

7 Tone: Is it Different?

LARRY M. HYMAN

1 Introduction

Except for a brief period in the late 1970s and early 1980s, tone has generally fallen outside the central concerns of theoretical phonology. During that period, the concepts and formalisms of Goldsmith's (1976a, b) autosegmental approach to tone provided the model to address other aspects of non-linear phonology including vowel harmony (Clements 1976, 1981), nasal harmony (Hyman 1982), and feature geometry (Clements 1985; Sagey 1986). In addition, autosegmental approaches to templatic morphology (McCarthy 1981), reduplication (Marantz 1982), and other aspects of prosodic morphology owed their inspiration to tone, which through the work of Pulleyblank (1986) provided important insights into the developing framework of lexical phonology and morphology (Kiparsky 1982b, 1985; Mohanan 1986). Most generative work prior to and during this period had centered around African tone systems (Leben 1973a; Hyman and Schuh 1974; Goldsmith 1976a; Clements and Ford 1979; Clements and Goldsmith 1984), two notable exceptions being Haraguchi (1979) and Yip (1980), who dealt with the tonal dialectology of Japanese and Chinese, respectively. Finally, Pierrehumbert (1980) developed an influential autosegmental approach to intonation based on English, which was subsequently applied to Japanese (Beckman and Pierrehumbert 1986) and many other languages since. While the autosegmental legacy is still quite alive, tone has not contributed as centrally to subsequent theoretical innovations in phonology. In the case of optimality theory (Prince and Smolensky 1993; McCarthy 2002), there have been some interesting applications, for example Myers' (1997) treatment of the Obligatory Contour Principle (OCP), but theoretical developments have largely been based either on segmental phonology or on stress, syllabification, reduplication, and other aspects of prosodic phonology and morphology.

My goal in writing this chapter is twofold. First, I propose to cover some of the aforementioned contributions that tone has made to phonological theory. Second, I wish to show that there is still much more for phonologists and others to learn from tone. I suggest that linguists should be very concerned about tone, for at least three reasons:

(i) Tone systems are found in approximately 50% of the languages of the world. The greatest concentrations of tone languages are found in Sub-Saharan Africa, East and Southeast Asia, South-central Mexico, and parts of Amazonia and New Guinea. While we have had access to information about the first three areas for some time, comparatively little has been available on tone in Amazonia and New Guinea until recently. The emerging picture is that these tone systems have interesting and diverse properties which complement the already varied African, Asian, and Mexican systems. The result is an extraordinary richness and a potential gold mine for future investigations.

(ii) The study of tone has influenced the history of phonology and promises to contribute further to our understanding of language in general, particularly interface issues. For instance, some of the most detailed and influential studies concerning the syntax-phonology interface have drawn from tonal alternations applying at the phrase level (Clements 1978; Chen 1987). More recently, several meetings have brought together scholars interested in the relation between tone, phrasal accent, and intonation, especially in the languages of Europe and East Asia (Germanic, Slavic, Basque, Chinese, Japanese, Korean, etc.), areas covered in some detail by Ladd (1996) and Gussenhoven (2004). Three collections of note are van der Hulst (1999), Jun (2005), and Riad and Gussenhoven (2007). Cutting across the components of grammar, both abstract and instrumental research have been concerned with how focus and other aspects of information structure, often marked by stress or intonation, are realized when there is a competing tone system. Some of this work has shown that focus is not necessarily prosodically marked in certain languages with tone (Downing 2007).

(iii) Tone systems have properties which SURPASS segmental and metrical systems. In Section 6 I conclude that tone can do everything that segmental and metrical phonology can do, but that the reverse is not true. This is especially true of the long-distance effects that tone exhibits both within and across words, as when the tone of one word migrates several syllables or words to its right. Since some tonal phenomena have no segmental or stress analogues, anyone who is interested in the outer limits of what is possible in phonology would thus be well-served to understand how tone systems work.

Despite the widespread occurrence of tone in the world's languages and the important contributions it has already made to our understanding of phonology and its interfaces, the lack of familiarity of some scholars with tone has allowed certain old misconceptions to persist. The rest of the chapter is organized as follows. In Section 2 I begin by defining tone and characterizing these persistent misconceptions. In Section 3 I discuss the autosegmental insight into tone. Section 4 deals with the question of whether tone has different properties from other phonology, either quantitatively (4.1) or qualitatively (4.2). Section 5

considers the issue of whether tone should sometimes be analyzed in accental terms. Throughout the discussion the focus will be on the question of what tone has to teach phonologists, and therefore, crucially, whether there are phonological properties that are found only in tone systems. The conclusion in Section 6 summarizes the key points of the preceding sections, concluding that the capabilities of tone do surpass those of either segmental or metrical phonology.

2 Defining Tone: Three Misconceptions

A logical place to begin is by raising an old, but most essential question, “What is tone?” How do we know if a language has tone? Two early definitions state that a language with tone is a language

- (1) a. “... having significant, contrastive, but relative pitch on each syllable” (Pike 1948: 3)
 b. ... in which both pitch phonemes and segmental phonemes enter into the composition of at least some morphemes. (Welmers 1959: 2)

While Pike originally saw tone as a contrastive feature on each syllable or other tone-bearing unit (TBU), Welmers’ definition insists on the MORPHOLOGICAL nature of tone: tone is not a property of syllables, as expressed by Pike, but rather of morphemes. Welmers correctly pointed out that not all morphemes need to have a tone – some may be toneless. Similarly, not all morphemes need to have a TBU – they may be “tonal morphemes” (Section 2, Section 3). We will see that there is great advantage in approaching tone from this morphological perspective. However, it is useful to update and slightly modify Welmers’ definition as in (2), and say that a language with tone is a language

- (2) ... in which an indication of pitch enters into the lexical realization of at least some morphemes. (Hyman 2001c: 1368; Hyman 2006: 229)

This statement defines tone in terms of individual morphemes whose tones may combine (and interact) in forming words. Since word-level tones may be assigned by rule, “lexical realization” refers to the output of lexical phonology, not necessarily underlying representations. For example, it could be argued that the Somali noun roots in (3) are underlyingly toneless:

- | | | | | | |
|-----|-----------|-----------|-----------------|----------|------------------|
| (3) | root | masculine | | feminine | |
| a. | /inan/ | ínan | ‘boy’ | inán | ‘girl’ |
| | /naʃas/ | náʃas | ‘stupid man’ | naʃás | ‘stupid woman’ |
| | /goray/ | góray | ‘male ostrich’ | goráy | ‘female ostrich’ |
| b. | /darmaan/ | darmáan | ‘colt’ | darmaán | ‘filly’ |
| | /ʒeesaan/ | ʒeesáan | ‘young he-goat’ | ʒeesaán | ‘young she-goat’ |
| | /dameer/ | daméer | ‘he-donkey’ | dameér | ‘she-donkey’ |

For these nouns, the masculine morpheme assigns a high (H) tone (´) to the penultimate vowel, while the feminine morpheme assigns an H tone to the final vowel (Hyman 1981; Saeed 1999). Toneless vowels which follow an H are realized low (L), while vowels which precede an H tone are realized mid (M). The contrast in (3a) is thus between [H-L] and [M-H], while in the [M-HL] and [M-MH] contrast in (3b), the last syllable contains a long vowel with either falling or rising pitch. Somali also illustrates the importance of the wording “indication of pitch” in the definition in (2), which is taken to mean tone features or any other analytical device whose only function is to characterize pitch. In the above analysis, an H tone is assigned to one of the last two vowels of the noun. Although an analyst may prefer to assign a diacritic accent (*) to the relevant vowel instead, the diacritic device would still be marking only H tone, which in fact may be entirely absent on a word, for example on both the subject noun and verb in the utterance *inan wáa dhaŕay* ‘a boy fell’. The definition in (2) would thus include Somali, but would exclude languages where all indications of pitch are introduced post-lexically, that is, at the phrase or utterance level.

While it is possible to view the above Somali roots as underlyingly toneless, and the gender markers as tonal morphemes devoid of a TBU, the most common situation is for morphemes to consist of both segmental and tonal features. This is seen in the minimal pairs, triplets, quadruplets, and quintuplets in (4)–(7) showing examples contrasting two, three, four, and five contrastive tone heights, respectively:

(4) Two levels: Dadibi (Papua New Guinea) (MacDonald and MacDonald 1974: 151)

- a. L (low) tone : wà ‘string bag’ nà ‘shoulder’
 b. H (high) tone : wá ‘edible greens’ ná ‘aunt’

(5) Three levels: Nupe (Nigeria) (Banfield 1914)

- a. L (low) tone : bà ‘to count’ wà ‘to scratch’
 b. M (mid) tone : bā ‘to cut’ wā ‘to extract’
 c. H (high) tone : bá ‘to be bitter’ wá ‘to want’

(6) Four levels: Chatino (Yaitepec) (Mexico) (McKaughan 1954: 27)

L (low) tone	Lower mid tone	Higher mid tone	High (H) tone
kù	kü	kū	kú
‘dove’	‘sweet potato’	‘I grind’	‘I eat’

(7) Five levels: Kam (Shidong) (China) (Edmondson and Gregerson 1992)

- ṭa¹¹ ṭa²² ṭa³³ ṭa⁴⁴ ṭa⁵⁵
 ‘thorn’ ‘eggplant’ ‘father’ ‘step over’ ‘cut down’

The contrasts in (7) show that it is sometimes difficult to give names and use accent marks for each tone level, in which case it is more practical to indicate

pitch levels by numbers (5 = highest, 1 = lowest). In addition to such level tones, some languages also have contour tones which either rise, fall, or both. The best known such example is Standard Mandarin, whose four-way contrast is frequently exemplified by means of the minimal quadruplet in (8).

