Tone inventories and tune-text alignments

Prosodic variation in 'hybrid' prosodic systems

Shelome Gooden, Kath-Ann Drayton and Mary Beckman
Department of Linguistics, University of Pittsburgh / University of the West Indies / Department of Linguistics, Ohio State University

The "hybrid" prosodic systems described for several Caribbean creoles challenge typologies that dichotomize between "intonation languages" and "tone languages" or between "stress" and "pitch-accent" languages. A more nuanced differentiation emerges if languages are compared in terms of questions concerning tone inventory and tune-text alignment, such as: Are the tunes of short utterances composed primarily of tone patterns specified to contrast words or of intonation patterns that are morphemes in their own right? What determines tune-text alignment at the lowest levels of the prosodic hierarchy? Should tones be anchored to rhythmically prominent syllables within focused constituents? This paper explores these questions for several languages with "hybrid" prosodic systems including some where the hybrid nature cannot plausibly be attributed to language contact.

1. Introduction

The prosodic systems of the Caribbean creole languages are interesting because they show a rich variety of patterns and include some languages that have been described as being "hybrid" or "mixed" types in one way or another. For example, Good (2006) describes Saramaccan as "hybrid" in the sense of having two clearly differentiated lexical strata. One stratum includes words of identifiably West African origin and is characterized by *lexical tone contrasts* similar to those observed in these languages, with the tone sequence on the syllables of any individual word corresponding systematically to the tone sequence on the cognate word in the substrate language. The other stratum includes words of identifiably European origin and is characterized by *accent contrasts* whereby a culminative high tone (an "accent") occurs on the tone-bearing unit that corresponds to the syllable with primary stress in the cognate word in the lexifier language. This kind of hybrid

Studies in Language 33:2 (2009), **354–394**. DOI 10.1075/sl.33.2.07goo
ISSN 0378–4177 / E-ISSN 1569–9978

system reflects the hybrid ancestry of the lexicon, and is very clearly the result of language contact. A rather different type of stratified or "hybrid" lexicon can be seen in Remijsen and van Heuven's (2005) description of the Curaçaoan variety of Papiamentu. In this case, there are lexical contrasts involving both differences in tone pattern for some minimal pairs and differences in accent placement for others, with the differences depending primarily on morpho-syntactic category and word length, regardless of the source language. Moreover, "accent" in this case does not mean placement of a culminative tone, but instead is a clear rhythmic prominence — i.e., stress. Rivera-Castillo (2006) describes phrasal patterns which suggest a comparably "mixed" system for the lexicon of the Aruban variety of Papiamentu as well. Finally, the Caribbean creole languages together also constitute a "mixed" system in yet another way. Specifically, alongside the two types of systems which mix "tone" and "accent" in the lexicon and phrasal phonology, there are also varieties which have been described as having "pure" accent systems. For example, although Lawton (1963) and Sutcliffe (1986) discussed Jamaican Creole (JC) in ways that have sometimes been interpreted as claiming that JC is "tonal" in the same way that Papiamentu is, Gooden (2003) shows that the contrasts involve intonational morphemes that depend on discourse function and are linked to stressed syllables in the same way that "pitch accents" are linked to stressed syllables of words under focal prominence in many dialects of English, Dutch, and several other languages.

These differences, both within and across the prosodic systems of Caribbean creoles thus present interesting challenges for research in prosodic typology. The most general statement of these challenges is that we need to be cautious in how we account for the variability that we observe. We want to know how we can account for the two types of variability across the lexicon in the Saramaccan and Papiamentu cases and also the variability across the different varieties when creoles such as Jamaican Creole are juxtaposed against creoles such as Saramaccan, which are both English lexicon creoles. Given what we know about the grammatical systems of these languages (see the GRR papers in this issue), it might seem plausible to suggest that the differences between the languages, like the variability within the Saramaccan lexicon, are the result of contact-induced change or substratum influence rather than of internal development or universal principles. That is, since many of the Caribbean creole languages are historically related and have similarly emerged out of contact between West African lexical tone languages and European accent languages, the typologist's first impulse may be to attribute the prosodic mixture to the original contact between West African languages and European languages (i.e. to processes of creole formation). In this paper, however, we will argue that this impulse needs to be curbed, pending a better understanding of the rich dimensions of variation observed in prosodic systems of non-creole languages

that have been studied in enough detail to note differences across contexts within any particular language variety and across varieties within any particular language family. That is, we will suggest that whenever a phonetically-grounded study is done with the kind of meticulous control of contexts that allows the linguist to go beyond surface appearances to tease apart the lexical and phrasal contributions to the tunes and rhythms of utterances, the language under study looks to have a "mixed" system. Therefore, taking a closer look at the different types of "mixed" systems that have been described for the better-studied non-creole languages can help to illuminate the question of origins of such mixed systems in the Caribbean creoles.

In this paper, then, we will review the many dimensions of variation that are evident across languages that we ourselves have studied. We will define concepts such as *tone*, *accent*, and *stress* within a framework that does not try to reduce these terms to the surface phonetic parameters that typically realize them, and we will illustrate them with examples from language groups that show a rich variety of prosodic variation across closely related language varieties. All of the languages that we will draw on are themselves "hybrid" in ways that are reminiscent of the situations described for Saramaccan, for Papiamentu or for the Caribbean creoles as a whole. However, the variation seen in many of these other languages cannot plausibly be attributed to the kind of language contact that gave rise to the two strata of the Saramaccan lexicon for example. Also, we will describe aspects of the prosodic systems of two typologically related Caribbean English creoles, JC and Trinidadian English Creole (TEC). We will use these two creole languages to show that variation is possible even within a particular prosodic type (in this case, prototypically "stress accent" systems) and suggest that at least for TEC this variability is due to much more recent situations of language contact than the ones that gave rise to the group of Caribbean creole languages.

We will conclude by suggesting that a more fruitful approach to understanding prosodic variation in Caribbean creoles is to acknowledge that prosodic systems themselves might be inherently "hybrid" because of the multiple functions that prosody typically plays in organizing utterances for the speakers and hearers of a language. A better understanding of the variation seen across and within languages thus depends on a shared understanding of these functions. Therefore, the sections of this paper are centered around concepts that help us understand these organizational functions of prosody. Specific challenges for the typologist are presented as they arise in the discussion.

The paper is organized as follows. Section 2 gives an overview of the Autosegmental Metrical framework of intonational phonology in order to make it clear what kinds of prosodic parameters might plausibly be affected by language contact. Section 3 discusses the importance of grounding our descriptions of prosodic

structure in a thorough understanding of the non-tonal cues to metrical structures and to the alignment of autosegments in general. Section 4 discusses the treatment of tone in the AM framework and Section 5 the function of tone across languages. Section 6 discusses the concept of 'tone language' in relation to the function of tone and the density of tone specification in languages. Section 7 examines how phonetic properties are manipulated to cue prosodic prominence in different languages. Section 8 discusses the concept of 'stress language' and describes some typical phonetic cues to stress prominence at different levels of the prosodic hierarchy. Section 9 looks at cross-linguistic differences in tonal association and how this relates to a classification of tone vs. stress language and Section 10 is the summary of the main arguments presented vis-à-vis prosodic variation and the classification of the prosodic systems of creoles.

2. A unified Autosegmental-Metrical account

The general framework that we adopt is the *Autosegmental-Metrical* (AM) account of tone and its relationship to prosodic organization. This name for the framework was coined by Ladd (1996), and it has been widely adopted in the subsequent literature on intonation and phrasing in many languages (see Gussenhoven 2004 for a recent review, and Jun 2005, for a collection of recent papers). Although the framework was developed first in describing intonation patterns in languages such as Swedish, English, and Italian (see, e.g., Bruce 1977; Pierrehumbert 1980, and Grice 1995), Pierrehumbert and Beckman (1988), Hayes and Lahiri (1991), Jun (1996), Gussenhoven and van Vliet (1999), and many others quickly showed that it is more generally applicable to the description of utterance tunes and their relationship to prosodic organization across a wide variety of languages.

A key fact to understand about this framework is the way in which it distinguishes between autosegmental content features and the metrical structures that license them. For example, characterizing the European lexifier stratum of Saramaccan in terms of *accent* as described above makes sense only if we have both the concept of a content element "high tone" (in opposition, say, to "low tone") and the separate concept of a structural domain (such as a prosodic word or larger intonational phrase) within which only one such element can occur. In a similar vein, Beckman and Edwards (1994) suggest that the concept of *stress* cannot be made coherent in a framework that does not separate the function of regulating content features so as to differentially impart prominence to some position (e.g., aligning a [+spread glottis] gesture in a particular way with some consonant within a prosodic word, or aligning a high tone with a particular syllable in an intonational

phrase) from the function of opposing prominent and weak positions via such regulation. The two halves of the term AM refer to this separation of functions.

Metrical. More specifically, then, one aspect of understanding the framework is to understand clearly what is meant by the *Metrical* part of the name for the framework. This term refers to the intersecting rhythms or layered prosodic constituents (segment, syllable, foot, etc.) projected from the content specifications (e.g. tones, consonant manner features) that are licensed at different positions. For example, we know that in all spoken languages, the low-level prosodic constituent "syllable" (abbreviated as σ) is composed of consonants (C) and vowels (V), which in turn are defined in part by content features such as the contrast between sonorant sounds such as [j] or [n] and obstruent sounds such as [ʃ] or [t]. These content features are licensed to occur in different positions of the syllable typically based on relative sonority. So more sonorant content features are licensed to occur at a syllable's head position (its nucleus) and less sonorant features are licensed to occur only at the edges (onset and coda). Generally speaking then, in many languages a very sonorant consonant such as [j], [l], or [n] can occupy the head position in a syllable, as long as there is a less sonorous edge consonant such as [ʃ], [t] or [b] separating this head element from the closest more sonorous neighboring autosegment. In American English for example, a syllable with a minimally sonorant head (e.g., the syllabic [l] in the second syllable of *apple* or the [ə] in the first syllable of *apply*) must group together with a syllable that has a more sonorant head segment (e.g., the inherently long, very open [æ] vowel in the first syllable of *apple* or the long typically diphthongal vowel in the second syllable of *apply*). This grouping also defines two higher-order prosodic constituents, typically the "foot" (abbreviated F) and the "prosodic word" (abbreviated ω), both of which are illustrated in Figure 1 using the words *apple, apply* and *peel* (for English) and *botn, Bantan* and *gyaadn* (for Jamaican Creole). Figure 1 shows the metrical structures that many researchers working within the AM framework would posit for these word pairs in these two language varieties.

Autosegmental. Given this understanding of metrical structure, the *Autosegmental* part of the name for the framework, then, refers to the specification of content properties that are autonomously segmented — i.e., that project as strings specified on independent tiers rather than being bundled together (cf. Chomsky and Halle 1968), in association to the metrical positions that license them in words. For example, in JC changing specifications of place and/or manner features can define as many as six autosegments for consonants for the monosyllabic prosodic word *bungks* [bʌŋks] 'to bounce; rubber band game'. There are changes in bilabial, velar and coronal features and changes in the degree of aperture/airflow giving a sequence of complete closure and stopping of air, changes in stricture for nasal airflow, complete closure again and modified aperture for sibilant airflow.

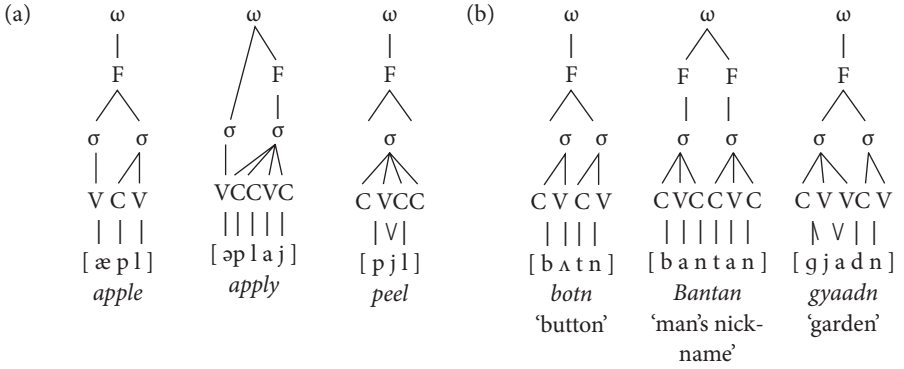

Figure 1. Contrasting word-internal prosodic structures illustrated with (a) American English and (b) Jamaican Creole

This means that in the phonological grammar of JC, place and voice features (like other features that readily undergo "assimilation processes") should be represented on a different autosegmental tier from manner features (Goldsmith 1990). This projection of the features onto different tiers allows us to account for the phonotactic constraints on syllable structure. For example, while the sequence $_{\sigma}$[tn can be tautosyllabic, n] $_{\sigma}$[t_must be heterosyllabic. Also, while $_{\sigma}$[gj is a possible onset in JC, $^{*}{}_{\sigma}$[jg is not.

