

Bruce, G. 1977. *Swedish Word Accents in Sentence Perspective*. Lund, Sweden: Gleerup.

Brugos, A., N. Veilleux, M. Breen, and S. Shattuck-Hufnagel. 2008. "The Alternatives (Alt) Tier for ToBI: Advantages of Capturing Prosodic Ambiguity." In *Proceedings of the Fourth International Conference on Speech Prosody*. https://www.isca-speech.org/archive/sp2008/papers/sp08_273.pdf.

Calhoun, S. 2006. "Information Structure and the Prosodic Structure of English: A Probabilistic Relationship." PhD diss., University of Edinburgh.

Calhoun, S. 2012. "The Theme/Rheme Distinction: Accent Type or Relative Prominence?" *Journal of Phonetics* 40:329–349.

Cangemi, F., and M. Grice. 2016. "The Importance of a Distributional Approach to Categoriality in Autosegmental-Metrical Accounts of Intonation." *Laboratory Phonology* 7 (1): 1–20.

Cole, J., T. Mahrt, and J. I. Hualde. 2014. "Listening for Sound, Listening for Meaning: Task Effects on Prosodic Effects on Prosodic Transcription." In *Proceedings of the Seventh International Conference on Speech Prosody*. https://www.isca-speech.org/archive/SpeechProsody_2014/pdfs/165.pdf.

Cole, J., Y. Mo, and S. Baek. 2010. "The Role of Syntactic Structure in Guiding Prosody Perception with Ordinary Listeners and Everyday Speech." *Language and Cognitive Processes* 25:1141–1177.

Cole, J., Y. Mo, and M. Hasegawa-Johnson. 2010. "Signal-Based and Expectation-Based Factors in the Perception of Prosodic Prominence." *Laboratory Phonology* 1:425–452.

Cole, J., and S. Shattuck-Hufnagel. 2016. "New Methods for Prosodic Transcription: Capturing Variability as a Source of Information." *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 7 (1): 8. http://dx.doi.org/10.5334/labphon.29.

Cooper, W., and P. Paccia-Cooper. 1980. *Syntax and Speech*. Cambridge, MA: Harvard University Press.

Crystal, D. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.

Delais-Roussarie, E., B. Post, M. Avanzi, C. Buthke, A. Di Cristo, I. Feldhausen, S.-A. Jun, et al. 2015. "Intonational Phonology of French: Developing a ToBI System for French." In *Intonation in Romance*, edited by Sonia Frota and Pilar Prieto, 63–100. Oxford: Oxford University Press.

Dilley, L. C. 2005. "The Phonetics and Phonology of Tonal Systems." PhD diss., MIT.

Dilley, L. C., M. Breen, M. Bolivar, J. Kraemer, and E. Gibson. 2006. "A Comparison of Inter-Transcriber Reliability for Two Systems of Prosodic Annotation: RaP (Rhythm and Pitch) and ToBI (Tones and Break Indices)." In *Proceedings of Interspeech 2006*. https://www.isca-speech.org/archive/archive_papers/interspeech_2006/i06_1619.pdf.

Dilley, L., and M. Brown, M. 2005. "The RaP (Rhythm and Pitch) Labeling System." Version 1.0. https://pdfs.semanticscholar.org/5f73/1dbcafb2b64da6eb15daa67718866bc74cc9.pdf.

Dilley, L., S. Shattuck-Hufnagel, and M. Ostendorf. 1996. "Glottalization of Word-Initial Vowels as a Function of Prosodic Structure." *Journal of Phonetics* 24:423–444.

Dupoux, E., N. Sebastián-Gallés, E. Navarrete, and S. Peperkamp. 2008. "Persistent Stress 'Deafness': The Case of French Learners of Spanish." *Cognition* 106:682–706.

Escudero, D., L. Aguilar, M. del M. Vanrell, and P. Prieto. 2012. "Analysis of Inter-Transcriber Consistency in the Cat_ToBI Prosodic Labelling System." *Speech Communication* 54:566–582.

Face, T., and P. Prieto. 2007. "Rising Accents in Castilian Spanish: A Revision of Sp_ToBI." *Journal of Portuguese Linguistics* 6:117–146.

Fiscus, J., J. Garofolo, M. Przybocki, W. Fisher, and D. Pallett. 1998. *English Broadcast News Speech (HUB4)*. Philadelphia: Linguistic Data Consortium.

Fletcher, J., and J. Harrington. 1996. "Timing of Intonational Events in Australian English." In *Proceedings of the Sixth Australian International Conference on Speech Science and Technology*, edited by P. McCormack and A. Russell, 611–615. Australian Speech Science and Technology Association, Canberra, Australia.