(8)	Tone I	(high level)	ma ⁵⁵	'mother'
	Tone II	(rising)	ma ³⁵	'hemp'
	Tone III	(falling-rising)	ma ²¹⁴	'horse'
	Tone IV	(falling)	ma ⁵¹	'scold'

By exploiting other contours as well as differences in phonation, for example, breathiness or glottalization, a language can have even more tonal contrasts on monosyllables, as in the following Trique (Itunyoso) [Mexico] examples (Dicario 2006):

(9)	Level	Falling	Rising	
	βe ⁴	li ⁴³	yāh ⁴⁵	'hair' 'small' 'wax'
	nne ³	nne ³²	yah ¹³	'plough (n.)' 'water' 'dust'
	nne ²	nne ³¹		'to tell lie' 'meat'
	nne ¹			'naked'

As many people do not speak a native language with tone, tonal contrasts such as the above can be quite unfamiliar to linguists and language learners alike:

Most language students, and even a shocking number of linguists, still seem to think of tone as a species of esoteric, inscrutable, and utterly unfortunate accretion characteristic of underprivileged languages – a sort of cancerous malignancy afflicting an otherwise normal linguistic organism. Since there is thought to be no cure – or even reliable diagnosis – for this regrettable malady, the usual treatment is to ignore it, in hope that it will go away of itself. (Welmers 1959: 1)

In this chapter we are concerned with whether and, if so, how tone is different from other aspects of phonology. While few phonologists would like to be identified with the above caricature, which Welmers (1973: 77) felt compelled to repeat 14 years later, there are occasional indications that a shocking number of linguists do indeed feel that tone is different from segmental phonology in rather dramatic ways. While we will conclude that there ARE important differences, let us first reject three rather extreme misconceptions about tone which are sometimes expressed:

(i) *Tone cannot be studied the same way as other phonological phenomena.* Upon encountering their first tonal experience, even seasoned field workers have asked me: "How can I tell how many tones my language has?" Each time this happens I am tempted to answer back with the rhetorical question: "How can you tell how many vowels your language has?" I have seen investigators try to discover the tonal categories by first eliciting long utterances, and then marking the relative

pitch changes between syllables, as one might initially do when approaching intonation. At best, this complicates the task. As in the case of voicing, nasality, vowel length, and other phonological contrasts, the normal technique is to first elicit individual words to determine the phonetic properties, and ultimately the phonemic contrasts. In the case of tone, this might then yield the tonal minimal pairs, triplets, quadruplets, and quintuplets seen in (4)–(8) above. Other languages may offer fewer minimal pairs or require specific contexts or *frames* in which the full range of contrasts can be discerned. For example, as seen in (10), the largely monosyllabic words of Hakha Lai [Burma] are pronounced with either HL falling (ˆ) or level L (ˉ) tone in isolation (Hyman and VanBik 2004):

(10)	in isolation		after ka=	'my'
	a. hmâa	'wound'	ka hmâa	'my wound'
		lûŋ	ka lûŋ	'my heart'
	b. kêe	'leg'	ka kêe	'my leg'
		kôoy	ka kôoy	'my friend'
	c. sâa	'animal'	ka sâa	'my animal'
		kâl	ka kâl	'my kidney'

However, as seen in the forms on the right, the falling tone nouns split into two classes when following toneless proclitics such as ka= 'my': the nouns in (10a) remain HL, while those in (10b) are pronounced with an LH rising (ˆ) tone. The natural conclusion to draw is that Hakha Lai has three underlying tones, /HL, LH, L/, and an LH → HL rule that applies after pause (and in certain other environments – see Hyman and VanBik 2004).

Such observations do not make tone different from other aspects of phonology – one has but to think of the languages which merge segmental contrasts on words in isolation, for example, the voicing contrast on German *Rat* 'advice' vs. *Rad* 'wheel'. In studying tone we need to be rigorous and comprehensive, just as we would have to be in attempting to analyze anything else. As we shall see, the issues that come up in the study of tone are quite complex, with more processes being available to tone, particularly at the phrase level, than to segmental phonology (cf. Section 4.2). However, since tones enter into paradigmatic contrasts very much like consonant and vowel features, we need not seek new methodologies. As Welmers (1959: 9) put it: "The more information we acquire about even the most complex tone systems, the more encouragement we receive that we already have the equipment needed to handle them." In short, tone CAN be studied just like other aspects of phonology.

(ii) *Tone cannot mark certain things.* A second misconception is that tone is used only to mark certain things. Most of the examples in (4)–(10) illustrate the LEXICAL function of tone: Different monomorphemic nouns, verbs, and so on, differ only in tone. However, the Somali examples in (3) show that tone can also be implicated in grammar, that is, it can have a MORPHOLOGICAL function. Further examples of morphological tone are seen in (11)–(13).

- (11) Tone marking person in Zapotec (Macuilianguis) (Mexico) (Broadwell 2000: 7)
- a. 1st person: bè-xáttā?-yà?-nà 'I ironed it'
 bè-gállá?-yà?-nà 'I hung it up'
- b. 3rd person: bè-xàttā?-nà-nà 'he ironed it'
 bè-gállá?-nà-nà 'he hung it up'
- (12) Tone marking tense-aspect in Nambikuára (Brazil) (Kroeker 1977: 129)
- a. present: xyău-nàřă 'he is staying' săsö-nàřă 'he is taking'
 b. past: xyàu-năřă 'he stayed' săsô-năřă 'he took'
- (13) Tone marking negation in Igbo (Aboh) (Nigeria) (elicited by the author)
- a. /ò jè kọ/ → ò jè kọ 'he is going'
 b. /ó jè kọ/ → ó jé kọ 'he isn't going' (H of subject /ó/ spreads onto /jè/)

In (11) we see that the first-person Zapotec verbs begin with an H tone, while the third-person verbs begin with L. In (12) the only difference between the present and past tense is tonal in Nambiquára (where $\hat{\text{~}}$ and $\check{\text{~}}$ mark HL falling and LH rising tones, respectively). Finally, in the Aboh dialect of Igbo, the only difference between the affirmative and negative utterances in (13) is the tone on the third-person subject pronoun /o/ (whose H spreads onto /jè/ 'go' in the second example). It is thus clear that tone can have a grammatical function as well as a lexical one.

The question is whether there are grammatical notions that tone cannot mark. In his Presidential Address at the 2004 Linguistic Society of America Annual Meeting, Ray Jackendoff proposed in passing the following alleged universal: "No language uses tone to mark case." The Maasai (Kenya) data in (14), however, provide a rather clear counterexample to this claim (Tucker and Ole Mpaayi 1955: 177–184):

- | | | | |
|------------|------------|------------|-------------|
| (14) | nominative | accusative | |
| class I: | èlòkònyá | èlúkúnyá | 'head' |
| | èncòmátá | èncómátá | 'horse' |
| class II: | èndéròni | èndèròni | 'rat' |
| | ènkòlòpà | ènkòlòpà | 'centipede' |
| class III: | òlmérègèsh | òlmérègèsh | 'ram' |
| | òlósówùàn | òlósówùàn | 'buffalo' |
| class IV: | òmótònyî | òmótònyî | 'bird' |
| | òsínkìrrî | òsínkìrrî | 'fish' |

In Maasai, native nouns usually consist of a gender prefix (masculine sg. *ol-*, feminine sg. *en-*) followed by a stem of one or more syllables. Except for class IV, nouns take different tones in nominative vs. accusative case. Although there are other complications, the four-syllable nouns which are cited show the following: in class I, the nominative has a single H on the last syllable, while the accusatives

have one L followed by all H syllables. In class II, nouns have a single H which is realized on the first stem syllable in the nominative vs. the second stem syllable in the accusative. In class III, the nominative has an H tone on the first and second stem syllables, while the accusative has an H only on the first stem syllable.

While Maasai is rather clear on the issue of tone marking case, the proposal is all the more surprising as so much of the tonal discussion in the 1960s and 1970s concerned the analysis of the so-called associative tone in certain West African languages such as Igbo (see Williamson 1986 and references cited therein). One analysis is that the associative marker consists of an H tonal morpheme which, as seen in the examples in (14), is assigned to the left in Central Igbo, but to the right in Aboh Igbo (Hyman and Schuh 1974: 98–99):

- (15) Central Igbo: àgbà + ´ + èṅwè → àgbá èṅwè ‘jaw of monkey’
 Aboh Igbo: ègbà + ´ + èṅwè → ègbà éṅwè ‘jaw of monkey’

In Igbo, the /L-L/ words *àgbà* ~ *ègbà* ‘jaw’ and *èṅwè* ‘monkey’ are pronounced L-L in isolation. As seen, according to the dialect, an intervening H tonal morpheme is realized either on the preceding or following TBU in the associative construction. But what is this “associative morpheme” if not a genitive case marker? The more pressing question is why anyone would seek to limit the kinds of constructions or semantic notions that tone can mark. The failure here is to appreciate the full morphological nature of tone: If a tone can be a morpheme, then it can do anything that a morpheme can do. This follows from the fact that tonal morphemes most commonly derive historically from earlier segmental+tonal morphemes whose segments have been lost (cf. Section 3). Thus, anything that can be marked by a segmental+tonal morpheme, can also be marked by a tonal morpheme. In other words, “tonal morphology . . . exhibits essentially the same range of morphological properties as in all of segmental morphology” (Hyman and Leben 2000: 588).