Association. A third aspect of the framework has to do with the relationship between an autosegment and the metrical position that licenses it; this abstract structural property is *association.* Differences in the association of the autosegment to the metrical structure yield differences in the interpretation of the resulting structure. This is illustrated in Figure 2 with a 'non-speech' analogy, to make the notion clear. The "segmental features" [slash] and [dot] allow the reader to interpret the "metrical structures" for [square] and [diamond]. Note the "phonological" ambiguity, such that the [dot] "edge autosegment" of [square] can be misparsed as the "head" of a [plus sign] "metrical structure" in (c). Structural constraints on where these segmental features can occur can be viewed then in terms of how they allow for defining the different shapes ("metrical structures") seen in (a) through (f). The analogous ambiguity in the grammar of speech is seen with a segment like [j], in that the palatal constriction which results in an extremely high second formant value, can be parsed either as the vowel [i] so that it is associated with the head of a syllable (as in *peel* in Figure 1a), or as a consonantal feature that is associated only with the edge of the syllable (as in *apply* in Figure 1a or *gyaadn* in Figure 1b). The appropriate interpretation thus comes through parsing the structure. These structural differences are language specific but, as is discussed below, can also be due to a misparse of the structure by the analyst. In much the

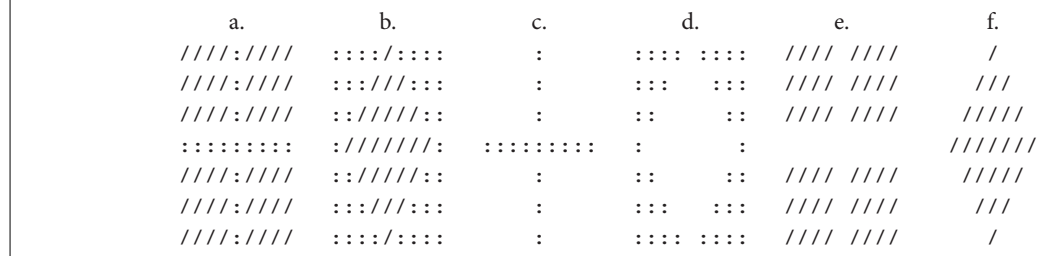

Figure 2. Three different grammars for associating [slash] vs [dot] texture autosegments to head vs edge positions in the metrical structures for [square] vs [diamond] (adapted from Beckman, 1996).

same way, the prosody of an utterance is a structure which is open to interpretation based on listeners' perceptions, which means that similar misparses of prosodic structure are possible.

The grammatical phenonemon of association thus presents us with the **first challenge** in describing the prosodic systems of creole languages. The kinds of parsing ambiguities illustrated in Figure 2 are especially likely in situations of language contact, where speakers of languages with similar inventories of autosegments but different association grammars can parse the content features completely correctly but misinterpret (or reinterpret) the metrical structures that they cue. This should be a cautionary note for typologists and creolists alike since field linguistics is typically a situation of language contact. As Hyman (2006:230) notes, "The question that is relevant in this context is whether we want to typologize according to the properties of prosodic systems or according to the analyses given to them by diverse linguists." Our understanding of how listeners resolve prosodic ambiguities and also of how they recover elided autosegmental features is crucial for our understanding of variation in phonological structure above the word.

In the following two sections, we will further amplify on how the segmental and tonal markings of prosodic structure are understood within this AM framework, highlighting common types of ambiguity that can lead to misunderstandings among linguists and also, more critically, to different parses among speakers. In some cases, the ambiguities are intrinsic to the long-term stock of word forms in a language, so that different speakers can (completely unawares to each other) have different parses for identical productions, in a way that has no lasting consequences for language transmission from one generation of speakers to the next. In other cases, the different parses can affect the speakers' (re)productions of each others' forms in ways that can lead to language change.

3. The phonetic grounding of segmental rhythms

In the preceding section, we described some constraints on the types of consonant autosegments that can occupy head positions in syllables of American English and JC as a way to introduce the notion that Autosegmental content is licensed by Metrical position. In this section, we will acknowledge and amplify on an unspoken problem with this description. The problem is that the technical terms *segment*, *consonant* and *vowel* are ambiguous. We can use them to refer to Metrical positions, as we did when we talked about [p] functioning as an "edge consonant" for the second syllable of *apple*. We can also use them to refer to the Autosegmental content features that are aligned with respect to those positions, as we did when we talked about the "sonorant consonant" [l] that is licensed to occur at the head position in this syllable. Moreover, since the Autosegmental content features are themselves grounded in language-universal physics and psychophysics, we can also use these terms to talk about the relevant phonetic properties that ground the Metrical structures, as we do when we say that in some dialects of American English, the [l] of *apple* is "vocalized" and realized as something more like the "vowel" [u].[1]

Acknowledging this three-way ambiguity is critical for understanding rhythmic differences across languages, such as the "merely tonal" accent system of Saramaccan versus the "stress" accent system of Curaçaoan Papiamentu. That is, in order to talk about the prosodic differences between these two creole varieties, we need to understand the phonetic grounding of those basic prosodic units — the traditional consonant and vowel segments that are named both by IPA symbols and by the C and V terminal nodes in the AM model of a spoken utterance — because consonant and vowel segments are what we might call the "leaf units" that define the rhythmic alternation we recognize as the prosodic shape of words. In many conservative creole languages this translates to a fairly regular rhythmic alternation between consonants and vowels as in Sranan *rata* 'rat' and *pusu* 'push'. In most other varieties, however, more complex syllable structures are more common. Whenever syllable structures are more complex, varying *alignment* patterns are required for the same set of autosegments in different metrical structures.

By *alignment* we mean the temporal relationship between the dynamics of autosegments from different tiers that are (partially) co-produced because of association to the same metrical position. For example, the glottal opening and closing gestures for the [-voice] autosegment align with the oral gestures for the sequence of place and manner features for the labial obstruent and lateral approximant in different ways in the English words *apply* and *plight* (where the /p/ marks the beginning of a foot or a prosodic word) compared to *apple* (where the /p/ marks a foot-medial syllable boundary), as shown in Figure 3. Note here the parsing ambiguity of the configuration in *polite*, where most speakers pronounce a "syllabic" [l],

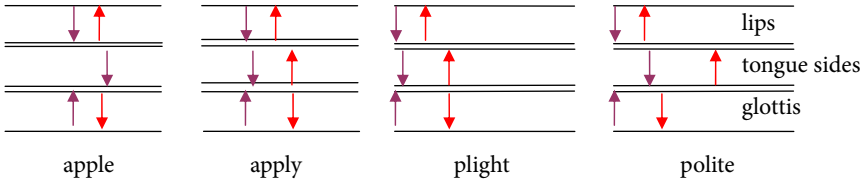

Figure 3. Alignment patterns for labial closure (down arrow) vs release (up arrow) gestures, lateral opening (down arrow) vs seal (up arrow) gestures, and glottal abduction (up arrow) vs adduction (down arrow) gestures for [pl] sequences associated to 4 different metrical structures.

making the word [pl̩.lait], as shown in the figure, but some speakers re-interpret the temporal delay between the "lips open" gesture of the [p] and the "tongue sides down" gesture of the [l] as an epenthetic reduced vowel, pronouncing the word as [pə.lait].

Sonority scaling is also important in discussing the phonetic grounding of segmental rhythms and the way that these rhythms are altered to mark different prosodic positions. By sonority, we mean that the manner feature autosegments that mark syllable structure are ordered on a continuum from the most open vowels to the most obstruent-like consonants. The properties that contrast the category types along this continuum from most to least sonorant segments can be "scaled up" or "scaled down" to help cue metrical structures above the syllable. For example, in many dialects of English, /p, t, k/ have longer closures and a pronounced interval of aspiration when they associate to the beginning edge of a foot or a prosodic word. The aspiration that distinguishes this manner class from the next higher class on the sonority scale (namely, /b, d, ɡ/) can be lengthened and intensified in words such as *plight* in Figure 3 where the [p] associates to the edge of a word that is also the edge of a foot, and so is "stronger" than the [p] in *apply*, which begins a foot but is not word-initial. This lengthening and intensification can be exaggerated further to mark the word as the prominent head of an intonation phrase (the nuclear stress) when the discourse context demands this (see Docherty 1992, inter alia). Also, English autosegments that can associate both with edge and head positions in syllables, such as the [l] in *plight* vs *polite*, show differences that correlate with differences in association. So [l]s are typically longer when associated to the head or coda positions of a syllable (see, inter alia, Price 1980; Browman and Goldstein 1988) and the component tongue tip and tongue body gestures are coordinated differently, allowing the more "vocalic" tongue body constriction to dominate in the filter, making for the "dark" allophone. Moreover, just as the contrast between more and less aspirated voiceless stops is continuous and gradiently related to the position in the prosodic hierarchy, the contrast between more consonantal and more vocalic variants of [l] is continuous (Sproat and Fujimura, 1993).

The sonority scaling of the American English [l] provides a first example of the ways in which structural ambiguity can lead to sound change. In some dialects, the different gestural coordination patterns between the apical and dorsal gestures that allows the dorsal gesture to dominate in syllable head and coda positions has been parsed as an absolute dominance of the more vocalic gesture, leading to "vocalized" [w] or [u] pronunciations. That is, the continuum of contrasts between more consonantal [l] at the beginnings of prosodic constituents and more vocalic [l] at the ends of prosodic constituents has been dichotomized and reinterpreted in terms of a phonemic split between a more central set of vowels in [kʉw] *coo* and [bəʊ] *bow* and a new set of back vowels in words such as [kuw] *cool* and [bow] and *bowl* (see, e.g., Dodsworth, 2005). The JC pronunciations of words such as *gyaadn* 'garden', *kyat* 'cat', *kyabidj* 'cabbage', and *gyaalik* 'garlic' are another analogous example, where the palatal release and CV transitions before the stressed low front vowel has been reparsed as a distinctive feature to be aligned with onset stop itself, in minimal pairs such as [kʲat] 'cat' versus [kat] 'cot' (Dyer and Wassink 2001) or [gʲaadn] 'garden' versus [gaadn] 'Gordon' (Irvine, 2004). This sound change is rooted in the reinterpretation of alignment patterns for a [j]-like onglide in both the low and the mid front vowel (i.e., the [ie] vowel) in words such as [giet] 'gate'.

Since this section discusses the phonetic grounding of word-level metrical structures in segmental alignment patterns, a note on phonetic representation is in order. In Figure 3, we have used idealized drawings ("gestural scores") to represent relevant aspects of the scaling and alignments of autosegments associated at different metrical positions of syllables and feet. This choice was dictated entirely by the complexity of exegeses for other representations such as ensembles of lip aperture traces measured from video recordings, linguo-palatal contact patterns measured using dynamic electropalatography and acoustic landmarks identified in spectrograms. We do not mean to imply that gestural scores are the only possible phonetic representation, or that articulatory representations are primary. Similar considerations dictate our choice of representation for another important element of any discussion of prosodic systems, the fundamental frequency (F_0), which we take up in the next section.

4. Tone in an Autosegmental-Metrical account

Tone features, like segmental features, are autosegments that are crucial to our understanding of prosodic structures above the word. We will use the fundamental frequency (F_0) contour as a phonetic representation of these autosegments. **Tones** are the local specification of pitch features as relatively short autosegments (about syllable-length), which are themselves independent of the vowel and consonant

autosegments with which they are aligned. In the AM approach, tones are treated as such regardless of their function in individual languages. We discuss further below how we can distinguish between the different functions of tone found across languages.

Like the manner specifications that give rise to the sonority scale, tone specifications are inherently arranged in a scale from low (L) to high (H) corresponding to target pitch values that speakers produce at different prosodic positions in an utterance. As with manner autosegments, tones are licensed by association to different positions in the metrical structure. Tones differ from other types of autosegments, however, in typically associating to a wide range of metrical positions across and within languages. (We will call these positions *association points*.) Like vowel features, for example, tones can associate to syllables or moras (i.e., head positions in syllables); when these syllables or moras also have a special status within a larger unit, the associated tones are the starred tones for that larger unit. Tones can also associate to the edges of higher-level constituents above the word such as the *accentual phrase* (AP) of Korean, French or TEC and the *intonational phrase* (IP) of English or Cantonese.

A tone or sequence of tones marking the edge of a higher-level constituent in this way is called a *boundary tone*. For example, in the TEC examples in Figure 4, the declarative in panel (a) is marked at the right edge by a low boundary tone (abbreviated as L%), and the polarity (yes-no) question in panel (b) is marked by a high boundary tone (abbreviated as H%). In both cases they mark the end of an intonational phrase.