Fodor, J. D. 1998. "Learning to Parse." *Journal of Psycholinguistic Research* 27:285–319.

Fodor, J. D. 2002. "Prosodic Disambiguation in Silent Reading." *Proceedings of the North East Linguistic Society* 32:113–132.

Frota, S. 2014. "The Intonational Phonology of European Portuguese." In Jun 2014b, 7–42.

Frota, S., P. Oliveira, M. Cruz, and M. Vigário. 2015. *P-ToBI: Tools for the Transcription of Portuguese Prosody*. Lisbon: Laboratório de Fonética, CLUL/FLUL. http://labfon.letras.ulisboa.pt/InAPoP/P-ToBI/.

Frota, S., and P. Prieto. 2015. *Intonation in Romance*. Oxford: Oxford University Press.

Garellek, M. 2013. "Production and Perception of Glottal Stops." PhD diss., UCLA.

Gee, J. P., and F. Grosjean. 1983. "Performance Structures: A Psycholinguistic and Linguistic Appraisal." *Cognitive Psychology* 15:411–458.

Godjevac, S. 2005. "Transcribing Serbo-Croatian Intonation." In Jun 2005c, 146–171.

Goldsmith, J. 1976. "Autosegmental Phonology." PhD diss., MIT.

Gordon, M. 2005. "Intonational Phonology of Chickasaw." In Jun 2005c, 301–330.

Grice, M., S. Baumann, and R. Benzmueller. 2005. "German Intonation in Autosegmental-Metrical Phonology." In Jun 2005c, 55–83.

Grice, M., and R. Benzmueller. 1995. "Transcription of German Intonation Using ToBI-Tones—The Saarbrücken System." *Phonus* 1:33–51.

Grice, M., M. D'Imperio, M. Savino, and C. Avesani. 2005. "Strategies for Intonation Labeling across Varieties of Italian." In Jun 2005c, 362–389.

Grice, M., and M. Savino. 2003. "Map Tasks in Italian: Asking Questions about Given, Accessible and New Information." *Catalan Journal of Linguistic* 2:153–180.

Gussenhoven, C. 2015. "Does Phonological Prominence Exist?" *Lingue e Linguaggio* 14:7–24.

Halliday, M. A. K. 1967. *Intonation and Grammar in British English*. The Hague, the Netherlands: Mouton.

Hayes, B. 1989. "The Prosodic Hierarchy in Meter." In *Perspectives on Meter*, edited by P. Kiparsky and G. Youmans, 203–260. New York: Academic Press.

Hirschberg, J., and J. Pierrehumbert. 1986. "Intonational Structuring of Discourse." *Proceedings of the 24th Meeting of the Association for Computational Linguistics*, 126–144.

Hirst, D. 1998. "Intonation in British English." In *Intonation Systems*, edited by D. Hirst and A. Di Cristo, 56–77. Cambridge: Cambridge University Press.

Hirst, D., and A. Di Cristo. 1998. "A Survey of Intonation Systems." In *Intonation Systems*, edited by D. Hirst and A. Di Cristo, 1–44. Cambridge: Cambridge University Press.

Hualde, J., and P. Prieto. 2016. "Towards an International Prosodic Alphabet (IPrA)." *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 7 (1): 1–25, especially 5. http://dx.doi.org/10.5334/labphon.11.

Indjieva, E. 2009. "Oirat Tones and Break Indices (O-ToBI). Intonational Structure of the Oirat Language." PhD diss., University of Hawaii.

Ito, K., and S. R. Speer. 2008. "Anticipatory Effect of Intonation: Eye Movements during Instructed Visual Search." *Journal of Memory and Language* 58:541–573.

Jun, S.-A. 1998. "The Accentual Phrase in the Korean Prosodic Hierarchy." *Phonology* 15 (2): 189–226.

Jun, S.-A. 2000. "K-ToBI (Korean ToBI) Labeling Conventions, Version 3.1." *UCLA Working Papers in Phonetics* 99:149–173.

Jun, S.-A. 2005a. "Korean Intonational Phonology." In Jun 2005c, 201–229.

Jun, S.-A. 2005b. "Prosodic Typology." In Jun 2005c, 430–458.

Jun, S.-A., ed. 2005c. *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Jun, S.-A. 2010. "The Implicit Prosody Hypothesis and Overt Prosody in English." *Language and Cognitive Processes* 25:1201–1233.

Jun, S.-A. 2014a. "Prosodic Typology: By Prominence Type, Word Prosody, and Macro-Rhythm." In Jun 2014b, 520–539.