The alternative view, that tone is better suited to express certain ideas rather than others, has, however, occasionally also been expressed:

In a tone language, tone is not a purely harmonic or musical element, it is the expression of a thought, of an idea; it belongs to the intellectual domain, such that we can formulate the following axiom: SEEK THE IDEA AND YOU WILL HAVE THE TONE. (Stoll 1955: 5; my translation, his emphasis)

Along these lines, Stoll suggests that H tone indicates “everything that is woman, female, feminine” while L tone indicates “everything that is masculine, male, man” (p. 156). While certain languages occasionally exhibit indications of tonal iconicity (cf. Ratliff 1992), there is no reason for H vs. L tones to signify feminine and masculine any more than voiceless vs. voiced consonants (with which these tones often correlate, respectively). Once again, tone is not different from other phonological features.

(iii) *Tone is expendable.* The third misconception is that tone is somehow less essential in a language than other phonological features. This is seen in the way

tones and tone systems are sometimes dealt with in the literature. First, many descriptive and pedagogical grammars do not analyze or present the tone system in any detail. Those that do may have a section on tone, but then elect not to transcribe tone in the rest of the study. Different excuses are variously provided: (i) the tones are not important because they do not have a heavy functional load (i.e. there are few minimal pairs); (ii) the tones are a typographical inconvenience (e.g. an accent mark gets in the way of marking nasalized vowels with a tilde); (iii) the tones have not yet been analyzed; (iv) in order to learn the tones you have to listen to a native speaker anyway; (v) native speakers prefer not to write the tones, and so on. There are, of course, practical orthographies that do not mark tone, just as there are those which fail to mark all of the segmental contrasts. While there is a growing experimental literature on whether it is advantageous to readers if tone is marked, and if so, when and how (see, for example, Bird 1999a, b), the minimal contrasts cited in the above numbered examples establish that tone is extremely important in many, if not most, languages which have a tone system.

The view that tone is expendable is not limited to those designing practical orthographies. In general linguistic work, even phonologists commonly cite language data without indicating the tones. The most egregious cases occur when linguists REMOVE the tones from tone-marked examples cited from other sources, sometimes adding in a footnote that they have done so because “the tones are not relevant to the current study.” Perhaps this is also encouraged by an awareness that stress, another prosodic property (which may be predictable or have a low functional load), is also often omitted from linguistic transcriptions, as it is from many practical orthographies. However, no other phonological feature is treated with such indifference as tone: contrastive voicing on consonants or contrastive rounding on vowels is never removed from original fully-marked examples because voicing/rounding “is not relevant to the current study.” On the other hand, there are speech situations where the speakers themselves omit the segments in favor of the tones:

... when, for some physical reason, it is inconvenient for a Nambikuára to separate his teeth, he may still participate in a conversation by talking through closed teeth using tone as his principle means of communication. (Kroeker 1977: 133–134)

The inherent importance of tone should therefore not be underestimated.

As I have tried to indicate, although the above three misconceptions are sometimes explicitly expressed, they are more frequently implicit in the way linguists go about doing their work. Whether overtly expressed or not, the evidence is that tone is thought of as something different from the rest of what phonologists – or linguists in general – study. These misconceptions are both extreme and wrong-headed. However, the question still remains. Is tone significantly different from other aspects of phonology? For example, is tone more independent from other phonological features than these latter are from each other? Is tone capable of greater variability than other features? Are there special processes which are found only in tone systems? Or, as Leben (1973b: 117) once put it:

Is tone such a special phenomenon that it must be viewed as a feature on morphemes or larger units in some languages, as a feature on syllables in others, and as a feature on segments in still others? If so, then there is something left to explain: namely, why tone, unlike any other linguistic entity we know anything about, is capable of this many different types of representation.

Such questions are taken up in the following sections.

3 The Autosegmental Insight

Over the past several decades there has been an exponential increase in the documentation, analysis, and theoretical understanding of tone systems from throughout the world. Numerous formal proposals have been made concerning tone features and feature geometry, register effects such as downstep, and the interaction between tone and stress (de Lacy 2002), among other issues. In addition, the treatment of globality issues in tone has led to the development of Optimal Domains Theory (Cassimjee and Kisseberth 1998a). However, as mentioned in Section 1, the greatest impact of tone on phonological theory occurred in the 1970s when it provided the model for autosegmental phonology (Goldsmith 1976a, b). As a preliminary to the question of whether tone is “different,” it is therefore appropriate to begin by considering the fundamental insight of autosegmental tonology, stated in (16).

- (16) Tones are SEMI-AUTONOMOUS from the tone-bearing units on which they are realized.

This view constituted a reaction to the “segmental” approach of standard generative phonology (Chomsky and Halle 1968), which represented consonants and vowel segments in terms of a single matrix of binary distinctive features. For example, the vowel /a/ could be characterized by the vertical array of the binary feature values [–cons], [+back], [–round], and [+low], as in (17).

- (17) a. High tone [á] b. Rising tone [ǎ]

$\left[\begin{array}{c} -\text{cons} \\ +\text{back} \\ -\text{round} \\ +\text{low} \\ +\text{HIGH} \end{array} \right]$	$\left[\begin{array}{c} -\text{cons} \\ +\text{back} \\ -\text{round} \\ +\text{low} \\ +\text{RISING} \end{array} \right]$
--	--

The question was how to represent the various level and contour tones in terms of features. An ad hoc response was to add tonal feature values to the segmental matrices in (17), for example, [+HIGH] for H tone, [+RISING] for rising tone (the capitals being used so as not to confuse tonal and vowel height features). However, there are two problems with the representations in (17).

The first problem stems from the inherent claim that tones are inseparable features on segments, whereas tones have considerable autonomy from their TBUs. As argued by Leben (1973a, b), some languages have a limited number of suprasegmental “tonal melodies” which must be abstracted away from the TBUs on which they are realized. Although his example was Mende (cf. Leben 1978), I cite examples in (18) which illustrate the corresponding five tonal melodies of closely related Kpelle (Liberia) (Welmers 1962: 86):

- (18) a. High throughout
 pá ‘come’ ɓóá ‘knife’
 láá ‘lie down’ píli ‘jump’
- b. Low throughout
 kpòò ‘padlock’ kpàki ‘loom’
 tònɔ ‘chisel’ tòlon ‘dove’
- c. High followed by low (low begins on the next vowel if there is one)
 yê ‘for you’ tôa ‘pygmy antelope’
 kpônɔ ‘door’ kâli ‘hoe’
- d. Mid throughout
 kpɔŋ ‘help’ sua ‘animal’
 see ‘sit down’ kali ‘snake’
- e. Mid with first vowel, then high followed by low
 tɛê ‘black duiker’ konâ ‘mortar’
 yuû ‘axe’ kpanâŋ ‘village’

In the above, I have adopted Welmers’ practice of using only one tone mark per word (M is unmarked). He thus writes /kâli/ for what is pronounced [káli] ‘hoe’, that is, H-L. Second, there is no difficulty reducing Kpelle to an underlying two-level system: The M that occurs in the MHL melody in (18e) is straightforwardly analyzed as an L which is raised before H (cf. Section 4.2), and the “mid throughout” melody in (18d) is underlyingly /LH/, as is seen when two “mid throughout” words occur in sequence:

In mid-mid, for the dialect being described here, the first mid has a slightly rising allotone . . . In some areas, the first mid is level, but the second mid begins a little higher and drops quickly to the level of the first. In still other areas, both phenomena occur: the first mid ends a little higher, and the second begins a little higher. In all cases, the conjunction of two mids is accompanied by an upward pressure. (Welmers 1962: 87, Note 2)

The important observation is that only five tone patterns (or “melodies”) are possible independent of the number or nature of the TBUs: /H/, /L/, /HL/, /LH/, /LHL/. If, on the other hand, each TBU were capable of carrying an independent, underlying /H/, /L/, rising, or falling tone, we would expect $4 \times 4 = 16$ combinations on two TBUs, rather than the five that are observed. In this sense the tones are autonomous from the TBUs.

The second problem with (17) has to do with the representation of contour tones. Features such as [RISING] and [FALLING] encode a change within the segment which is not made explicit. In a two-level tone system, a rising tone typically acts as if it is a sequence of L+H realized on a single TBU, while a falling tone acts as if it is an H+L sequence on a single TBU. We see this in the following noun forms from Mende (Sierra Leone) (Leben 1978), which are arranged according to the same five tonal melodies as in Kpelle:

(19)		base noun		+ hu 'in'	+ ma 'on'
a.	/H/	kó	'war'	kó-hú	kó-má
b.	/L/	bèlè	'trousers'	bèlè-hù	bèlè-mà
c.	/HL/	mbû	'owl'	mbú-hù	mbú-mà
d.	/LH/	mbǎ	'rice'	mbà-hú	mbà-má
e.	/LHL/	nyàhâ	'woman'	nyàhá-hù	nyàhá-mà

As seen, the two locative enclitics =*hu* and =*ma* take their tone according to the tone of the preceding noun. When the noun ends in an H or L, as in (19a, b), the enclitic takes the same tone. When the noun ends in a falling (HL) tone, as in (19c, e) or a rising tone, as in (19d), the two parts of the contour "split": the first part goes on the last syllable of the noun, while the second goes on the enclitic. We thus see that when there is an extra available syllable, a falling tone maps as an H-L sequence and a rising tone maps as an L-H sequence. This then provides the evidence that contours should not be characterized by features such as [FALLING] and [RISING] but rather as sequenced level tone features, as in (20).