Although Figure 4 illustrates IP boundary tones in utterance-final position, the IP boundary can also occur utterance internally, as in Figure 5. Here the longer utterance which ends with a low boundary tone also has two internal IP boundaries marked (in this case) by high boundary tones (H%) on the words *standard* and *to.* There are also phrase tones, which like boundary tones mark the edges of higher-level prosodic constituents. In all three of the utterances in Figures 4 and 5, there is an accentual phrase constituent (AP) below the IP, that is marked by an H tone. This H seems to be a phrase-level tone that is licensed by the AP constituent edge, but it also seems to be associated in some way to the head of the prosodic word, so that its alignment domain is larger. That is, the high pitch is realized over many syllables when the other tone in the AP is at distance from the edge, in a way that has been called *interpolation* or *spreading* in different traditions within the field.2

The tone following the H of the AP exemplifies a different kind of association point. In each of the APs, the H is followed by a L* tone which associates to the stressed syllable of the head word (e.g., ***ca****labash* and *wa****ter*** in the two APs in Figure 4). Starred tones, which are tones associated with stressed syllables in this

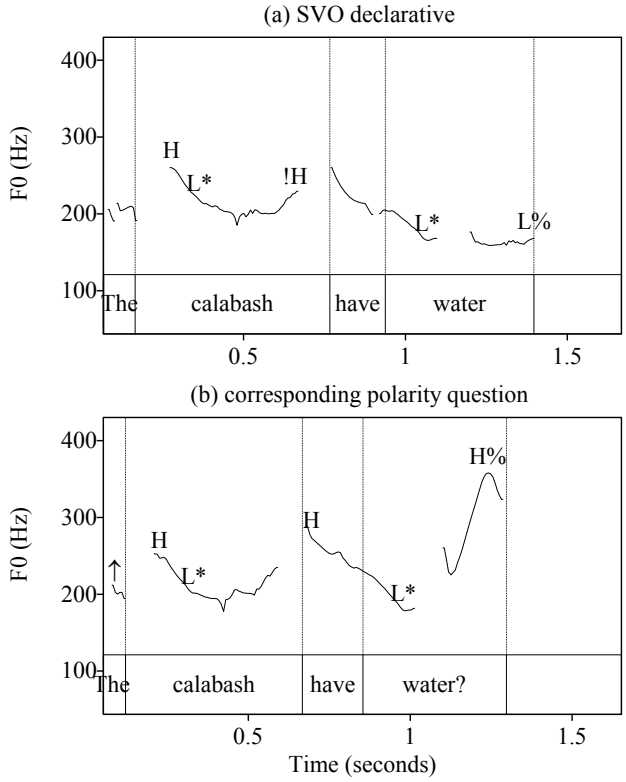

Figure 4. F_0 contours of TEC broad focus declarative (a) *The calabash (gourd) has water (in it)* and polarity (yes-no) question (b) *Does the calabash (gourd) have water (in it)?*

way, can be called *pitch accents*. More generally, the star ("*") diacritic is used to differentiate such head-marking association points from the "pure" edge-marking association points designated by the "%" diacritic for boundary tones. The fact that AP-level H does not have either diacritic in these figures indicates the ambiguity (or possibly bivalent nature) of its association point(s).

Another common case when diacritics such as "*" and "%" are not used to designate association points is in languages with very dense distributions of tone. In these languages, one particular low-level prosodic constituent or prosodic position is designated as the "minimum tone-bearing unit" (TBU) — i.e., the smallest prosodic constituent that licenses the occurrence of a tonal specification. This term is particularly useful when the grammar of the language dictates that each constituent at this level must have an associated tonal autosegment. So in Cantonese, where the TBU is the syllable, every syllable has a specification for at least one tonal autosegment.[3] Cantonese is an example of a maximally dense distribution of tones relative to TBU. Most other Sinetic languages have sparser distribu-

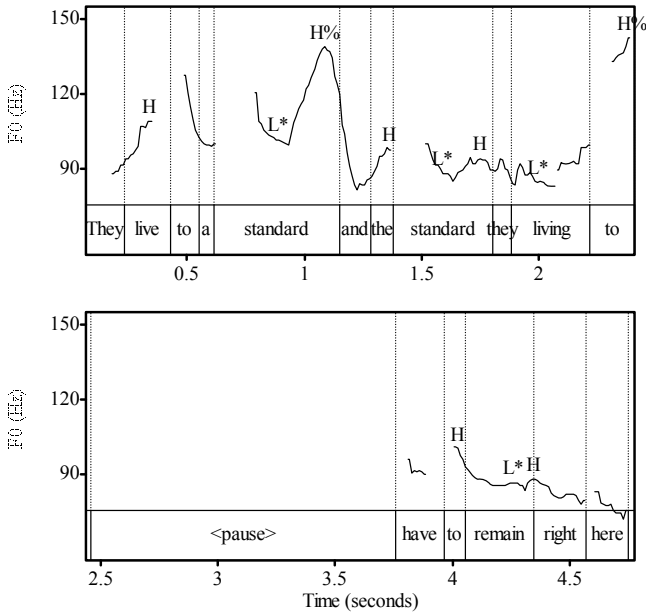

Figure 5. F_0 contours of TEC declarative showing sentence internal IPs (Indo-Trinidadian speaker, recorded 1970s)

tions, with Beijing Mandarin representing the opposite extreme of many "neutral tone" syllables (see, e.g., Kratochvil 1998, and other descriptions cited in Peng et al. 2005).

In addition to such local tone specifications anchored at particular association points, grammars also specify pitch features for larger constituents, typically referring to parameters of the *pitch range* — i.e., top and/or bottom values of the tone scale as a whole. For example, many languages mark polarity (yes-no) questions by shifting the entire scale upward over the last (or only) IP in an utterance, so that L tones can be higher than H tones earlier in the utterance. This is seen in the JC example in Figure 6a. The polarity question is produced in a high overall pitch range (compare the declarative in Figure 6b) and the pitch is raised even higher still for the terminal rise of the polarity (yes-no) question intonation.

The pitch range can be compressed also, by shifting the top of the range downward at some fixed point in a prosodic constituent, i.e., *downstep*. For example, in standard (Tokyo) Japanese, this downstep occurs at each lexically specified HL, so that each subsequent H tone is lower than the H of the triggering HL sequence

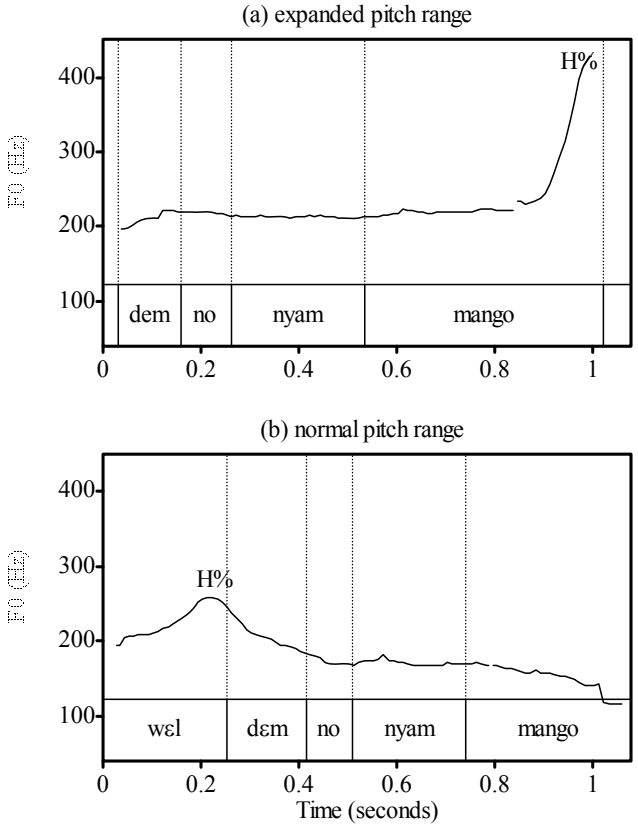

Figure 6. F_0 contours of JC utterances in raised pitch range, *Don't they eat mangoes?* (6a) and normal pitch range, *They don't eat mangoes* (6b).

until the pitch range is reset to mark the beginning of the next intonational phrase, (Figure 7).

In the Japanese example in Figure 7, downstep is not explicitly tagged, because it is predictable from the accentual phrasing and the distribution of accents. However, languages can also specify downstep as a contrastive feature for pragmatic effect. In the TEC utterances in Figure 8, for example, the H on *does* is downstepped (tagged with the "!" diacritic as !H) relative to the previous high in both the declarative and the polarity question, to express the subordination of this word in the information structure of the utterance.

Long-range pitch range manipulations such as the downstep in Figure 8 are just one way in which the differences in F_0 values among H and L sequences can signal prominence relationships. Another mechanism involves *tone scaling* within the local pitch range. That is, as with the sonority scale, the relevant acoustic prop-

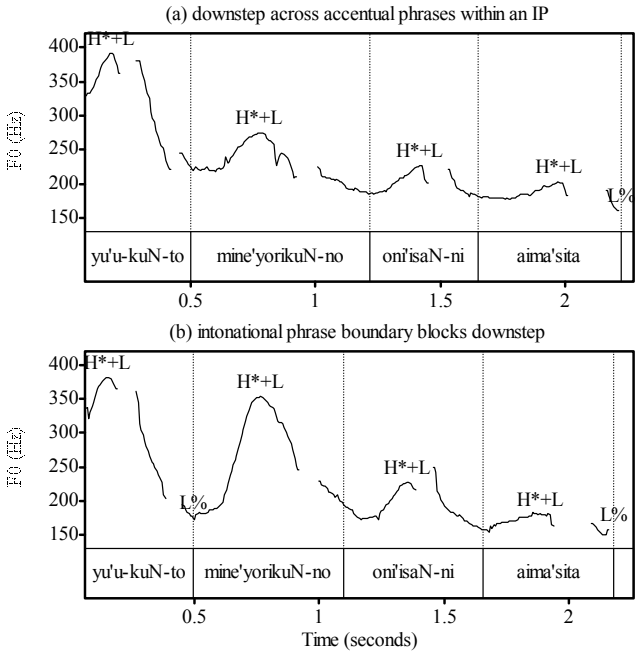

Figure 7. F_0 contours for two utterances of the sentence *Yu'u-kun to Mine'ori-kun no oni'isan-ni aima'sita*. which can be parsed as 'I met with the older brother of Yuu and Mineyori' (one person) or 'I met with Yuu and the older brother of Mineyori.' (two people), depending on whether the utterance is produced as one intonational phrase (IP), with three occurrences of downstep (7a), or as two intonational phrase, with downstep blocked at the IP boundary after *Yuu-kun-to* (7b) (example from Eda 2004).

erty that defines the tone scale (i.e., F_0) can be manipulated to cue higher-level metrical structure. For example, in standard (Tokyo) Japanese a L that marks the edge of an IP is lower in the local pitch range than the L that marks the edge of an IP-medial AP.

We ended the preceding section by describing two cases where ambiguities in the prosodic function of sonority scaling effects for some segmental property have been reinterpreted prosodically, leading to sound change and prosodic differences across related language varieties. That is, we described [l]-vocalization in certain Midlands dialects of American English in terms of reinterpretations of the effects of scaling the sonority of the dorsal constriction in American English [l] toward the more vocalic end of the continuum. A reinterpretation in the opposite direction has led to the emergence of a contrastively palatalized dorsal stop in the case of stressed historic [a] in JC. The comparable cases of ambiguity and reinterpretation of tonal autosegments are ubiquitous. They are easier to describe and understand, however, if we first discuss the functions of tone association and tone scaling.

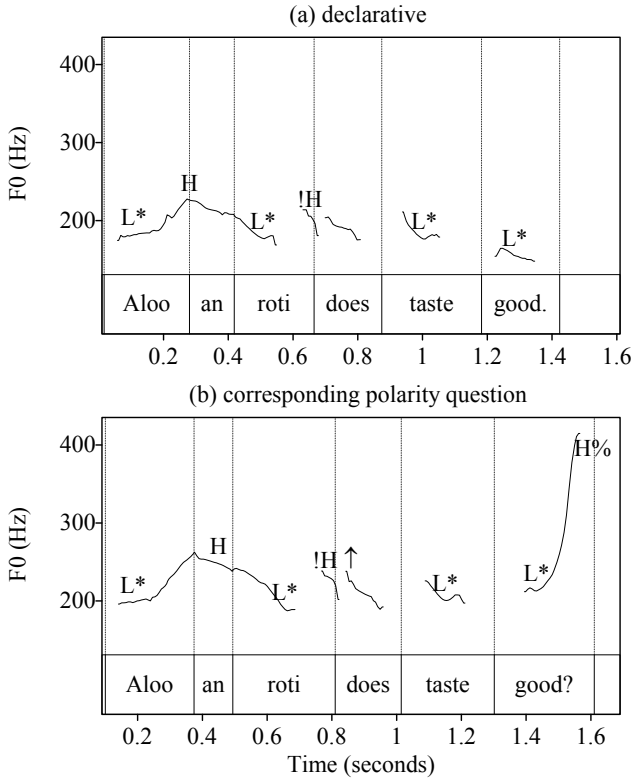

Figure 8. F_0 contours of TEC utterances *Aloo and roti taste good* (a) and *Do aloo and roti taste good?* (b).