Jun, S.-A., ed. 2014b. *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Jun, S.-A., and J. Bishop. 2015. "Priming Implicit Prosody: Prosodic Boundaries and Individual Differences." *Language and Speech* 58 (4): 459–473.

Jun, S.-A., and J. Fletcher. 2014. "Methodology of Studying Intonation: From Data Collection to Data Analysis." In Jun 2014b, 493–519.

Jun, S.-A., and S. Kim, S. 2004. "Default Phrasing and Attachment Preferences in Korean." In *Proceedings of Interspeech 2004*. https://www.isca-speech.org/archive/archive_papers/interspeech_2004/i04_3009.pdf.

Jun, S.-A., and C. Koike. 2008. "Default Prosody and RC Attachment in Japanese." *Japanese-Korean Linguistics* 3:41–53.

Jun, S.-A., and M. Oh. 2000. "Acquisition of 2nd Language Intonation." *Proceedings of International Conference on Spoken Language Processing* 4:76–79.

Karlsson, A. 2014 "The Intonational Phonology of Mongolian." In Jun 2014b, 187–215.

Khan, S. D. 2008. "Intonational Phonology and Focus Prosody of Bengali." PhD diss., UCLA.

Khan, S. D. 2014. "The Intonational Phonology of Bangladeshi Standard Bengali." In Jun 2014b, 82–117.

Kjelgaard, M. M., and S. Speer. 1999. "Prosodic Facilitation and Interference in the Resolution of Temporary Syntactic Closure Ambiguity." *Journal of Memory and Language* 40:153–194.

Klatt, D. H. 1975. "Vowel Lengthening Is Syntactically Determined in a Connected Discourse." *Journal of Phonetics* 3:129–140.

Ladd, D. R. 1983. "Phonological Features of Intonational Peaks." *Language* 59:721–759.

Ladd, D. R. 1990. "The Metrical Representation of Pitch Register." In *Papers in Laboratory Phonology I*, edited by J. Kingston and M. Beckman, 35–57. Cambridge: Cambridge University Press.

Ladd, D. R. (1996) 2008a. *Intonational Phonology*. 2nd ed. Cambridge: Cambridge University Press.

Ladd, D. R. 2008b. "Review of Prosodic Typology by Jun 2005c." *Phonology* 25 (2): 372–376.

Leben, W. 1973. "Suprasegmental Phonology." PhD diss., MIT.

Leben, W. 1976. "The Tones in English Intonation." *Linguistic Analysis* 2:69–107.

Lee, E.-K., and D. Watson. 2011. "Effects of Pitch Accents in Attachment Ambiguity Resolution." *Language and Cognitive Processes* 26 (2): 262–297. https://linguistics.ucla.edu/people/jun/papers%20in%20pdf/ASAPoster2016LeeandJun.pdf.

Lee, H., and S.-A. Jun. 2016. "Types of Errors in English Prosody Made by Native Korean Speakers." Poster presented at the Acoustical Society of America, Honolulu, Hawaii, December 1. https://linguistics.ucla.edu/people/jun/papers%20in%20pdf/ASAPoster2016YooandJun.pdf.

Lehiste, I. 1973. "Phonetic Disambiguation of Syntactic Ambiguity." *Glossa* 7 (2): 107–121.

Liberman, M. 1975. "The Intonational System of English." PhD diss., MIT.

Liberman, M., and J. Pierrehumbert. 1984. "Intonational Invariance under Changes in Pitch Range and Length." In *Language Sound Structure*, edited by M. Aronoff and R. Oehrle, 157–233. Cambridge, MA: MIT Press.

Lieberman, P. 1967. *Intonation, Perception and Language*. Cambridge, MA: MIT Press.

Maekawa, K., H. Kikuchi, Y. Igarashi, and J. Venditti. 2002. "X-JToBI: An Extended J-ToBI for Spontaneous Speech." In *Proceedings of the Seventh International Conference on Spoken Language Processing*. https://www.isca-speech.org/archive/archive_papers/icslp_2002/i02_1545.pdf.

Mayo, C., M. Aylett, and D. R. Ladd. 1997. "Prosodic Transcription of Glasgow English: An Evaluation Study of GlaToBI." In *Proceedings of the ESCA Workshop on Intonation: Theory, Models and Applications*. https://www.cstr.ed.ac.uk/downloads/publications/1997/esca2.pdf.

McGory, J. 1997. "Acquisition of Intonational Prominence in English by Seoul Korean and Mandarin Chinese Speakers." PhD diss., Ohio State University.

Mennen, I. 1998. "Stability of Tonal Alignment: The Case of Greek Prenuclear Accents." *Journal of Phonetics* 26 (1): 3–25.