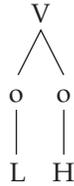
(20)	a. Falling tone [â]	b. Rising tone [ǎ]
	$\left[\begin{array}{c} -\text{cons} \\ +\text{back} \\ -\text{round} \\ +\text{low} \\ \text{V} \end{array} \right]$ $\begin{array}{c} \diagdown \quad \diagup \\ \text{H} \quad \text{L} \end{array}$	$\left[\begin{array}{c} -\text{cons} \\ +\text{back} \\ -\text{round} \\ +\text{low} \\ \text{V} \end{array} \right]$ $\begin{array}{c} \diagdown \quad \diagup \\ \text{L} \quad \text{H} \end{array}$

As proposed by Goldsmith (1976a, b), the H and L tones (or tone features) are represented on a separate tonal tier. Since they are both linked to the same TBU, the result is a contour tone. Other evidence that contours consist of two (occasionally more) independent tones linked to the same TBU can be cited from tone systems throughout the world. On the other hand, Yip (1989, 2002) argues that contour tones should be analyzed as units in certain Chinese dialects where the sequenced tone features appear to function as units. In such cases a rising tone would be represented roughly as in (21a), where the two tone features, L and H, are linked to a single tonal node.

(21) a. contour tone



b. tone cluster



(TBU)

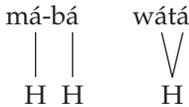
(tonal node)

This is what Yip refers to as a true contour tone vs. the more common “tone cluster” representation in (21b), where each tone has its own tonal node.

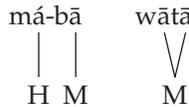
With such representations, we now understand the meaning of the “semi-autonomy of tone”: Tones are on a separate tier, but they are linked to their TBUs by association lines. Originally the proposal was that there were two tiers, a segmental tier and a tonal tier, but further elaborations were proposed to capture (i) the feature geometry of segments (Clements 1985; Clements and Hume 1995); (ii) the feature geometry of tones (Bao 1999; Snider 1999; Yip 1995, 2002), and (iii) the moraic and syllabic structures which organize the segments and serve as the TBUs to which the tones link. All of this was made possible by the autosegmental insight that tones are semi-autonomous from their TBUs. Support for this position has come from three general observations.

The first argument is that there is a NON-ISOMORPHISM between the two in the sense that the tones and TBUs do not necessarily synchronize: As we saw in (20) and (21), more than one tone can link to the same TBU, in which case we obtain a contour tone. The reverse, where one tone links to two TBUs, is also possible. As a result, Kukuya (Congo) (Paulian 1975; Hyman 1987) contrasts two kinds of H-H words. As seen in (22a), both *má-bá* ‘they are oil palms’ and *wátá* ‘bell’ are pronounced H-H in utterance-medial position:

(22) a. Medial



b. Prepausal



In (22b), however, the two words are realized differently before pause. As seen, there is an H → M rule which affects the last H feature before pause, not just the last H TBU. The contrasting representations in (22a), which had no equivalence in pre-autosegmental tonology, provide the structural difference that results in the surface opposition of H-M vs. M-M before pause. This difference largely correlates with a morphological difference: *má-bá* consists of two morphemes while *wátá* consists of one. In general, the double representation seen in *má-bá* is possible only when each H belongs to a different morpheme, while the branching representation of *wátá* is expected of single morphemes. There are exceptions in both directions, but in general the Obligatory Contour Principle (OCP) (Leben

- (25) a. /kíti/ 'animal' b. kíti ditó 'uncle's animal'
 /n̄juši/ 'chicken' n̄juši ðitó 'uncle's chicken'

Both of the nouns in (25a) are toneless, pronounced identically with a low falling contour before pause or an L tone. However, as seen in (25b), the realization of /ditó/ 'uncle' is different after the two nouns. In the case of 'uncle's animal', the three toneless TBUs are pronounced on a level mid pitch. In the case of 'uncle's chicken', the floating L of 'chicken' links to the first syllable of 'uncle' and the two toneless TBUs of *n̄juši* continue to be pronounced with a low falling contour. As seen in the underlying forms in (25a), nouns such as /n̄juši/ 'chicken' have a floating L tone after them (which derives from the loss of Proto-Mixtec final glottal stop, see Longacre 1957; Dürr 1987).

While the effects of lexical floating tones can be discerned in tone systems from all parts of the world, they are particularly common in Mexico and West Africa. Besides Peñoles, other Mixtec languages show the effects of a lexical floating L, for example, Atatláhuca (Mak 1953); a lexical floating H, for example, Chalcatongo (Hinton *et al.* 1991) and San Miguel el Grande (Pike 1948, Goldsmith 1990: 20–27); or both, for example, Acatlán (Pike and Wistrand 1974), Jacaltepec (Bradley 1970), and Magdalena Peñasco (Hollenbach 2004). Within Africa, the Grassfields Bantu languages are particularly well known for their complex floating tone systems (Voorhoeve 1971; Hyman and Tadadjeu 1976). In Aghem (Cameroon) (Hyman 1979b), although the two nouns *kí-fú* 'rat' and *kí-wó* 'hand' are pronounced identically as H-H in isolation, they have different effects on the tone that follows:

- (26) a. $\begin{array}{ccc} \text{kí-fú} & & \text{kí-m̄} \\ | & & | \\ \text{H} & & \text{L} \\ | & & | \\ \text{kí-wó} & & \text{kì-m̄} \\ | & & | \\ \text{H} & & \text{L} \end{array}$ 'one rat' 'one hand'
- b. $\begin{array}{ccc} \text{fú} & \text{kín} & \\ | & | & \\ \text{H} & \text{H} & \\ \text{wó} & & \text{↓kín} \\ | & & | \\ \text{H} & \text{L} & \text{H} \end{array}$ 'this rat' 'this hand'

As indicated by the dotted line in (26a), the H tone of the root *-fú* spreads onto the prefix of the numeral 'one' (pronounced *kí-m̄* in isolation), whose L tone then delinks. The H tone of the root *-wó* 'hand', however, does not spread. This is because it is followed by a floating L which belonged to a historically lost second syllable (cf. Proto-Bantu **-bókò* 'hand'). While much of the older work on Mixtec languages would have simply divided up H-H nouns into an arbitrary class A vs. class B, the floating L analysis has the advantage that it naturally accounts for the additional difference observed in (26b). Here the demonstrative *kín* 'this' (which conditions the deletion of the noun class prefix *kí-*) is realized H after *fú* 'rat', but as a downstepped H after *wó* 'hand'. As we saw in (23) and (24), a floating L frequently conditions downstep, as it does in Aghem. The floating L of /kí-wó/ is thus not circular. It blocks H tone spreading onto a following L and

conditions downstep on a following H, something which would not be naturally captured by a class A/B diacritic account.

It should be clear from the foregoing that the autosegmental representations, which express the traditional intuition of the semi-autonomy of tone, provide a more explanatory analysis in many cases. This does not mean that all tonal phenomena will have the properties illustrated in the preceding paragraphs. In some languages, when a vowel is deleted, its tone shows a stability effect, as above, while in others, for example, Shilluk (Gilley 1992: 164), the tone is deleted with the vowel. In addition, some languages may not distinguish intra- vs. heteromorphemic representations as in (22a). Worse yet, they may violate the OCP and show contrasts WITHIN morphemes. Odden (1982, 1986), for example, argues that in Shambala *nyóká* 'snake' has one doubly linked H tone vs. *ngó^htó* 'sheep', which, in violation of the OCP, has two underlying H tone features, the first causing the second to downstep. Similar issues arise concerning underlying L tone in Dioula (Odienné) (Ivory Coast) (Braconnier 1982):

- | | | | | | |
|------|----|--------------|----|----------|----------------------|
| (27) | a. | before pause | b. | before H | |
| | | sèbè | | sèbé | 'paper' |
| | | tùrù | | túró | 'oil' |
| | | kàràkà | | kàràká | 'bed' |
| | | sùmàrà | | sùmárá | 'soumbala' (a spice) |

As seen in (27a), the four monomorphemic nouns are pronounced all L before pause. When followed by an H, as in (27b), both the bisyllabic and trisyllabic nouns show two patterns: either one or two L tone syllables become H. The analysis may depend on how the rule is formulated. If the rule is as in (28a), the observed differences can be represented in one of two ways:

- | | | | | | | | | |
|------|----|------|---|------|---|---------|---|--------|
| (28) | a. | L | → | H | / | {//, L} | — | H |
| | b. | sèbe | | туру | | karaka | | sumara |
| | | | | ∨ | | ∨ | | ∨ |
| | | L L | | L | | L L | | L L |
| | c. | sèbe | | туру | | karaka | | sumara |
| | | | | ∨ | | | | ∨ |
| | | L | | L | | L | | L |

In (28b) a noun may have one vs. two L tone features, where only the L feature which precedes the H is raised to H. Depending on one's theoretical assumptions, this analysis potentially has two problems: (i) the forms with two Ls violate the OCP intramorphemically; (ii) the rule in (28a) is formulated as a feature-changing rule rather than a tone spreading rule – which is the most common way to express tonal assimilations (Section 4.1). In response to the first problem we might instead propose the representational differences in (28c). As seen, there is now only one /L/, linked either to one or two TBUs. As also seen, this /L/ can be preceded

features. We might therefore say that *tone is like segmental phonology in every way – only more so!* The phrase “only more so” can have two meanings: (i) QUANTITATIVELY more so: tone does certain things more frequently, to a greater extent, or more obviously (i.e. in a more straightforward fashion) than segmental phonology. (ii) QUALITATIVELY more so: tone can do everything segments and non-tonal prosodies can do, but segments and non-tonal prosodies cannot do everything tone can do. In fact, both of these characterizations are correct, as we shall see in the following two subsections.