5. The morphosyntactic functions of tone

It is crucial to understand that all of these aspects of tone can be used for different morphosyntactic functions, from the categorical choice of local tone specifications and their association points to the continuous manipulation of global pitch range or local pitch scaling value. Moreover, this variation in morphosyntactic function can occur both across and within languages and represents another way in which prosodic systems can be said to be 'hybrid'. For example, as noted in the preceding section, in standard (Tokyo) Japanese the scaling of L boundary tones differentiates the IP boundary from the AP boundary, while in standard Mandarin Chinese, tone scaling is more saliently involved in lexical contrast, with the L of tone 3 ("dipping tone") being phonetically lower than the L of tone 2 ("rising tone").[4]

Figure 9 illustrates contrasting morphosyntactic functions for the choice of tone and tone association point within a language. In these utterances, the rise at the beginning of each of the four lexical words is a fixed LH sequence that is insert-

ed by the phrasal phonology and functions only to mark the beginning of the accentual phrase. The inflection point that is tagged with the L boundary tone (L%) near the end of the sentence also is a fixed shape that marks the end of the accentual phrase. The H% versus LH% tones that mark the end of the intonational phrase, by contrast, are taken from a rich inventory of pragmatic (tone) morphemes that convey such discourse functions as surprised echo question in Figure 9 (top) versus emphatic confirmation of old information (bottom), as described by Venditti, Maeda and van Santen (1998), among others. These example utterances are from Eda (2004), who shows that L2 speakers and L1 Japanese speakers of other dialects have difficulty interpreting these tunes.

Tokyo Japanese also assigns a culminative HL tone sequence (an "accent") to a lexically specified TBU within some APs, and the presence and the location of this TBU (if present) contrasts ordinary lexical words, as illustrated in Figure 10. In other words, the presence of the HL in these words is lexically specified. The first AP (the object NP) in each of the first two sentences in 10 contains an accented

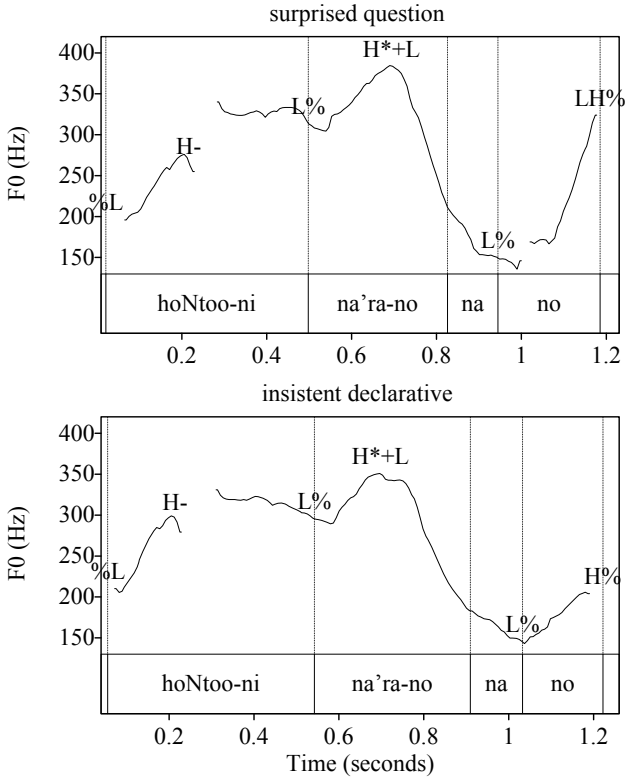

Figure 9. F_0 contours for utterances of the sentence *Hontoo-ni Na'ra-na no* 'Really the one from Nara', contrasting two IP boundary tones in Tokyo Japanese.

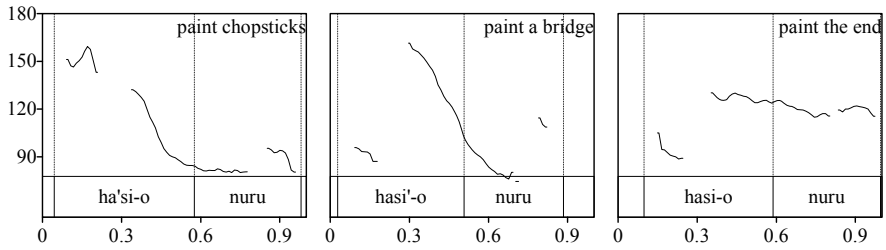

Figure 10. F_0 contours for three Tokyo Japanese sentences that contrast in *akusento* pattern of the object noun AP.

lexical item, but the first AP in the last sentence contains no accented lexical item, hence the lack of downstep on the verb *nuru*. This effect of the lexical tone pattern of a preceding word on the tone shape of a following word also interacts with (and indirectly reflects) the alignment between the prosodic organization of the utterance and other types of linguistic relationships, such as the syntactic structure of the sequence or the larger discourse function. For example, if the discourse context puts narrow focus on the verb, there can be an apparent reorganization of the prosodic structure, so that the verb can be realized in its own expanded pitch range (see Venditti, Maekawa, and Beckman, 2008, for a discussion of the various focus-marking mechanisms). Thus, while the tone pattern on the object noun is in part lexically specified, making Japanese a "tone language" by Hyman's (2006) typology, the tone pattern of the sentence as a whole cannot be described without referring to "post-lexical" tones that demarcate higher order prosodic constituents. The tone pattern on the verb, then is a complex hybrid product of the lexical and phrasal tone specifications. Since the contrasting tone patterns that result resemble patterns with a purely "lexical" or a purely "post-lexical" function in other languages, Japanese tone patterns are prone to opposite patterns of assimilation and misinterpretation by speakers of Mandarin Chinese versus speakers of English.

The fact that speakers of other dialects have difficulty interpreting the tunes in Figure 9 also speaks to the hybrid nature of Japanese tone structures. That is, while all dialects share the core functions of using tonal autosegments to demarcate constituents at two levels of the prosodic hierarchy, the tone patterns are highly dialect specific at both levels. Linguists have long recognized that the lexically-specified aspects of AP-level patterns vary dramatically across dialects (see, e.g., Haraguchi, 1977), but they are only beginning to appreciate the dialect-specific nature of other aspects of the tune (see, e.g., Igarashi, 2007). The H% boundary tone in Figure 9b, for example, seems to be especially characteristic of younger speakers of the Tokyo dialect, who use it utterance medially to mark particular types of focus structures in ways that are conducive to re-interpretation as a "secondary" lexical accent rather than as a boundary tone (see Venditti et al. 2008).

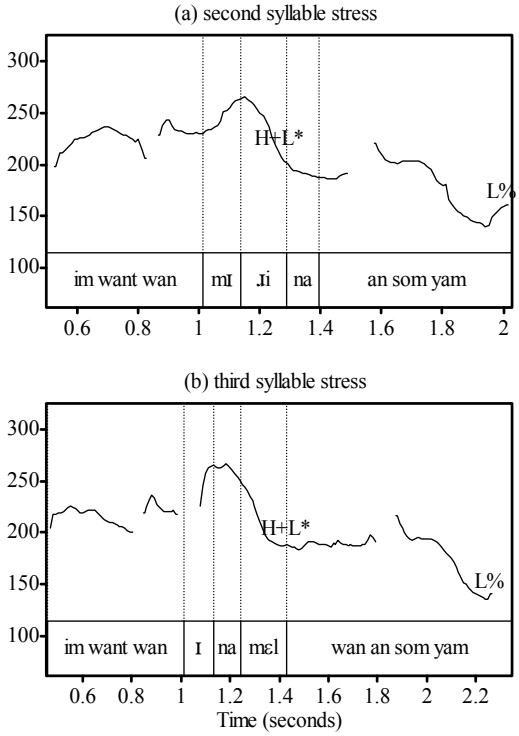

Figure 11. F_0 contours of two Jamaican Creole words that contrast in terms of the location of stress. In the sentences *He wants some undershirts and some yams* (a) and *He wants some enamel ones and some yams* (b).

In comparison, the JC examples in Figure 11 also differ in terms of location of a HL pitch sequence. In this case, however, the location of the HL cues the difference in the metrical structure for penultimate syllable stress in *mi**rii**na* 'undershirt' (a) versus final syllables stress in *ena**mel*** (b); this contrast is not attributable to a lexical specification of tone.

JC also has IP boundary tones marking higher level constituents at phrase edges and utterance internally. In Figure 12 the H boundary tone (H%) has various functions and cues different meanings as a function of its association point within the utterance and of the overall pitch range assigned to the IP that licenses it. The H% at the end of *well* marks the phrase edge in (a) where we have a physical pause marking the phrase break in the F_0 contour as much as it does in (b) where there is only a virtual pause. In both cases the pragmatics is the same, a softened declarative, *Well, they don't eat mangoes.* The H boundary tone in (a) is realized at a higher pitch than in (b) because of the break. In (c) however, the high pitch excursion at the end of the phrase for the high boundary tone (H%) cues a different

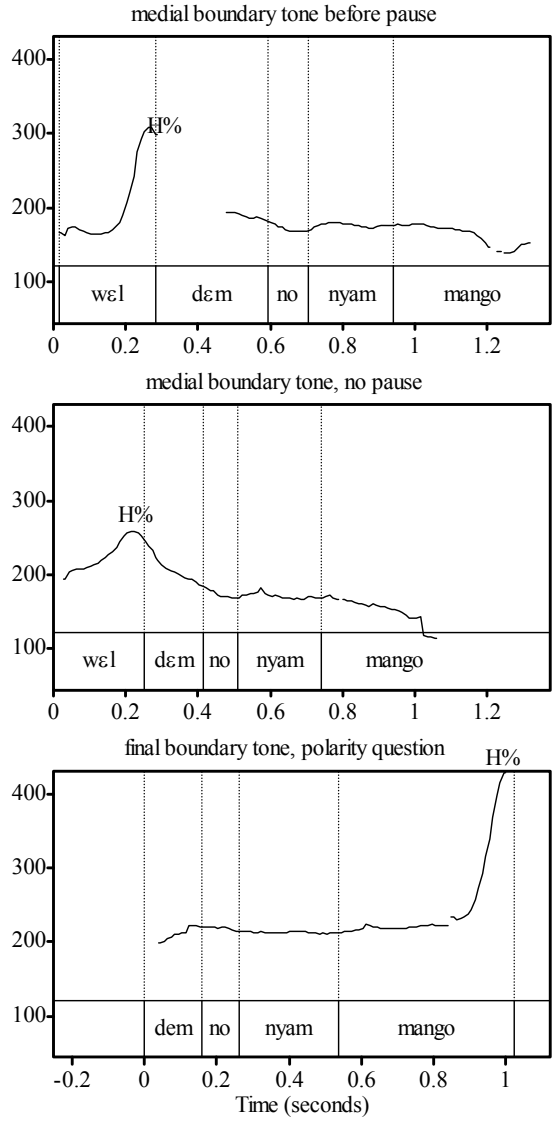

Figure 12. F_0 contours of productions of *Wel dem no nyam mango* 'Well, they don't eat mangoes' and *dem no nyam mango?* 'They don't eat mangoes?!'

meaning, that of surprised question. In all three cases though, the H% marks a phrase boundary.

The languages we have looked at so far share some similarities. Like Tokyo Japanese, JC, and TEC, Beijing Mandarin also has IP boundary tones that are pragmatic (tone) morphemes in their own right, illustrated in Figure 13. Like all

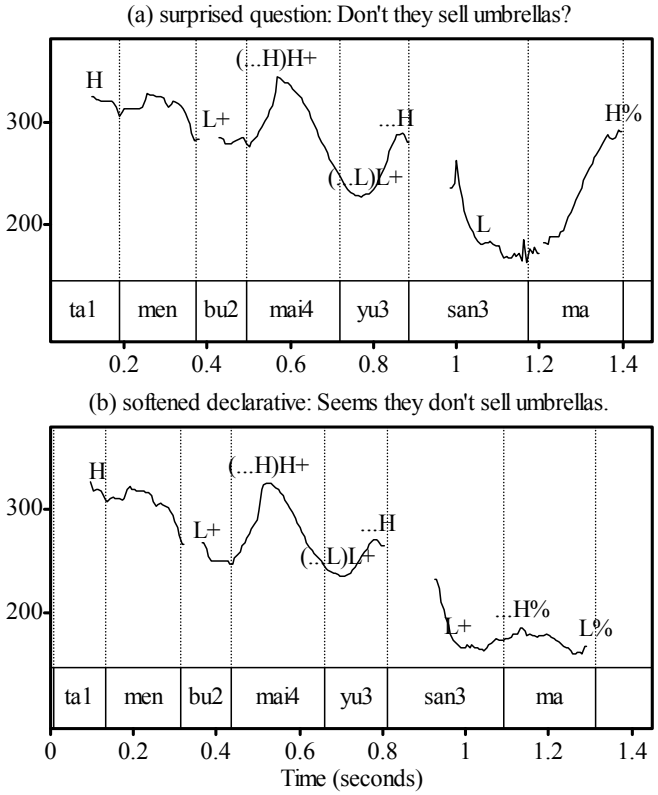

Figure 13. F_0 contours for two utterances of *Ta1-men bu4- mai4 yu3-san3 ma*. 'Don't they sell umbrellas.' produced with contrasting boundary tones by a Beijing Mandarin speaker (from Lee 2000, cited in Peng et al. 2005).

three of the other languages, Mandarin Chinese also has tones that are associated to specific syllables within the phrase, but the location of these association points can be specified lexically. For example, there is no tone specified with the plural morpheme *-men* of *tāmen* 'they' (where *tā* is 'he/she/it'). This is an "unstressed" or "neutral tone" syllable lexically. (We will discuss this further in Section 6.) However, very importantly, unlike in JC and TEC, or even in Japanese, the tone patterns that are associated to syllables in Mandarin primarily have a lexically contrastive function. There are four contrasting tone patterns that function exactly like vowel or consonant features in differentiating lexical words, as illustrated in Figure 14.