Mennen, I. 2004. "Bi-Directional Interference in the Intonation of Dutch Speakers of Greek." *Journal of Phonetics* 32:543–563.

Mennen, I. 2007. "Phonological and Phonetic Influences in Non-Native Intonation." In *Non-Native Prosody: Phonetic Descriptions and Teaching Practice*, edited by J. Trouvain and U. Gut, 53–76. Berlin: Mouton de Gruyter.

Mennen I. 2015. "Beyond Segments: Towards a L2 intonation Learning Theory." In *Prosody and Language in Contact*, edited by E. Delais-Roussaire, M. Avanzi, and S. Herment, 171–188. Heidelberg, Germany: Springer-Verlag.

Millar, J., J. Vonwiller, J. Harrington, and P. Dermody. 1994. "The Australian National Database of Spoken Language." *Proceedings of the ICASSP* 94:97–100.

Mo, Y., J. Cole, and E. K. Lee. 2008. "Prosody Perception by Naive Listeners: Evidence from a Large Multi-Transcriber Reliability Study." Poster presented at the Eighty-Second Annual Meeting of the Linguistic Society of America, Chicago, IL. January 3–6.

Nespor, M., and I. Vogel. 1986. *Prosodic Phonology*. Dordrecht, the Netherlands: Foris.

O'Brien, M., and U. Gut. 2010 "Phonological and Phonetic Realisation of Different Types of Focus in L2 Speech." In *Achievements and Perspectives in the Acquisition of Second Language*

Speech: New Sounds 2010, edited by K. Dziubalska-Kołaczyk, M. Wrembel, and M. Kul, 205–215. Frankfurt, Germany: Peter Lang.

Ostendorf, M., P. J. Price, and S. Shattuck-Hufnagel. 1995. *The Boston University Radio News Corpus*. Technical Report ECS-95–001. Boston: Boston University.

Ostendorf, M., I. Shafran, S. Shattuck-Hufnagel, L. Carmichael, and W. Byrne. 2001. "A Prosodically Labeled Database of Spontaneous Speech." In *Proceedings of the ISCA Tutorial and Research Workshop on Prosody in Speech Recognition and Understanding*. https://www.isca-speech.org/archive_open/archive_papers/prosody_2001/prsr_022.pdf.

Palmer, H. E. 1922. *English Intonation with Systematic Exercises*. Cambridge: Heffer.

Peng, S-H., C. Marjorie, C.-Y. Tseng, T. Huang, O. J. Lee, and M. Beckman. 2005. "Towards a Pan_Mandarin System for Prosodic Transcription." In Jun 2005c, 230–270.

Pierrehumbert, J. 1980. "The Phonology and Phonetics of English Intonation." PhD diss., MIT.

Pierrehumbert, J., and M. Beckman. 1988. *Japanese Tone Structure*. Cambridge, MA: MIT Press.

Pierrehumbert, J., and J. Hirschberg. 1990. "The Meaning of Intonational Contours in the Interpretation of Discourse." In *Intensions in Communication*, edited by P. R. Cohen, J. Morgan, and M. E. Pollack, 271–311. Cambridge, MA: MIT Press.

Pitt, M. A., K. Johnson, E. Hume, S. Kiesling, and W. Raymond. 2005. "The Buckeye Corpus of Conversational Speech: Labeling Conventions and a Test of Transcriber Reliability." *Speech Communication* 45:89–95.

Price, M. P. J., S. Ostendorf, S. Shattuck-Hufnagel, and C. Fong. 1991. "The Use of Prosody in Syntactic Disambiguation." *Journal of Acoustical Society of America* 60:2956–2970.

Prieto, P. 2014. "The Intonational Phonology of Catalan." In Jun 2014b, 43–80.

Prieto, P., L. Aguilar, I. Mascaró, F. J. Torres-Tamarit, and M. Vanrell. M. 2009. "L'etiquetatge prosòdic Cat_ToBI*." Estudios de Fonètica Experimental* 18:287–309.

Prieto, P., and M. Ortega-Llebaria. 2009. "Do Complex Tones Induce Syllable Lengthening in Catalan and Spanish?" In *Phonetics and Phonology: Interactions and Interrelations*, edited by M. Vigário, S. Frota, and M. J. Freitas, 51–70. Amsterdam: John Benjamins.

Prieto, P., and P. Roseano, eds. 2010. *Transcription of Intonation of the Spanish Language*. Munich: Lincom.

Schafer, A., J. Carter, C. Clifton, and L. Frazier 1996. "Focus in Relative Clause Construal." *Language and Cognitive Processes* 11:135–163.