4.1 *Quantitative Differences Between Tonal and Non-tonal Phonology*

In this section I discuss the following properties of tones, as they apply within the word domain: spreading, local shifts, non-local shifts, plateauing and polarity. Each of these is frequently attested in tone systems, but only the first is robustly attested in segmental phonology. I begin with the most common tonal process, tone spreading, whose properties are summarized in (31).

(31) Horizontal assimilation (or tone spreading) (Hyman 1975: 223)

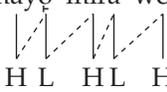
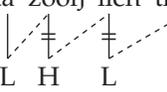
a.	Natural	b.	Unnatural
	L-H → L-LH		L-H → LH-H
	H-L → H-HL		H-L → HL-L
	(perseverative)		(anticipatory)

As seen, the term “horizontal assimilation” refers to cases where a full tone spreads onto a neighboring TBU. (Register effects or “vertical assimilations” are discussed in Section 4.2.) As is well known, even when there is no tone spreading, tonal targets tend to be reached late within their TBU: “Late realization of tonal targets has been demonstrated both for languages in which tones are lexical . . . and for those in which they are intonational . . .” (Kingston 2003: 86). As a consequence, phonological tone spreading tends to be perseverative, and where anticipatory, spreading is much rarer and has a quite different character (Hyman 2007b). This stands in marked contrast to what is usually said about segmental assimilations:

In regular conditioned sound changes, the conditioning factor is far more frequently a sound which follows than one which precedes. (Greenberg 1957: 90)

I examined 365 segmental assimilatory rules culled from 60 languages . . . documented in the Stanford Phonology Archive. 195 of these rules involved anticipatory assimilation of a segment to a following segment. 89 of these involved the perseverative assimilation of a segment to a preceding segment . . . The conclusion must be that segmental assimilation is generally anticipatory . . . (Javkin 1979: 75–76)

As seen in (32a), alternating sequences of input Hs and Ls undergo both H tone-spreading (HTS) and L tone spreading (LTS) in Yoruba (Laniran and Clements 2003: 207):

- (32) a. /máyo mírà wé/ [máyo mǐ rǎ wě] ‘Mayomi bought books’

- b. /kà zòonj lièn thùm/ [kà zòonj lièn thùm] ‘my three big monkeys’


Since phrase-internal contour tones are permitted in Yoruba, the result is a sequence of falling and rising tones. This contrasts with the situation in (32b) from Kuki-Thaadow (NE India, Burma) (Hyman 2010). Since this language does not allow phrase-internal contours, LTS and HTS condition delinking of the original tone except on the final syllable. The result is a bounded tone SHIFT: both the first H and the second L in the input in (32b) are realized only on the following TBU.

A similar relation between tone spreading and shifting is seen in the closely related Nguni Bantu languages in (33), where the processes are unbounded:

- (33) a. Ndebele b. Zulu
ú-kú-hlek-a u-kú-hlek-a ‘to laugh’
ú-kú-hlék-is-a u-ku-hlék-is-a ‘to amuse (make laugh)’
ú-kú-hlék-ís-an-a u-ku-hlek-ís-an-a ‘to amuse each other’

In (33a) the underlying (underlined) H tone of the initial prefix /ú-/ spreads up to the antepenultimate in Ndebele (Zimbabwe) (Sibanda 2004). The result is an H tone sequence spanning several syllables. However, as seen in (33b), the same H tone SHIFTS to the antepenultimate syllable in Zulu (South Africa) (Downing 1990: 265). In this case unbounded spreading + delinking has produced unbounded tone shift. Ndebele, thus, represents the older situation.

We can assume that both bounded and unbounded spreading have analogues in segmental phonology, where the most natural comparison is with processes such as vowel harmony. However, while it is very common for an underlying tone to shift several syllables to another position within the word (or onto a subsequent word, as will be seen in Section 4.2), there are very few cases reported where a segmental feature has this property. One such case in progress comes from Makonde (Mozambique) (Liphola 1999, 2001). As seen in (34), a process of vowel height harmony converts the applicative suffix /-il-/ to [-el-] after the mid root vowels /e/ and /o/:

- (34) underlying VH-harmony V-lengthening V-reduction
a. /ku-pet-il-a/ → ku-pet-el-a → ku-pet-eel-a ~ ku-pateela ‘to separate for’
b. /ku-pot-il-a/ → ku-pot-el-a → ku-pot-eel-a ~ ku-pateela ‘to twist for’

While the height harmony process is widespread in Bantu, as is phrase-penultimate vowel lengthening, Makonde appears unique in allowing /e/ and /o/ to reduce to [a] in pre-penultimate position. As seen, both inputs are potentially realized

b.	bet-idik-a	'lower'	bet-elek-ele	/bet-ilik-ile/
	kel-umuk-a	'turn around'	kel-omok-ele	/kel-umuk-ile/
			M ----- M	

Mid height plateauing represents a response to the precariousness of word-final /e/, which must either spread to another post-root vowel or be peripheralized to [i]. Thus, unlike most other Bantu languages, Yaka does not allow bisyllabic noun stems of the form CVCe (where V = any vowel). While H tone plateauing is quite frequent, the Yaka process is quite unique. It does, however, show that segmental plateauing is possible, however rare.

The last process to be considered in this subsection is polarity. Although having a number of manifestations, the best known case of tonal polarity occurs when an affix or clitic takes the opposite tone of its base or host. An example of this occurs in Margi (Nigeria) (Pulleyblank 1986: 203):

- (43) a. hògyì gù 'you are a Higi'
 b. màrgyí gù 'you are a Margi'

As seen, the subject clitic /gù/ 'you sg.' takes the opposite or polar value of the tone that precedes it: H after an L tone vs. L after an H. Note that in such cases, which are quite frequent, despite different possible solutions, it is often difficult to assign a unique underlying tone to the polar morphemes. The same point applies to cases of polar boundary tones: /H/ and /L/ are realized as LH and HL utterance-initially in Ticuna (Colombia) (Montes-Rodrigues 1995) and as HL and LH utterance-finally in Thlantlang Lai (Burma) (Hyman 2007a: 14). In these languages the boundary tone is polar to the adjacent lexical tone.

Since it would be arbitrary to propose a specific underlying tone in the above cases, tonal polarity differs from tonal (or segmental) dissimilation, where one of two identical specified features dissimilates (Hyman and Schuh 1974: 100). Thus, the unusual and mysterious low vowel dissimilation /CaCa/ → CeCa in Woleian and Marshallese and /CaCa/ → CiCa in Ere (Blust 1996) do not seem to be parallel. However, it is not out of the question that certain apparent polar effects owe their existence to a historical dissimilatory process. Consider, for example, the following alternating H tone pattern in Kirundi (Goldsmith and Sabimana 1986):

- (44) -sab- 'ask for' -báz- 'ask (question)'
- | | | | |
|----|-------------------------|-------------------------|-----------------------|
| a. | ku-sab-a | ku-báz-a | 'to ask' (infinitive) |
| | ku-bi-sab-a | ku-bí-baz-a | -bi- 'them' |
| b. | ku-bí-mu-sáb-a | ku-bí-mu-báz-a | -mu- '(to) him' |
| | ku-bí-mu-kú-sab-ir-a | ku-bí-mu-kú-bar-iz-a | -ku- '(for) you' |
| | ku-há-bi-mú-ku-sáb-ir-a | ku-há-bi-mú-ku-bár-iz-a | -ha- 'there' |
- 'to ask him (for) them for you there'

Ignoring the infinitive prefix *ku-* we see that up to three H tones may appear from the tone span that includes the object prefixes and the verb root. While Goldsmith and Sabimana account for the alternating H-L pattern in metrical terms, another

way to conceptualize the pattern is to assume an earlier H tone plateau (as Furere and Riiland 1983 report for closely related Kinyarwanda), to which a left-to-right H-H → H-Ø dissimilatory process subsequently applies. Either way, what makes the Kirundi facts interesting are the forms in the left column of (44a). When a toneless verb root occurs with one or no object prefix, the whole infinitive is toneless (vs. the corresponding forms involving an H verb root). However, as soon as a second object prefix is added in (44b), we obtain not one H, but two. It would seem that as soon as one H is introduced by the morphology, it must be alternated within the tone span. While the Kirundi facts are reminiscent of alternating stress, I am unaware of any parallel case where a segmental feature is assigned on an alternating basis.