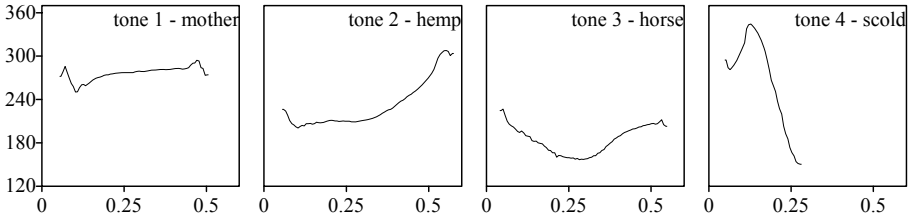

Figure 14. F_0 contours for citation form utterances of four morphemes that share the same [ma] CV autosegments but differ in tone; produced by a female speaker of (Taiwan Guoyu) standard Mandarin.

6. Tone function versus tone density in defining "tone language"

In traditional typologies, the tone patterns that contrast in Figures 10 and 14 are singled out as constituting "tone proper" or prototypical tone. For example, Hyman (2006:229) says, "A language with tone is one in which an indication of pitch enters into the lexical realization of at least some morphemes," and he explicitly excludes the kind of tonal morphemes that contrast in Figures 9 and 13 on the grounds that they are "post-lexically assigned" to the sentence in its discourse context. An "*intonation-only language*" (or "intonation language") then would be one in which an indication of pitch enters into the realization of only pragmatic morphemes such as the boundary tones in Figure 12.

Some researchers treat prosodic typologies as a continuum, with "pitch accent" languages such as Japanese in the middle. This continuum goes from languages with most tones coming from the lexicon, some tones from the lexicon and some from the discourse pragmatics, and all tones from the pragmatics. Note that this continuum of lexical and post-lexical tone specification is distinct from the continuum of density of tone specification described earlier. However, in most descriptions, the two continua are collapsed, so that "pitch accent language" can refer to either of two different phenomena. It can refer to languages with a sparse distribution of tones such that the number of words in a typical utterance can be identified from the tone pattern alone (as in Tokyo Japanese). However, it can also refer to languages in which some tones are licensed by prominence (i.e., there are starred tones associated with stressed syllables) but the tones shapes for these prominence-marking sequences are lexically contrastive (as in Stockholm Swedish, standard Serbian, or Curaçaoan Papiamentu). This collapsing of the two continua is related to Hyman's (2006) suggestion that while the coarse-grained distinction between tone languages and stress languages is feasible, a finer distinction to include pitch-accent languages is more problematic since many of these putative pitch-accent languages incorporate elements of both traditional tone and stress-

accent languages. We agree with Hyman that "pitch-accent language" is not a truly helpful term, but we would go further, to suggest that even the coarse-grained distinction between "tone language" and "stress-accent language" is not helpful either, since there are any number of ways in which a language can fit Hyman's criterion for a tone language ("a language in which an indication of pitch enters in the representation of at least some morphemes"). Some of these ways may not be conducive to there being a robust alternation between metrically strong and metrically weak syllables, so that it is impossible to pick out an "obligatory head" syllable that imparts "wordhood" to each word. For example, it is difficult to conceptualize any way in which one syllable can be marked as the obligatory head of a word in a language such as Cantonese, where practically every syllable is a morpheme and every syllable preserves its lexical tone even when consonant and vowel autosegments are quite reduced by the process that Wong (2006) calls "syllable fusion". However, Beijing Mandarin and the majority of the languages that Hyman lists as having tone and not stress are metrically quite unlike Cantonese.

The **second challenge** then is that, because there are several relevant dimensions of difference, there are many different ways of allocating languages to the different putative types and hence many different ways in which language change can lead to variable types within a language group. Intonation languages can use tone to contrast pragmatic functions in different ways (e.g., Gussenhoven 2004: 45–46), and the ways to be a tone language are even more varied (Hyman 2006, inter alia). For example, by Hyman's criterion, Limburgian Dutch, Kyungsan Korean, and Panjabi are tone languages, whereas standard Dutch, Chonnam Korean, and Hindi are intonation languages. Given this variation, it is not helpful to simply invoke any of the traditional types in differentiating contact-induced change from "normal" internal variation. In fact, the continuum of variation in tone specification discussed above is observable in contact-induced language varieties as well. This means, one has to look deeper, to see exactly which dimensions are involved in the variation observed within a language group and across neighboring language groups. More specifically, it is important to look at cases where the segmental rhythms of words and phrases leave their tunes open to more than one parse (multiple parses) of their internal segmentation into tonal autosegments or to the tones' association patterns or to their functions.

In this context, it is useful to look at cases of relatively recent change as a result of language contact. Hualde et al. (2002) for example, show that some dialects of Basque (a language with a Tokyo Japanese type tone system) have become intonation languages of the same type as neighboring dialects of Spanish (but different from French). The variation between the Japanese-like system and the Spanish-like system is attributable to language contact.

The TEC examples in Figures 15 and 16 also illustrate effects of recent contact. Figures 15 and 16 are from 2 TEC speakers recorded in the early 1970s. The speech of the Indo-Trinidadian speaker is clearly different from that of the Afro-Trinidadian speaker, with L* accents on stressed syllables in the former as opposed to L+H* accents on stressed syllables in the latter. Interestingly, contemporary recordings of Afro-Trinidadian TEC speakers, have similar L* accents on stressed syllables as in the earlier Indo-Trinidadian recordings. Figure 4, illustrates this in the speech of a female speaker who grew up in a community with Indo-Trinidadians. Data from Drayton (2006) suggests that other Afro-Trinidadians who interact with Indo-Trinidadians also have L* accents instead of L+H* accents on stressed syllables. This suggests a shift in tone specified for a common prosodic function of marking stressed syllables as opposed to a shift in the structure that is marked with an associated change in the type of prosodic system itself. The presence of the L* in the speech of contemporary Afro-Trinidadian speakers is thus plausibly attributable to more recent contact with Indo-Trinidadian TEC speakers who themselves are influenced by Hindi/Bhojpuri.[5] Hindi itself has L* accents on prominent syllables (Harnsburger 1999) as do Indian English dialects that are influenced by Hindi (Pickering and Wiltshire 2000). Downstepping in the contemporary TEC

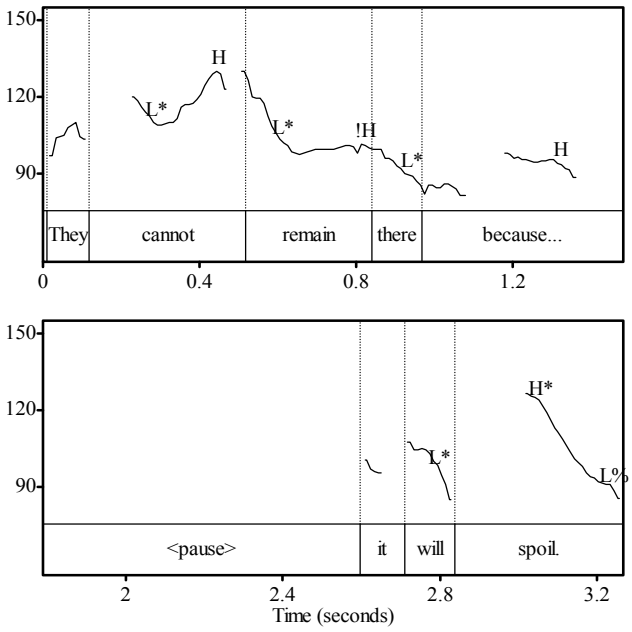

Figure 15. F_0 contour of Indo-Trinidadian TEC speaker with L* accents on stressed syllables. (recorded 1970s)

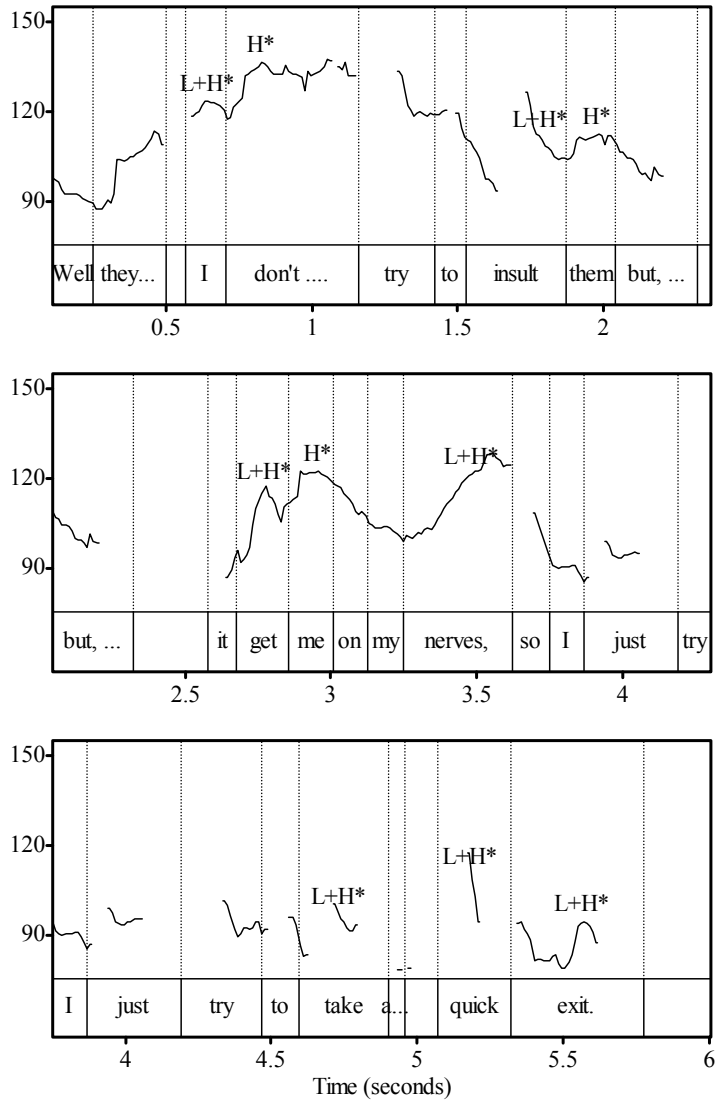

Figure 16. F_0 contour of Afro-Trindadian with L+H* on stressed syllables (recorded 1970s).

utterances in Figures 4 and 8 is also similar to those in the utterance produced by the older Indo-Trinidadian speaker in Figure 15.

7. Prominence-marking mechanisms

In the preceding sections we have focused on the description of tones, their association patterns, and their morphosyntactic functions; we have alluded in some places to such prominence-related notions as *stress* and *pitch-accent*. One of the most pervasive sources of confusion in comparing prosodic systems across languages is that traditional typologies often treat stress as a kind of autosegmental feature rather than as part of the metrical structure. In this section, we will lay out our understanding of stress, as just one of the structural ways in which prominence relationships can be marked prosodically.

It is useful to begin by describing a very different mechanism, which we will call *edge prominence*. Edge prominence refers to the prominence associated with the peripheral positions in a prosodic domain and there are a variety of ways in which these prominences are realized in languages that use this as a salient mechanism.

One manifestation of edge prominence is *pitch range reset*. This is illustrated in Figure 17, where there is a particularly steep pitch rise at the IP boundary between *ya'ne-no* 'roof' and *maNnaka-ni* 'right smack in the middle' — which is the new information in this discourse context. Venditti et al. (2008) identify this kind of reset as one of the most reliable markers of focal prominence in Tokyo Japanese.

The pitch range reset at the beginning of higher-level constituents such as the IP in Tokyo Japanese often coincides with hyperarticulation of consonants and vowels. That is, consonant and vowel autosegments can be scaled toward the edge values of the sonority scale. The beginning of the IP is associated with more stop-like pronunciations for /b, d, g/ and longer release intervals for /p, t, k/, by comparison to AP-medial /b/ and /d/ (which are very lax and are sometimes even approximants) and to AP-medial /g/, which is usually [ŋ].