Schafer, A., S. Speer, and P. Warren. 2005. "Prosodic Influences on the Production and Comprehension of Syntactic Ambiguity in a Game-Based Conversation Task." In *Approaches to Studying World-Situated Language Use*, edited by J. C. Trueswell and M. K. Tanenhaus, 209–225. Cambridge, MA: MIT Press.

Schafer, A., S. Speer, P. Warren, and S. White. 2000. "Intonational Disambiguation in Sentence Production and Comprehension." *Journal of Psycholinguistic Research* 29:169–182.

Schepman, A., R. Lickley, and D. R. Ladd. 2006. "Effects of Vowel Length and 'Right Context' on the Alignment of Dutch Nuclear Accents." *Journal of Phonetics* 34:1–28.

Scott, D. 1982. "Duration as a Cue to the Perception of a Phrase Boundary." *Journal of the Acoustical Society of America* 71:996–1007.

Selkirk, E. 1986. "On Derived Domains in Sentence Phonology." In *Phonology Yearbook 3*, edited by C. Ewen and J. Anderson, 371–405. Cambridge: Cambridge University Press.

Silverman, K., M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert, and J. Hirschberg. 1992. "ToBI: A Standard for Labeling English Prosody." In *Proceedings of the 1992 International Conference on Spoken Language Processing*. https://www.isca-speech.org/archive/archive_papers/icslp_1992/i92_0867.pdf.

Snedeker, J., and E. Casserly. 2010. "Is It All Relative? Effects of Prosodic Boundaries on the Comprehension and Production of Attachment Ambiguities." *Language and Cognitive Processes* 25:1234–1264.

Speer, S., and A. Foltz. 2015. "Implicit Contrastive Prosody: Individual Differences in Processing and Production." In *Explicit and Implicit Prosody in Sentence Processing: Studies in Honor of Janet Dean Fodor*, edited by L. Frazier and E. Gibson, 263–285. Berlin: Springer.

Speer, S. R., M. M. Kjelgaard, and K. M. Dobroth. 1996. "The Influence of Prosodic Structure on the Resolution of Temporary Syntactic Closure Ambiguities." *Journal of Psycholinguistic Research* 25:249–271.

Stirling, L., J. Fletcher, I. Mushin, and R. Wales. 2001. "Representational Issues in Annotation: Using the Australian Map Task Corpus to Relate Prosody and Discourse Structure." *Speech Communication* 33:113–134.

Thorsen, N. 1980. "A Study of the Perception of Sentence Intonation: Evidence from Danish." *Journal of the Acoustical Society of America* 67:1014–1030.

Thorsen, N. 1983. "Two Issues in the Prosody of Standard Danish." In *Prosody: Models and Measurements*, edited by A. Cutler and D. R. Ladd, 27–38. Heidelberg, Germany: Springer.

Trofimovich, P., and W. Baker. 2006. "Learning Second Language Suprasegmentals: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech." *Studies in Second Language Acquisition* 28 (1): 1–30.

Ueyama, M., and S.-A Jun. 1998. "Focus Realization in Japanese English and Korean English Intonation." *Japanese-Korean Linguistics* 7:629–645.

Venditti, J. 1997. "Japanese ToBI Labeling Guidelines." *Ohio State University Working Papers in Linguistics* 50:127–162.

Venditti, J. 2005. "J_ToBI Model of Japanese Intonation." In Jun 2005c, 172–200.

Wightman, C., S. Shattuck-Hufnagel, P. Price, and M. Ostendorf. 1992. "Segmental Durations in the Vicinity of Prosodic Phrase Boundaries." *Journal of Acoustical Society of America* 91:1707–1717.

Wong, W. Y. P., M. K. M. Chan, and M. E. Beckman. 2005. "An Autosegmental-Metrical Analysis and Prosodic Annotation Conventions for Cantonese." In Jun 2005c, 271–300.

Yoo, H.-J., and S.-A. Jun. 2016. "Distribution of Accentual Phrase-Medial Tones in Seoul Korean." Poster presented at the Acoustical Society of America, Honolulu, Hawaii December 2.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data is available.
Names: Barnes, Jonathan, 1970– editor. | Shattuck-Hufnagel, Stefanie, editor.
Title: Prosodic theory and practice / edited by Jonathan Barnes and Stefanie Shattuck-Hufnagel.
Description: Cambridge, Massachusetts : The MIT Press, 2022. | Includes bibliographical references and index.
Identifiers: LCCN 2021000764 | ISBN 9780262543170 (paperback)
Subjects: LCSH: Prosodic analysis (Linguistics)
Classification: LCC P224 .P739 2022 | DDC 414/.6—dc23
LC record available at https://lccn.loc.gov/2021000764