In this section we have examined several tonal phenomena as they apply roughly within the word domain. In each case it was suggested that the tonal process is more natural and frequent than its segmental analogue. One way to look at this is to say that tone is less restricted than segmental phonology. This interpretation receives support from a common restriction in segmental phonology which seems rarely to apply to tone: root control (Clements 1981). Whereas many harmony processes involve the assimilation of (underspecified) affixes to segmental features of the root, even word-level tonal assimilations seem largely to apply across-the-board. To take just one example, it has been oft noted that prefixes rarely, if ever, condition vowel harmony on a following root (Hall and Hall 1980: 227n). However, it is quite commonplace for the tone of a prefix to spread onto the following root, as seen in the Aghem examples in (45):

- (45) a. HTS: /kí-kɔ̀ɔ̀/ → kí-kɔ̀ɔ̀ ‘cutlass’
 b. LTS: /kì-tɛ̀ɛ̀/ → k̀ì-tɛ̀ɛ̀ ‘cricket’

It would appear that the perseverative tendency for tonal targets to be realized late overrides any counteracting tendency for prefixal tones to assimilate to roots. Since segmental assimilations were said to have an apparent anticipatory bias, we make the following statistical prediction: Roots will tend to assimilate to the tones of prefixes, but to the segmental features of suffixes. While we should not expect this prediction to be without exception, I believe that the asymmetry is quite real (Hyman 2002).

4.2 *Qualitative Differences Between Tonal and Non-tonal Phonology*

In the preceding section, the comparison of tonal to segmental processes was largely limited to the word domain. In this section we will consider tonal properties that apply across words. As will be seen, once we do so, the differences become more pronounced. We begin by considering “register effects” and then move on to discuss long-distance processes that apply at the phrase level.

In Section 4.1 it was said that tones tend to perseverate in “horizontal” assimilations. This is true only if we are looking at full tone assimilation. A second

possibility is for tones to undergo “vertical assimilation” or register adjustments. In this case, the asymmetries are quite different, as summarized below:

(46)

		Compression	
a. Input:		Anticipatory	Perseverative
L-H		M-H	L-M
H-L		? M-L	? H-M

		Expansion	
b. Input:		Anticipatory	Perseverative
L-H		? ↓L-H	? L-↑H
H-L		↑H-L	H-↓L

As seen in (46a), either tone of a /L-H/ input can undergo a pitch register adjustment and ultimately become a third level tone: the L may be raised to M, the H may be lowered to M, or as we saw in Kpelle in (18d), both may apply in which case /L-H/ is realized [M-M]. The result is tonal COMPRESSION: the difference in the interval between the two output tones becomes smaller than in the input. An input /H-L/ does not show this effect. As indicated by the question marks, it is quite unusual for the H to be lowered or the L to be raised. In fact, as seen in (46b), the opposite effect of EXPANSION is observed: /H-L/ (but not /L-H/) may undergo raising of the H or lowering of the L. The raising process is seen in the Engenni (Nigeria) example in (47a) (Thomas 1974: 12):

(47)

a. /únwónì/ ‘mouth’ ↓ [únwónì]	b. /únwónì + ólíló/ ‘mouth of a bottle’ ↓ [únwón ólíló]
--------------------------------------	---

In Thomas’ analysis a single H TBU is raised to a “top” or superhigh (˩) tone before an L. In (47b) the L tone vowel /i/ is elided, but still has the raising effect on the preceding tone. The result is a surface contrast between the H and superhigh tone.

H tone raising before L is a quite widespread phenomenon, also occurring for example in Gurma (Burkina Fasso) (Riailand 1981, 1983), Kiriimi (Cahi) (Tanzania) (Hyman 1993), Edopi (Indonesia (Papua)) (Kim 1996), and Chinantec (Lealao) (Mexico) (Rupp 1990). Tesfaye and Wedekind (1990: 360) report that in Shinasha (Ethiopia) an H-L drop is realized “with about four semitones” while an L-H rise “is realized as a pitch increase of only two, sometimes three semitones.” The question is whether there is a physiological reason for such differences or whether H-raising is a strategy for maximizing the tonal space – or both? Many languages have “downdrift” or “non-phonemic” downstep whereby the second H of an H-L-H sequence is realized on a lower pitch than the first. If there are enough transitions from H to L to H, Hs which are late in the utterance may become quite low. Since

it increases the tonal space, raising an H before an L could thus be useful as a pre-planning counterforce to processes which lower tones (Rialland 2001).

Some support for this interpretation is obtained from languages which raise H tone in anticipation of a contrastive downstep (\downarrow H). In languages which contrast H-H vs. H- \downarrow H there can be multiple downsteps within an utterance, with each \downarrow H being produced at a lower pitch level than the previous H. Starting at a higher level could thus be quite useful. In this context consider the following pairs of examples from Amo (Nigeria) (Hyman 1979a: 25) and Luganda (personal notes with Francis Katamba):

(48) Non-local H \rightarrow \uparrow H in anticipation of (long-distant) phonemic downstep

- | | | |
|----|---|--------------------------------------|
| a. | kité úkóómí fínáwà | 'the place of the bed of the animal' |
| | kí \uparrow té úkóómí fíká \downarrow lé | 'the place of the bed of the monkey' |
| b. | à-bá-síb-á kígùùndú | 'the ones who tie up Kigundu' |
| | tè- \uparrow bá-síb-á \downarrow kígùùndú | 'they do not tie up Kigundu' |

In each pair the second utterance contains a \downarrow H which is lacking in the first. As indicated, but not usually transcribed, a sequence of Hs is quite audibly raised in anticipation of the downstep. Thus, between the first and second syllables, there is a step up of [+1] in the examples lacking a \downarrow H vs. [+2] in the examples having a later \downarrow H. While the raising of H before L appears to be local, it is striking how early the first H is raised in anticipation of the non-local H- \downarrow H interval. As we shall see below, other such long-distance interactions are quite typical of phrasal tonology.

The properties of what I refer to as canonical downstep (Clements 1979; Hyman 1979a) are as follows: (i) H contrasts with \downarrow H only after another (\downarrow)H; (ii) \downarrow H establishes a "ceiling effect" until the register is re-set: thus, a following H will be pronounced at the same pitch level rather than higher; (iii) there is a theoretically unlimited number of downstep pitch levels (H- \downarrow H- \downarrow H- \downarrow H- \downarrow H...). While such systems are best known from Sub-Saharan Africa, they are in fact found in languages from throughout the world, for example, Kuki-Thaadow (NE India, Burma) (Hyman 2010), Kairi (Rumu) (Papua New Guinea) (Newman and Petterson 1990), Mixtec (Coatzospan) (Mexico) (Pike and Small 1974), Tatuyo (Colombia) (Gomez-Imbert 1980), among many others. Downstep is, thus, a very natural tonal phenomenon. The question, then, is what, if anything, corresponds to downstep in segmental phonology? The brief answer: nothing. Again, we might look to vowel height for a parallel. Recall from (39) the transfer of vowel height features from root to prefix in Esimbi. The eight underlying root vowels are exemplified in (49) and as they are realized in their singular and plural forms:

- | | | | | |
|------|----------|----------|------------------|------------------|
| (49) | URs | gloss | class 3 sg. /u-/ | class 6 pl. /a-/ |
| a. | /-tili/ | 'end' | u-tili | o-tili |
| | /-wúsu/ | 'fire' | u-wúsu | o-wúsu |
| b. | /-yembe/ | 'song' | o-yimbi | ε-yimbi |
| | /-góro/ | 'foot' | o-gúru | ɔ-gúru |
| | /-námə/ | 'tongue' | o-ními | ɔ-ními |

c.	/-sɛmɛ/	'grain'	ɔ-simi	a-simi
	/-gòɔ/	'disease'	ɔ-gùnu	a-gùnu
	/-kara/	'rope'	ɔ-kiri	a-kiri

When occurring with the class 3 singular prefix /u-/, there are no complications: the vowel height of the root simply transfers to the prefix (and all root vowels are pronounced [+high]). However, when the prefix is plural class 6 /a-/, we observe that it is one step lower than the corresponding singular. This is obtained by fusing the transferred root vowel height with the lower vowel height of /a-/. (Although this would predict that /a-/ should be realized a- or A-, secondary processes modify these impermissible outputs to o-, e-, and ɔ-.) The fact that there is a step-wise lowering of the prefix might suggest that /a-/ functions like a downstep marker operating on vowel height (Hyman 1988: 263). However, neither the Esimbi facts nor any other such scalar segmental process shows the properties of canonical tonal downstep: Thus, there is no language where hypothetical $\downarrow i, \downarrow u$ (perhaps pronounced [ɪ, ʊ] or [e, o]) contrast only after /i, u/, as $\downarrow H$ contrasts only after another H. There also is no "ceiling effect" on subsequent vowels, such that $C^{\downarrow}iCe$ and $C^{\downarrow}eCe$ are pronounced [CeCe] and [CɛCæ], respectively – and there certainly is no such effect on subsequent words in the phrase, as in the case of tonal downstep. Since tone and vowel height otherwise share properties, for example, their gradiance along a single F0 vs. F1 dimension, the only conclusion to draw is that register effects such as downstep make tone qualitatively different from segmental features.