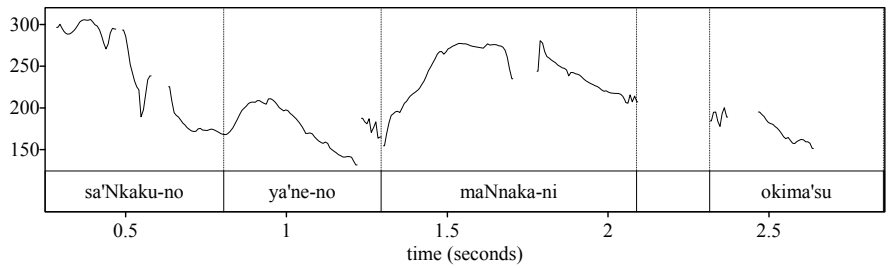

Figure 17. F_0 contour for Tokyo Japanese utterance *(Sono pi'Nku-no ma'do-o) sa'Nkaku-no ya'ne-no maNnaka-ni okima'su.* 'I am placing (that pink window) smack in the middle of the triangular roof.' (Example from the 'house-building' task used to elicit spontaneous narrative in Venditti and Swerts 1996.)

These characteristics can be exaggerated even more at the beginnings of constituents that are in focus in the discourse structure (e.g., new information). That is, intonational phrasing (and dephrasing of subsequent material) along with hyperarticulation of segments at the edge can be used as a focusing mechanism in Japanese (and in many languages).

Note that this exaggerated pitch range and hyperarticulation at the edge of the focal constituent should not be confused with the phenomenon of "accent" in the language. In a Tokyo-type system, "accent" refers to the lexically specified TBU for the H of the HL word melody, which has a culminative distribution in the AP. The words *sa'Nkaku* and *ya'ne* then are "accented" despite being out of focus, whereas *maNnaka* is "unaccented" despite the *focal prominence* imparted by the steep LH edge rise.

In some other languages, such as English, focal prominence is marked instead by associating focus-marking tones to the stressed syllable of the head word of the focal constituent and then suppressing or reducing comparable tones on all analogous TBU in following constituents, as illustrated in Figure 18. This docking of the accent tone on the stressed syllable is predicated on the idea that the designated

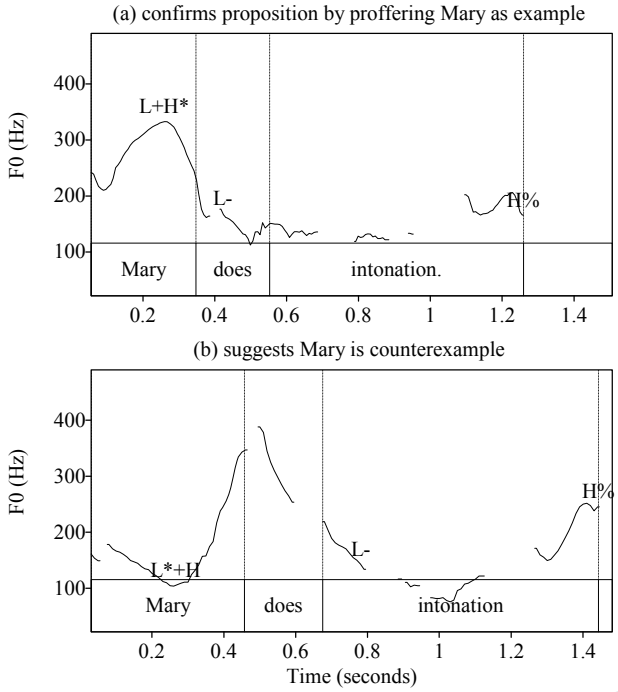

Figure 18. F_0 contour for two utterances of the sentence *Mary does intonation*, produced as hypothetical responses to the proposition that *Only crazy people work on intonation*.

TBU must be prominent at lower levels of the prosodic hierarchy as well (e.g., the head of a foot or a prosodic word). The designated TBU is then the *nuclear stress* and the associated tone shape the *nuclear accent*. In both utterances in Figure 18, then, there is a nuclear pitch accent giving focal prominence to *Mary*, with the L+H* pitch accent (top) offering *Mary* as a confirming example and the L*+H pitch accent (bottom) suggesting a counter-example.

The counterpart to the hyperarticulation of edge segments under edge prominence is the possible hyperarticulation of segments6 in nuclear accented syllables (see, e.g., de Jong 1995). In the English prosodic system, there is also an intimate interaction with vowel quality. In the native (Germanic) vocabulary, there is a lexical alternation between "stressed" syllables with full vowels which can be association points for pitch accents and "unstressed" syllables with reduced vowels or syllabic consonants which normally cannot be association points for pitch accents. When such unstressed syllables must function as an association point under conditions such as metalinguistic correction, where they should have focal prominence, the segmental specification is "promoted" upward in an implication hierarchy of prominence markers. For example, the nucleus is produced as a full vowel, as in *This whisky wasn't* EXported *from Ireland; it was* DEported! where [dəˈpoɹtɛd] becomes [ˈdipoɹtɛd] so that it can bear the nuclear pitch accent (Bolinger 1981). In these circumstances we are able to get pronunciations we would not otherwise get.

8. The phonetic grounding of stress

When markers of prominence are associated to some syllables but not others as described above, the language has a *stress system*. The stress markers must be arranged in an implicational scale such that presence of the prominence marker at one level implies presence of prominence markers at all lower levels. Conversely, lack of prominence at one level precludes the prominence markers at all higher levels. The autosegmental markers of this prominence relationship typically differ for different levels of the prominence hierarchy. The English hierarchy is illustrated in Figure 19 using the representational device of the "metrical grid" to note the stress levels for the syllables in the utterances in Figure 18, with a description of the associated prominence markers (following analyses of Vanderslice and Ladefoged 1972; Bolinger 1981; Beckman and Edwards 1994, inter alia; see Gooden 2003 for similar discussions on JC).

Since phonological markers of prominence differ at different levels of the hierarchy, there are no direct acoustic correlates for stress. Instead, the phonetic properties associated with stress at any level are parasitic on the phonetics of the

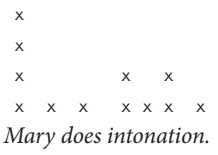

Figure 19. Metrical grid representation showing stressed and accented syllables.

relevant prominence marker (cf. Beckman and Edwards 1994: "Stress is not an autosegment"). At the top levels in Figure 19, for example, F_0 cues stress but only very indirectly, because the pitch accent that marks stress at this level is chosen from a rich inventory of pragmatic morphemes, so the associated tone can be H, as in the first utterance in Figure 18, or it can be L, as in the second. The native listener parses the F_0 contour and its alignment with the words in terms of some legal string of tones, and the anchoring sites for tone sequences, parsed as pitch accents, are perceived as accented (cf. Fry 1958). Other acoustic parameters such as duration and vowel quality come into play because they are correlates of the stress markers at lower levels of the hierarchy — e.g., the listener does not misinterpret the H in the second utterance in Figure 18 as a pitch accent, because the reduced vowel in *does* cannot be an association point. The acoustic correlates of stress are themselves not automatic but will differ across languages. In Swedish, for example, the lowest level of the hierarchy is defined by the prosodic contrast between long syllables (with a VV or VC rhyme) and short syllables. Association of a lexically-specified H+L* or H*+L "word accent" defines the next higher level. In Mandarin Chinese, on the other hand, the lowest level of the hierarchy is defined categorically by the lack of tone specification ("neutral tone") and only secondarily by hypoarticulation (i.e., segmental lenition including vowel reduction). The next level of the hierarchy then is marked by various tone-scaling phenomena such as downstep. Stress-related information is thus derived from a complex combination of acoustic cues associated with the duration, amplitude and F_0 of syllabic sequences within an utterance. All of this points to the fact that stress itself is not a suprasegmental feature.

Note, too, that even related languages with structurally identical stress systems can differ in the tone shapes that mark accented syllables, as well as in typical patterns of tone scaling and tonal alignment, among other things. Research on English varieties provides evidence that there are phonetic differences for the same pragmatic tone sequences (cf. Grabe et al., 2000; Arvaniti and Garding, 2007). A similar situation holds in the Caribbean creole languages that have stress systems. Stressed syllables in both JC and TEC are longer than unstressed syllables and are realised with a low F_0, so at first pass the HL sequences in Figure 20 appear to be identical tone shapes. However, when they are parsed with reference to the

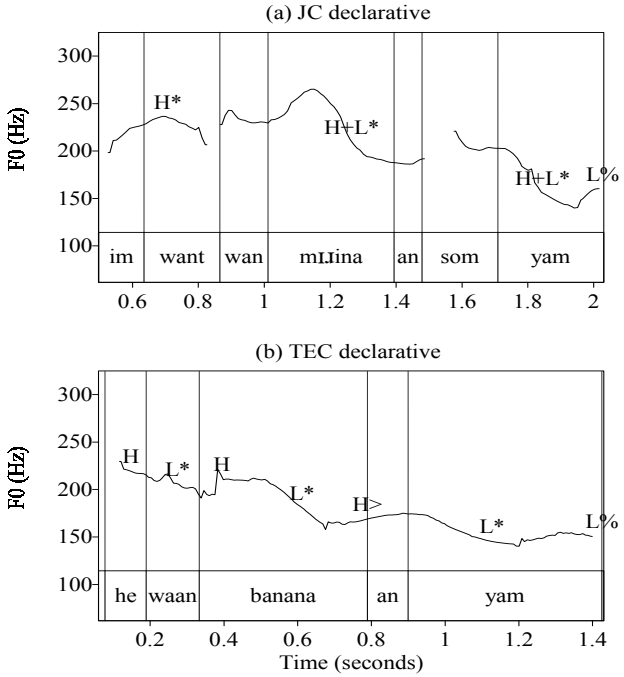

Figure 20. F_0 contours of JC and TEC declaratives, showing different association points for superficially similar HL sequences (a) *He wants an undershirt and some yams* (b) *He wants banana(s) and yam(s).*

prosodic structure, we see the differences in the association of the H of the HL sequence as a leading tone of a bitonal pitch accent in JC (Gooden 2003) and as a phrase marker for the AP in TEC (Drayton 2004), just as in Hindi (Harnsberger 1999), Bengali (Hayes and Lahiri 1991), and Indian English (Pickering and Wiltshire 2000).

The comparison between JC and TEC declarative intonations in Figure 20 highlights another important aspect of the AM theory. In terms of the stress hierarchy, the two varieties are similar. Both languages have an implicational scale of prominence markers, with pitch accents associating only to syllables that are produced with "strong" vowel and consonant allophones at a lower level of the hierarchy. However, their phrasal hierarchies differ, with TEC (but not JC) having a tonally marked prosodic constituent below the IP. Evidence for positing the AP in TEC is seen when we compare the accents on the word *aloo* in intonational-phrase medial versus phrase initial position (Figure 21). In phrase initial position the H tone is absent, which is not predicted if the H were a part of a bitonal sequence.

In JC and TEC, as well as in all varieties of English, the pitch accent shapes constitute an "intonational lexicon" of pragmatic morphemes rather than a stock

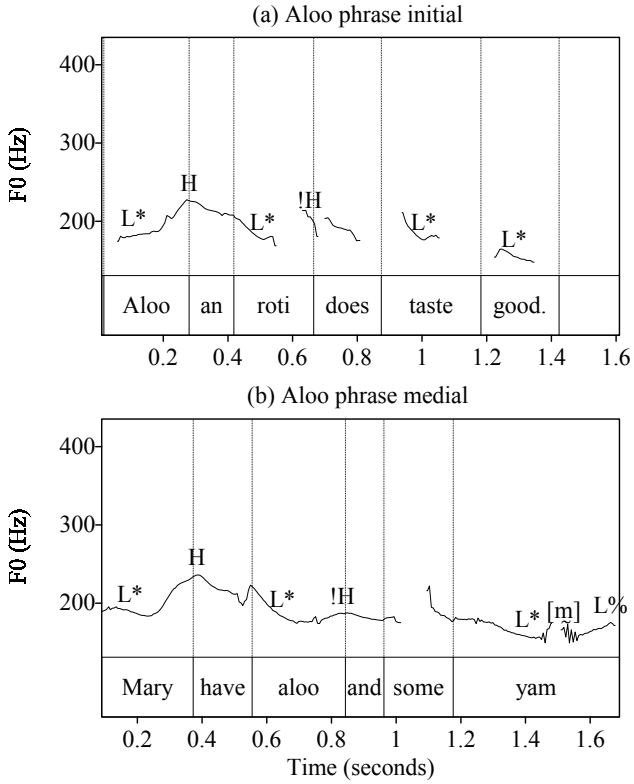

Figure 21. F_0 contours for (a) Aloo and roti taste good and (b) Mary has aloo and some yams.

of tonal autosegments on par with vowels and consonants. In other languages, by contrast, the tones that associate to stressed syllables at some high level of the stress hierarchy are not drawn from an inventory of paradigmatically contrasting pragmatic morphemes, but instead are lexically specified autosegments. Even these languages that have both tone and stress may combine these in very different ways making them different prosodically. In Beijing Mandarin, for example, the tones that associate to stressed syllables at a level analogous to level 3 in Figure 19 are the inventory of four lexical tone patterns that contrast the morphemes in Figure 14. In Stockholm Swedish, similarly, there are two tone shapes that mark stressed syllables at a level homologous to stress level 2 in the English hierarchy in Figure 19, but these pitch accents are lexically specified and contrast words rather than constituting pragmatic morphemes in their own right. For example, in Stockholm Swedish both *anden* 'the duck' and *anden* 'the ghost' have primary stress on the first syllable, but the pitch accent that marks this stress on 'the duck' is H+L* (i.e., falling onto a L tone target on the stressed syllable) whereas the pitch accent

on 'the ghost' is H*+L (falling from a H tone target on the stressed syllable). (Importantly in citation utterances, there will be a peak on the second syllable as well, but this is a "phrase accent" that marks the last word in the IP or in a constituent bearing focal prominence — see Bruce 1977, 1990.)