In fact, perhaps the most significant difference between tone and segmental phonology concerns the ability of tonal processes other than register to apply at long distances at the phrase level. Recall from (33) that Ndebele spreads, while Zulu shifts an H tone to word-antepenultimate position. In other Bantu languages an H tone spreads or shifts to a designated syllable in a following word. The words in the Shambala utterance in (50a) are all underlyingly toneless, and are therefore pronounced all L (Philippson 1998: 320):

- (50) a. mawe magana mane na= milongo mine '440 stones'
 b. magí mágána matátú ná= mílóngo mine '340 eggs'

However, in (50b) the two underlying H tones (originating on the underlined vowels of /magí/ 'eggs' and /matátu/ 'three') spread to the penultimate syllable of the following phonological word (or clitic group). Corresponding to the word-level spreading vs. shifting difference of Ndebele vs. Zulu in (33), the following Giryama examples show a long-distance rightward shift or displacement to the penultimate syllable of the following word (Philippson 1998: 321; cf. Kisseberth and Volk 2007):

- (51) a. ku-tsol-a ki-revu 'to choose a beard' /-tsol-/ 'choose'
 b. ku-on-a ki-révu 'to see a beard' /-ón-/ 'see'

As in Shambala, the words in (51a) are both toneless, and the phrase is pronounced all L. In (51b), the verb root /-ón-/ ‘see’ carries an H tone which is displaced onto the penultimate of the toneless word /ki-revu/ ‘beard’. It is examples such as these which motivate Yip (2002: 133) to remark that “the most striking property of African tone is its mobility”. One looks in vain for a phrasal spreading or displacement of a segmental property in this way: Vowel-, consonant, and nasal harmonies, which appear to most closely mimic the effects of tone spreading, are typically limited to a word-size domain (which may include clitics). In the rare cases where vowel harmony hops over a word boundary, the affected target is typically a grammatical morpheme, as in Kinande, or the process instead represents a local coarticulatory effect, as in Nez Perce (Aoki 1966). In (51b) we have a robust case of an H tone shifting from one lexical word to another – something which is exactly duplicated in a number of other Bantu languages, for example, Digo (Kisseberth 1984: 163–164) and Zigula (Kenstowicz and Kisseberth 1990: 175).

While anticipatory spreading and shifting were said to be rare at the word level, phrasal anticipation appears to be quite natural. Consider the following forms from Tiriki [Kenya] (Paster and Kim 2007) in (52).

- (52) /xu-molom-el-a/ ‘to speak for’ /xu-rhúmùl-il-a/ ‘to hit for’
- | | | | |
|----|-----------------------|--------------------------------------|--------------------|
| a. | xù-mòlòm-èl-à mù-lìmì | xù-rhúmùl-il-à mù-lìmì | /mu-limi/ ‘farmer’ |
| b. | xù-mólóm-él-á mú-línà | xù-rhú [↓] múl-íl-á mú-línà | /mu-lína/ ‘friend’ |

Both /xu-molom-el-a/ and /mu-limi/ are underlyingly toneless and pronounced all L in isolation, while /xu-rhúmùl-il-a/ and /mu-lína/ have an underlying H on their first root syllable, which spreads onto the prefix: [xù-rhúmùl-il-à], [mú-línà]. In (52a), where the infinitive is followed by toneless /mu-limi/, no further change is observed. In (52b), however, the H of /mu-lína/ is anticipated not only onto its prefix /mu-/, but also onto all of the toneless TBUs of the preceding word. As seen, the toneless verb /xu-molom-el-a/ becomes all H, while the H of /mu-lína/ is anticipated up to the H of /xu-rhúmùl-il-a/. Where the two H tones meet, we observe the indicated downstep. The same anticipatory process will apply through multiple words and phrases: /xu-rhúmùl-il-a + mu-limi + mu-lína/ → [xù-rhú[↓]múl-íl-á mú-límì mú-línà] ‘to hit the friend for the farmer’.

Recall the Luganda H tone plateauing process which was illustrated word-internally in (40b). As seen in (53), the process actually applies across words within certain postlexical tonal domains (Hyman *et al.* 1987: 89):

- (53) a. y-a-láb-à ‘he saw’ bi-kópò ‘cups’ by-aa= Wálúsìmbi ‘of Walusimbi’
 H L H L H L
- b. y-a-láb-á bí-kópó by-áá= Wálúsìmbi ‘he saw the cups of Walusimbi’
 H ----- H ----- H L
 L L
 ↓ ↓
 Ø Ø

As indicated, each of the three words in (53a) has an H to L pitch drop. However, when combined in (53b), the whole sequence from the first to last H is pronounced on an H tone plateau. This is obtained by deleting Ls which occur between Hs, followed by plateauing.

As seen in the inputs in (53a), Luganda contrasts H, L, and Ø at an intermediate stage of representation (at the output of the word phonology). Any word-level Ls which occur between Hs are deleted within the proper tonal domain, no matter how distant the TBUs of the surrounding Hs may be. Another deletion process which takes place at a distance occurs in Peñoles Mixtec (Daly and Hyman 2007). In this language the underlying tones are /H/, /L/, and /Ø/, with an OCP(L) constraint triggering the following L tone deletion rule in (54a):

- (54) a. $L \rightarrow \emptyset / L _$
 b. $\text{ɿ}^N \text{d}ij\text{-ni-kwe-}š\text{i} \text{ kada-kwe-}š\text{i} \text{ ɿ}^N \text{ ɿ}^N \text{ ɕ}u^N \rightarrow \text{ɿ}^N \text{d}ij\text{-ni-kwe-}š\text{i} \text{ kada-kwe-}š\text{i} \text{ ɿ}^N \text{ ɿ}^N \text{ ɕ}u^N$
 one alone-only-pl-she POR.do-pl-she one one work L Ø
 ‘Only one of them will do each of the jobs.’

This rule of L tone deletion applies across any number of toneless TBUs which may intervene. The rule thus applies to the second underlying /L/ in (53b), where there are 12 intervening toneless TBUs occurring between it and the preceding /L/. Like the cases of tone spreading, shifting, and plateauing, such extreme deletion at a distance is without parallel in segmental phonology.

While the above examples have to do with natural tonal processes which apply at a distance, another widespread phrasal phenomenon concerns cases where the tones of certain constructions are uniquely determined by the tones of the first morpheme or word. A well-known example concerns Shanghai compounds and other tightly bound constructions (Zee and Maddieson 1979: 109). As seen in the two examples in (55), this also is obtained by a two-step deletion + spreading process. All but the first tone is deleted, after which the second of the two features of the tonal contours is assigned to the second syllable. In the examples, the third syllable acquires a default L tone.

- (55) a. ‘illuminate’ ‘symbol’ ‘machine’ → ‘camera’
 $\begin{matrix} \text{tsɔ} & + & \text{ɕiã} & + & \text{tɕi} & \rightarrow & \text{tsɔ} & \text{ɕiã} & \text{tɕi} & \rightarrow & \text{tsɔ} & \text{ɕiã} & \text{tɕi} \\ \wedge & & \wedge & & \wedge & & \wedge & & & & | & | & \vdots \\ \text{M} & \text{H} & \text{M} & \text{H} & \text{M} & \text{H} & \text{M} & \text{H} & & & \text{M} & \text{H} & (\text{L}) \end{matrix}$
- b. ‘sky’ ‘studies’ ‘terrace’ → ‘observatory’
 $\begin{matrix} \text{t}^h\text{i} & + & \text{vəŋ} & + & \text{dɛ} & \rightarrow & \text{t}^h\text{i} & \text{vəŋ} & \text{dɛ} & \rightarrow & \text{t}^h\text{i} & \text{vəŋ} & \text{dɛ} \\ \wedge & & \wedge & & \wedge & & \wedge & & & & | & | & \vdots \\ \text{H} & \text{L} & \text{L} & \text{H} & \text{L} & \text{H} & \text{H} & \text{L} & & & \text{H} & \text{L} & (\text{L}) \end{matrix}$

While the tone melodies of languages such as Kpelle, seen earlier in (18), are normally a property of words, the Shanghai data show that tones may also be mapped over phrases. In fact, it is quite common for the tones of syntactically

conditioned constituents to be determined by the underlying tone or tone pattern of the first word. Based on their tonal behavior, Efere (2001b: 158–159) sets up the four following classes A–D in Izon (Bumo, Nigeria):

- (56) class schema tone pattern determined by the A–D class of the phrase-initial word
- | | | |
|---|------------|--|
| A | (L) H + H | all TBUs in the phrase = H |
| D | (L) H + HL | first word = all H, H spreads one TBU to the right; other TBUs = L |
| B | (L) H + L | first word = all H; subsequent TBUs = L |
| C | (L) HL + L | first word keeps its HL drop, remaining TBUs = L |

The (L) in parentheses refers to the fact that vowel-initial words can begin H or L, whereas consonant-initial words begin H. Since only the tones of the first word are relevant, we can assume, as in Shanghai, that the tones of non-initial words are first deleted. The four classes are illustrated in (57) in the frame . . . /náná kímí/ ‘man who owns/has . . .’ (whose underlying H tones are deleted):

- (57) A (L) H + H bélé ‘pot(s)’ → bélé náná kímí (H spreading to end)
 D (L) H + HL ikíé ‘friend’ → ikíé nánà kìmì (H spreading one TBU)
 B (L) H + L wárí ‘house’ → wárí nà nà kìmì (no H spreading)
 C (L) HL + L sérí ‘scarf’ → sérí nà nà kìmì (no H spreading)

As indicated, classes A, B, and D have a first word which is all H (class C words, which have a pitch drop, are largely borrowings). They differ by the degree to which this H affects subsequent words within the tonal phrase: all the way to the end (A), one TBU onto the second word (D), no spreading (B).