In Bruce's (1977) seminal monograph on Swedish stress and pitch accent, he showed how the intonation patterns of sentences can be analyzed as a string of tone targets contributed by two different components of the grammar. Some tones in the string are "pre-linked" in the lexicon (i.e. underlying) and contrast two classes of words, the so-called "Accent 1" words versus the "Accent 2" words. Other tones in the string are word tones and sentence-level intonation word accents in sentence perspective.

Remijsen and van Heuven (2005) make a similar argument for Curaçao Papiamentu, with the added difference that this variety of Papiamentu has a lexical contrast between words with no lexically specified tone (like the "unaccented" words of Basque and Japanese) and words with a lexically specified falling tone. That is, they set up a contrast between words that are realized in citation form with a rising F_0 (an intonational prominence-marking morpheme) on the stressed syllable and words that are realized in citation form with a lexically specified fall on the stressed syllable and the rising F_0 that marks the intonational prominence displaced to dock on the final syllable.

To return to the point that we tried to make near the beginning of Section 6, we can rephrase it as follows. There are any number of ways in which in which a language can fit Hyman's criterion for a tone language. Some of these ways are completely compatible with there being a robust implicational scale for prominence marking of the sort that defines one syllable as the "obligatory head" of a word and that word as the "obligatory head" of a phrase. In Swedish, for example, every word must have a long syllable (i.e., a syllable with a long vowel or a coda consonant), which can associate to a prominence-marking tone. At one level of the metrical hierarchy, that tone is the lexical H+L* or H*+L that picks out the syllable with the obligatory culminative prominence that is termed "accent" (hence the term "word accent"). At a higher level of the metrical hierarchy, one word in each intonational phrase must be picked out as the obligatory head word that also bears the "phrase accent" (H-). In this language tone association defines two distinct levels of metrical prominenence. In Papiamentu, by contrast, the head syllable at the word level is obligatorily marked with a durational prominence. It is marked also at that level with a tonal prominence only if it is a "lexically accented word" (i.e., "Type II" in Remijsen and van Heuven's terms). In both cases, there is a clear implicational scale of prominence markers, but tone figures in that scale in different ways at different levels of the hierarchy in the two languages. If we focus on the tone association facts, then the resemblance is lost, and Papiamentu looks

more like Japanese, despite the clear differences in the metrical systems of the two languages and in the functions that the tone patterns play relative to the metrical structures of words.

9. Cross-linguistic differences in tonal association and prominence marking

In this section we briefly recapitulate some of the things that we have learned in reviewing the research on these differences in tonal association across languages, and discuss them in terms of the (we think not completely helpful) distinction between "stress languages" and "tone languages" in traditional prosodic typologies.

A *first* relevant observation is that there is an intimate link between tonal markers of syllable prominence at the top of the stress hierarchy and the prosodic expression of focus. In English, focal prominence is marked by the association of an IP edge tone (a "phrase accent") after the last pitch accent in the focused constituent, which blocks any following pitch accents. The last pitch accent is then the "nuclear accent" and that accented syllable is the head of the IP. In Swedish, focal prominence is marked by the association of a H phrase accent and downstep at each subsequent word accent (Bruce 1982). In Beijing Mandarin, focal prominence is marked by a clear realization of the lexical tones in the focused constituent followed by a drastic compression of the pitch range (Jin 1996). In TEC we see a reduced pitch range over the initial AP in the narrow focus declarative, whereas under broad focus, the initial AP is produced in a higher pitch range (see Figure 22); in both cases the L* simply marks the stressed syllable under focus. There are also segmental effects of a lengthened vowel in the focussed stressed syllable. The H tone that marks the edge of the post focal AP in the narrow focus utterance is also prominent, being higher than the analogous AP in the broad focus utterance. In addition, the stressed syllable in the post focal utterance is deaccented. When the focused constituent occurs in utterance final position as in Figure 23, the H tone of the AP is not seen as the L% takes precedence. This suggests that it is the vowel length and presence of the L* that are both important for narrow focus.

A *second* observation concerns the syntagmatic constraints at a lower level of the stress hierarchy. In all languages that have been identified to date as having a prominence hierarchy that is structurally equivalent to the one described for English in Figure 19, there is some level of stress that is part of the definition of the prosodic word (see Hyman's 2006, *Obligatory Head* constraint). In English, for example, a pitch accent can only associate to a syllable with a full, unreduced vowel and every prosodic word must have (at least) one such metrically strong syllable. When a function word such as *the*, that normally is pronounced with a reduced

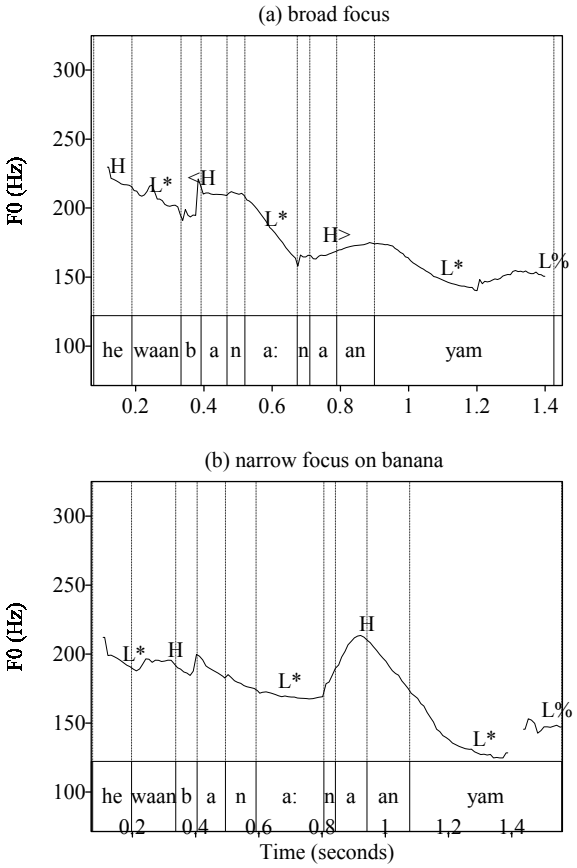

Figure 22. F_0 trace of broad focus utterance (a) *He wants bananas and yams* and narrow focus utterance (b) *He wants bananas (not apples) and yams.*

vowel (or even with no clear vowel nucleus), is put into focus, it is given a nonce pronunciation with the full vowel that occurs in *heat* or in *hut*. In Swedish, every prosodic word must have a syllable that bears a pitch accent. In Mandarin Chinese, every prosodic word must have a syllable with a fully realized lexical tone. Syllables without lexical tone are either suffixes, such as the *–men* suffix in the word *tā-men* 'they' in Figure 13, or function words, such as the negative marker *bu* or the particle *ma* in Figure 13. In fast speech, the vowels in such toneless syllables can be reduced or even deleted.

Depending on the typologist, the term *stress language* can mean any one of following: (a) It can refer to any language that has a hierarchy of syllable prominences structurally like the one in Figure 19; (b) It can refer to the subset of languages in (a) that define the prosodic word in terms of a culminative Obligatory Head syllable that is marked for prominence by tone association at some higher level of the hier-

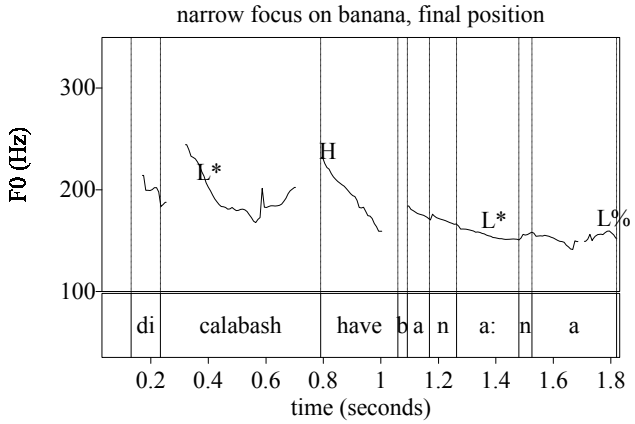

Figure 23. F_0 contour of *The calabash has bananas (in it)*, produced with narrow focus on banana

archy; (c) It can refer just to the subset of languages in (b) in which the tones that associate to stressed syllables come from an inventory of pragmatic morphemes.

In many languages, the obligatory head syllable is constrained to occur at a fixed position or within a fixed range of positions relative to the prosodic word edge. For example, in Greek, the stressed syllable (the syllable to which pitch accents can associate) must occur on one of the last three syllables of the word. In English, the majority of words have stressed first syllables at the level where stress is marked by an associated full vowel (grid level 3). In Shanghainese, similarly, words are stressed on the first syllable. When monosyllabic morphemes are combined into compound words, lexical tones on all syllables except the first are deleted and the vowels and consonants can undergo comparable reduction processes. For example, high vowels in non-initial syllables can be devoiced or deleted (Zee 1990).

The **third challenge** then is that since languages with stress systems often use a constrained distribution of stress to demarcate lexical words, the misparsing of edge prominence and pitch range reset as markers of stress in languages without stress accent is a very common contact-induced reanalysis. Moreover, because many field linguists are first-language speakers of languages with syllable-prominence systems, the same misparsing can influence their phonological analyses. As Hyman (2006: 246) points out, "Researchers who operate under the assumption that all languages have S[tress]A[ccent] may read stress into the phonetic variations they hear or observe instrumentally." Thus, linguists who are speakers of English and other Germanic dialects often analyze the LH sequence that aligns to the beginning edge of many accentual phrases in Hexagonal French7 as a pitch accent placed variably on the first or second syllable of the phrase, although the alignment facts do not support this analysis (Welby 2006).

The **fourth challenge** is that the grammar of focus marking can involve manipulations of phrasing, pitch range, tone scaling, and/or tone association patterns — i.e., an intricate interplay of prominence markers at several levels of the prosodic hierarchy. When words are studied in citation utterances, they are "in focus" and it is impossible to sort out which aspects of the utterance melody are due to lexically specified structure and which to the grammar of focus marking. For example, traditional analyses of Swedish word prosody interpreted the phrase accent as a part of the lexical specification of words bearing the H*+L pitch accent. It is therefore imperative that the analyst examine words in longer utterances with varying pragmatic functions in the discourse, alongside those in citation form. In a similar fashion, even where subsystems or strata of languages exhibit different prosodic properties, we also have to pay attention to the pattern that is reflective of the bulk of the lexicon. Until higher-level metrical structure and the grammar of focus-marking are better studied across more diverse languages, any gross typology of "word prosody" may be premature.

10. Summary

As we have discussed here, there is considerable variability in prosodic systems cross-linguistically such that attributing prosodic variation to language contact can be challenging. The main points we want to emphasize here are the following. First, parsing ambiguities of phonological structures can occur in situations of language contact and can lead to structural changes. The phonetic cues to the phonological categories can be similarly misparsed. Second, intonation languages can use tone to contrast pragmatic functions in different ways and even traditional tone languages are not a homogeneous group. This reinforces the idea that prosodic variation does not necessarily result from language contact but can be the result of system internal variation. Third, phonetic cues to prosodic domains in non-stress languages may be misinterpreted as markers of stress by speakers of stress languages, and vice versa. This means that the diachronic source of a stress contrast in a creole language need not have been in the lexicon of any of the contributing lexifier languages. Stress can arise instead from the reinterpretation of word-level tonal melodies. The source is difficult to determine from historical/diachronic data since prosodic information is not typically recorded. In investigating the possibilities of contact-induced change on the prosodic system(s) of Creole languages, the analyst must shift from the purely diachronic- and corpora-based approaches to theoretically grounded methods that incorporate experimental techniques tested on synchronic data. At the same time using synchronic data to argue for contact-induced/diachronic change from creole formation should be done very cautiously.

Fourth, discourse functions such as focus marking often involve an intricate interplay of prominence markers at several levels of the prosodic hierarchy. This makes it necessary to examine prosodic properties of target words in several different prosodic contexts and in different discourse contexts in making classifications. A great deal of work remains to be done in classifying and describing the contemporary prosodic systems of these languages and the description of discourse-level intonational patterns lags very far behind the descriptions of grammatical patterns. Consequently, very little is known about the development of the prosodic systems of Caribbean creoles (but see Devonish 1989, 2002). The only way to learn more is through systematic studies of multiple varieties which will eventually lead to more solid evidence on the development of prosodic systems in specific creole varieties. Furthermore, we suggest that phonological change and more specifically prosody does not operate in the same way as does grammatical (morphological and syntactic) change. This means we must also refine our methodologies for examining putative contact-induced change. The analyst must therefore explore all the known explanations for language change; that is, internal development, universal principles, contact-induced change (substrate, superstrate or adstrate influence) in describing the prosodic system of contact languages (as for other languages for which the language contact hypothesis is proposed).

Notes

1. We can appreciate this three-way ambiguity when we examine notorious "exceptions" to universals such as sonority sequencing principles. For example, when we follow Dell and Elmedlaoui (1998) in saying that a voiceless obstruent segment can occupy the head position in a syllable in Tamazight Berber, we are tacitly acceding to a particular account of the syllable-internal Metrical structure and a particular account of how the Autosegmental content features that contrast sounds such as [t] to sounds such as [l] are aligned to syllable-internal positions. When we instead follow Coleman's (1996) alternative description of these "exceptional" syllables, we are tacitly assuming a different account of the Metrical organization, one that parses the stop release as a separate segmental unit at the bottom of the prosodic hierarchy. The phonetic description of these syllabic stop consonants as obligatorily released is unchanged.

2. An astute reviewer points out that many phonologists use both terms, to refer to two different types of transition from one tone target to the next. In such differentiated usage, *interpolation* refers to transitions over tonally unspecified metrical positions between the association points for two adjacent tone targets, whereas *spreading* refers to a process of copying the tone specification for the first target onto one or more potential association points up to the association point for the second of two adjacent tone targets. The criteria for differentiating between these two types of transition are highly theory-specific and disentangling the assumptions behind different theories to be able to be more specific about the transition from the H to the L* in the TEC accentual phrase is beyond the scope of this paper.

3. The convention is to notate two tone targets per syllable, using Chao's (1930) numerical scale, where 1 corresponds to L and 5 corresponds to H.

4. This is reflected in the standard Chao-type notation of these tones as being 35 for tone 2 as opposed to 214 for tone 3.

5. Mohan (1990) discussed a koineization of the Indic dialects into a variety called Trinidad Bhojpuri which, like other Bhojpuri varieties, shares a lexicon with and is related to Hindi. Also many Indo-Trinidadians have continuous religious and cultural exposure to the Hindi language.

6. Of course, hyperarticulation is not always possible under these circumstances.

7. French spoken in metropolitan France, *l'Hexagone*.

References

Arvaniti, Amalia & Gina Garding. 2007. Dialectal variation in the rising accents of American English. *Papers in Laboratory Phonology* 9, Jennifer Cole & José I. Hualde (eds.), 547–576. Berlin: Mouton de Gruyter.

Beckman, Mary E. 1996. The parsing of prosody. *Language and Cognitive Processes* 11: 17–67.

Beckman, Mary E. & Jan Edwards. 1994. Articulatory evidence for differentiating stress categories. *Laboratory Phonology* 3: 7–33.

Browman, Catherine P. & Louis Goldstein. 1988. Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45: 140–155.

Bruce, Gösta. 1977. *Swedish word accents in sentence perspective*. Lund: Gleerup.

Bruce, Gösta. 1982. Developing the Swedish intonation model. Department of Linguistics Working Papers. Lund University.

Bruce, Gösta.1983. Accentuation and timing in Swedish. *Folia Linguistica* 27(1–2): 221–23.

Bruce, Gösta. 1990. Alignment and composition of tonal accents: Comments on Silverman and Pierrehumbert's paper. *Papers in Laboratory Phonology* 1: 107–114.

Bolinger, Dwight.1981. Two kinds of vowel, two kinds of rhythm. *Intonation and its parts: The melody of language*, Dwight Bolinger (ed.), 1986. Stanford CA: Stanford University Press.

Chao, Yuan-Ren. 1930. A system of tone letters. *Le Maître Phonetique* 30: 24–27.

Chomksy, Noam & Morris Halle. 1968. *The sound pattern of English*. New York NY: Harper and Row.

Coleman, John. 1996. Declarative syllabification in Tashlhit Berber. *Current trends in phonology: Models and methods,* Jacques Durand & Bernard Laks (eds.), 177–218. Salford: European Studies Research Institute, University of Salford.

de Jong, Kenneth. 1995. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 91: 491–504.

Dell, François & Mohamed Elmedlaoui. 1996. Nonsyllabic transitional vocoid in Imdlawn Tashlhiyt. *Current trends in phonology: Models and methods,* Jacques Durand & Bernard Laks (eds.), 217–244. Salford: European Studies Research Institute, University of Salford.

Devonish, Hubert. 1989. *Talking in tones: A study of tone in Afro-European Creole languages*. London: Karia Press.

Devonish, Hubert. 2002. *Talking rhythm, stressing tone: Prominence in Anglo-West African Creole languages.* Kingston: Arawak Press.
Docherty, Gerard J. 1992. *The timing of voicing in British English obstruents*. Dordrecht: Foris.
Dodsworth, Robin. 2005. Linguistic variation and sociological consciousness. PhD dissertation, Ohio State University.
Drayton, Kathy-Ann. 2004. Course project paper for Ling 795T (Practicum on Intonation), Ohio State University.
Drayton, Kathy-Ann. 2006. Word Level prosody Trinidadian English Creole: A phonetic analysis. Seminar paper, University of the West Indies: St. Augustine.
Dyer, Judy & Alicia Beckford Wassink. 2001. Taakin braad and talking broad: Changing indexicality of phonetic variants in two contact situations. *Texas Linguistic Forum* 44(2): 288–301.
Eda, Sanae. 2004. Processing of intonation patterns in Japanese: Implications for Japanese as a foreign language. PhD dissertation, Ohio State University.
Frota, Sonia. 2000. *Prosody and Focus in European Portuguese: Phonological Phrasing and Intonation*. New York NY: Garland.
Fry, Dennis. B. 1958. Experiments in the perception of stress. *Language and Speech* 1: 205–123.
Goldsmith, John A. 1990. *Autosegmental and metrical phonology.* Oxford: Basil Blackwell.
Good, Jeff. 2006. A twice-mixed creole? Tracing the history of a prosodic split in the Saramaccan lexicon. Paper presented at the LSA/SPCL Meeting Special Panel Session on Prosodic Descriptions of Creole Languages: Implications for Creole Formation, January 6–7. Albuquerque, New Mexico.
Gooden, Shelome A. 2003. The phonology and phonetics of Jamaican Creole reduplication. PhD dissertation, Ohio State University.
Gooden, Shelome A. 2005. Proposal for special session on *Prosodic descriptions of Creole languages: Implications for Creole formation*. Annual Meeting of the Society for Pidgin and Creole Linguistics, 6–7 January 2006, Albuquerque, NM.
Grabe, Esther; Brechtje Post; Francis Nolan & Kimberley Farrar. 2000. Pitch accent realization in four varieties of British English. *Journal of Phonetics* 28: 161–185.
Grice, Martine. 1995. *The intonation of interrogation in Palermo Italian: Implications for intonation theory*. Tuebingen: Max Niemeyer.
Gussenhoven, Carlos. 2004. *The phonology of tone and intonation.* Cambridge: CUP.
Gussenhoven, Carlos & Peter van der Vliet. 1999. The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics* 35: 99–135.
Haraguchi, Shôsuke. 1977. *The tone pattern of Japanese: An autosegmental theory of tonology*. Tokyo: Kaitakusha
Harnsberger, James D. 1999. The role of metrical structure in Hindi intonation. Paper presented at the 19th South Asian Linguistics Analysis Roundtable, University of Illinois, Urbana-Champaign.
Hayes, Bruce & Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9: 47–96.
Hualde, José; Gorka Elordieta; Iñaki Gaminde & Raika Smiljani. 2002. From pitch-accent to stress-accent in Basque. *Laboratory Phonology* 7: 547–584.
Hyman, Larry M. 2006. Word-prosodic typology. *Phonology* 22: 225–257.
Igarashi, Yosuke. 2007. Typology of prosodic phrasing in Japanese dialects. Paper presented at the Workshop on Intonational Phonology of Understudied or Fieldwork Languages. ICPHS 2007. Saarbrücken, Germany.

Irvine, Alison. 2004. A good command of the English language: Phonological variation in the Jamaican acrolect. *Journal of Pidgin and Creole Languages* 19: 41–76.
Jin, Shunde. 1996. An acoustic study of sentence stress in Mandarin Chinese. PhD dissertation, Ohio State University.
Jun, Sun-Ah. 1996. *The phonetics and phonology of Korean prosody: Intonational phonology and prosodic structure.* New York NY: Garland.
Jun, Sun-Ah (ed.). 2005. *Prosodic typology: The phonology of intonation and phrasing.* Oxford: OUP.
Kratochvil, Paul. 1998. Intonation in Bejing Mandarin. *Intonation systems: A survey of twenty languages*, Daniel Hirst & Albert Di Cristo (eds.), 417–431. Cambridge: CUP.
Ladd, D. Robert.1996. *Intonational phonology*. Cambridge: CUP.
Lawton, David. 1963. Suprasegmental phenomena in Jamaican Creole. PhD dissertation, University of Michigan.
Lee, Ok Joo. 2000. The pragmatics and intonation of ma-particle questions in Mandarin. MA thesis, Ohio State University.
Mohan, Peggy. 1990. The rise and fall of Trinidad Bhojpuri. *International Journal of the Sociology of Language* 85: 21–30.
Peng, Shu-hui; Marjorie K.M. Chan; Chiu-yu Tseng; Tsan Huang; Ok Joo Lee & Mary E. Beckman. 2005. Towards a pan-Mandarin system for prosodic transcription. *Prosodic typology: The phonology of intonation and phrasing*, Sun-Ah Jun (ed.), 230–270. Oxford: OUP.
Pickering, Lucy & Caroline Wiltshire. 2000. Pitch-accent in Indian-English teaching discourse. *World Englishes* 19:173–183.
Pierrehumbert, Janet B. 1980. The phonology and phonetics of English intonation. PhD dissertation, Massachusetts Institute of Technology.
Price, P.J. 1980. Sonority and syllabicity: Acoustic correlates of perception. *Phonetica* 37: 327–343.
Remijsen, Bert & Vincent J. van Heuven. 2005. Stress, tone and discourse prominence in the Curaçao dialect of Papiamentu. *Phonology* 22: 205–235.
Rivera-Castillo, Yolanda. 2006. Subsystem interface and tone typology in Papiamentu. Paper presented at the LSA/SPCL meeting special panel session on *Prosodic Descriptions of Creole Languages: Implications for Creole Formation*, January 6–7. Albuquerque NM.
Selkirk, Elisabeth. 1995. Sentence prosody: Intonation, stress and phrasing. *The handbook of phonological theory,* John Goldsmith (ed.), 550–569. Cambridge: Blackwell.
Sproat, Richard & Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21: 291–311.
Sutcliffe, David. 1986. Jamaican Creole tonality and the implications for syntactic analysis. *Society for Caribbean Linguistics Occasional Paper* 2.
Vanderslice, Ralph & Peter Ladefoged. 1972. Binary suprasegmental features and transformational word-accentuation rules. *Language* 48: 819–838.
Venditti, Jennifer & Marc Swerts. 1996. Prosodic cues to discourse structure in Japanese. *Proceedings of the International Conference on Spoken Language Processing (Philadelphia, October 1996)* 2: 725–728.
Venditti, Jennifer J.; Kazuaki Maeda & Jan P. H. van Santen. 1998. Modeling Japanese boundary pitch movements for speech synthesis. *Proceedings of the Third ESCA workshop on speech synthesis*, 317–322. Jenolan Caves, Australia.
Venditti, Jennifer J., Kikuo Maekawa & Mary E. Beckman. 2008. Prominence marking in the Japanese intonation system. *Handbook of Japanese linguistics*, S. Miyagawa & M. Saito (eds.) Oxford: OUP.

Welby, Pauline. 2006. French intonational structure: Evidence from tonal alignment. *Journal of Phonetics* 34: 343–371.

Wong, Wai-Yi Peggy. 2006. Syllable fusion in Hong Kong Cantonese connected speech. PhD dissertation, Ohio State University.

Zee, Eric. 1990. Vowel devoicing in Shanghai. *City University of Hong Kong Papers in Linguistics* 2: 69–104.

Authors' addresses

Shelome Gooden
Department of Linguistics
University of Pittsburgh
2830 Cathedral of Learning
Pittsburgh, PA 15260, USA

sgooden@pitt.edu
kadray@yahoo.com
mbeckman@ling.osu.edu