While Williamson (1988) uses a system of tonal melodies + floating tones to account for the variations in such tone patterns found in related Ijoid languages and dialects, at some point it becomes quite difficult to predict the phrasal patterns from specific underlying tones on the first word which determines them. As a case in point, in Urarina (Peru), tone is also determined by the first word of the phrase, which (Olawsky 2006) groups into the four classes A–D:

- (58) class tone pattern determined by A–D class of the phrase-initial word
- | | |
|---|--|
| A | first word = L; H is assigned to initial syllable of following word |
| B | first word = L; H is assigned to 3rd syllable of following word |
| C | first word = L; H is assigned to last syllable of final word of phrase |
| D | first word keeps its H tone when a word follows, all the rest = L |

In isolation, Urarina words generally have a single final H. Whereas the first word is all H in three out of the four classes in Izon, in Urarina the first word of a phrase is L except in class D (which also includes some words that have their H on the penultimate rather than final syllable). As summarized above and illustrated

when preceding the trisyllabic verb *rɯ.a.kaa* ‘carries/carried (3sg.)’ below, the difference concerns the placement of an H tone on the next word:

- (59) A raaná ‘peccary (sp.)’ → raana rú.a.kaa ‘he has carried peccary’
 B obaná ‘peccary (sp.)’ → obana rɯ.a.káa ‘he has carried peccary’
 C reemaé ‘dog’ → reemae rɯ.a.káá ‘he has carried dog’
 D makusajarí ‘pepper’ → makusajarí rɯ.a.kaa ‘he has carried the pepper’

The above represents the general case. One complication is that class B will assign the H to the second syllable of a trisyllabic word whose last syllable is short. It is possible to treat class C as assigning a toneless pattern (with the phrase-final mora being due to a default rule) and class D as borrowings and exceptions. This would still leave the difference between class A and class B. One is tempted to start with an H assigned to one syllable which then shifts two or three syllables to the right, but there does not appear to be any evidence for this, hence Olawsky’s class A–D approach.

The same conclusion is suggested by the facts from the Move dialect of Yagaria (Papua New Guinea). As indicated in (60), according to Ford (1993: 196–197) words either have stable (S) tones or belong to one of three “unstable” tone classes (U_1 – U_3):

- (60) a. S hógà ‘left’ → hógà kàyàlè ‘left pig’ (no change + all L)
 S fáipái ‘white’ → fáipái kàyàlè ‘white pig’ (no change + all L)
 b. U_1 lòlé ‘two’ → lòlè kàyàlé ‘two pigs’ (all L + L-L-H)
 U_3 fèlá ‘wild’ → fèlà kàyálé ‘wild pig’ (all L + L-H-H)
 U_2 kòlí ‘scared’ → kòli káyálé ‘scared pig’ (all L + H-H-H)

In the above examples the second word is /kàyálé/ ‘pig’, which becomes all L after a stable tone word, as in (60a). In (60b), all three unstable tone words have L-H tone in isolation. As seen, they have different effects on the next word: U_1 places a single H on the last syllable, U_2 assigns H to the last two syllables, and U_3 assigns H to all three syllables. One interpretation might be to view the second word as becoming toneless after a stable tone word (and hence all L). Unstable tone words would lose their H tone, and assign a single H to one syllable of the next word, which then spreads onto following syllables. In this case, we would obtain /kayalé/ (after U_1), /kayále/ (after U_2), and /káyale/ (after U_3). It is not clear how these different H tone placements could be predicted from different underlying representations on the first word. While the assignment of a single H to the next word in Urarina and Yagaria might suggest “accent-like” behavior, it is important to note that stress-accent systems do not appear to show such effects. Thus, there appears to be no language where class A words assign a final stress to the next word, class B a penultimate stress, and class C an initial stress. Rather, phrase-level arbitrary classes seem to be a tonal phenomenon.

The final evidence that more is going on than a simple mapping from the first word to the phrase is found in Wuxi, a Northern Wu dialect that has been studied

in detail by Chan and Ren (1989). The following table is reproduced from Chen (2000: 323) (cf. also Yip 1989: 156–159):

(61)

1st σ			2nd σ		
			T3, 4, 5, 6	T1, 2	T7, 8
A LLH	T3	→	C LHH		
	T6	→	D LHL		
B HHL	T1	→	A LLH		
	T4	→	A LLH	D LHL	
	T7q	→	A LLH	D LHL	B HHL
C LHH	T5	→	B HHL		
D LHL	T2	→	B HHL		
	T8q	→	A LLH	D LHL	

Synchronic Wuxi contrasts four surface tones (A–D), which, as indicated, are characterized in terms of four tone features (which Yip 1989 reanalyzes, respectively, as L+LH, LH+L, H and L+HL, where LH and HL are tonal contours in the sense of (21a)). These tones derive from various mergers of the Middle Chinese eight categories (T1–T8), where T7q and T8q refer to the two tones on stopped syllables. Wuxi, like Shanghai, maps a single tone pattern over a phrasal domain. However, as seen in (61), the exact shape of the pattern is determined not only by A–D identity of the first tone, which represents mergers of the Middle Chinese tones, but also by a three-way contrast in the historical identity of the second tone. There thus is considerable arbitrariness from a synchronic point of view. As Chan and Ren (1989) point out, Wuxi tone sandhi was originally right-dominant (whereby the first tone changes), but then became left-dominant, deleting the second tone, as in Shanghai. Chen (2000: 325) provides the following sample derivation to illustrate:

- (62) a. na + dā ‘milk candy’
 HHL LHL (base tones)
 LHL LHL (first tone undergoes sandhi)
 b. LHL Ø (deletion of second tone)
 LH HL (tone spread)

In (62a) the initial HHL (B) tone changes into an LHL (D) tone. This is followed in (62b) by first deleting the LHL (D) tone of the second morpheme, which then

allows the one LHL tone to map over the two syllables. The result is an alternation between the /HHL/ of the first syllable with a /LHL/ which it maps over the two syllables. The Wuxi situation demonstrates the extent to which synchronic tonal properties can encode the history of a language, particularly at the phrase level. It would be extremely interesting to know the historical origins of the Izon, Urarina, and Yagaría systems.

We have been concerned with the fact that while a phrase-initial word can assign an arbitrary tone to a following word, this does not seem to be true either of segmental phonology or of stress. There is no language where, say, the first word arbitrarily assigns a [+nasal] to different syllables of the next word, nor are there stress classes that assign different stresses to a following word. It should be noted that the examples cited above are all head-final, such that it is an initial modifier or complement which determines the tonal properties of the phrase. This contrasts with grammatical feature assignments which normally go from head to dependent. In this context consider the highly unusual tonal agreement from Barasana (Colombia) (Gomez-Imbert and Kenstowicz 2000: 438–439), which has an underlying contrast between bimoraic morphemes which are all H vs. HL. As seen in (63), possessive pronouns condition tonal agreement on the following noun:

(63)

	~kúbú (H) 'shaman'	~bídi (HL) 'pet'
~bádí (H) 'our'	~bádí ~kúbú	~bádí ~bídí
~ídà (HL) 'their'	~ídà ~kúbù	~ídà ~bídi

When the pronoun is H, as in the case of ~bádí 'our', ~bídi 'pet' changes to H. Similarly, when the possessive pronoun is HL, as in the case of ~ídà 'their', ~kúbú 'shaman' changes to HL. Since agreement is normally a grammatical concept, one might be tempted to consider the above facts not as tonal agreement, but as grammatical agreement (where H and HL are exponents of grammatical features). However, the agreement in Barasana goes from modifier to head, hence again, in the wrong direction. It is tempting to instead see the examples in (55)–(63) as cases where the head (and potentially other non-initial words) undergoes reduction, followed by additional processes of tone spreading, tone assignment, tone agreement and so on.

To summarize this section, we have seen that tone is capable of reaching deeply across word boundaries for both grounded and not-so-grounded processes, for example, downstep anticipation, non-local H spreading, shifting and plateauing, OCP effects, tone mapping. Segmental phonology does not have such a long reach, but rather is restricted to local adjustments, as when the last segment of one word interacts with the first segment of the next. Stress offers more possibilities, particularly when the stress of one word is retracted or deleted so as to avoid a clash with the stress of another word. It too, however, does not show all of the parallels

illustrated in the above discussion. If tone is really different, then these findings have the following implication for word-prosodic typology: any system that does what only tone can do is tone. The issue of typology is taken up in the following section.

5 Tone vs. Accent

While I have treated each of the phenomena and all of the examples in the preceding sections as strictly tonal, there has been a tendency to view languages which restrict the distribution of their tones as “accentual.” For example, since Urarina assigns a single H within the appropriate phrasal domain, might this H be a “pitch accent?” However, since most of the properties discussed in Section 4.2 do not have analogues in stress-accent systems, their identification as accentual phenomena is not obvious. In this section, I will argue that the phenomena in question are typical of tone systems, particularly those which place restrictions on the distribution of their tones.

Within the generative tradition, the study of word-prosodic typology was greatly influenced by McCawley (1968b, 1970), who attempted to set up a principled distinction between tone vs. pitch-accent systems based both on distributional properties and rule types (tones tend to assimilate; accents tend to dissimilate or reduce). A survey of subsequent literature reveals that the terms “accent,” “pitch accent,” and “tonal accent” have generally been used to refer to tone systems which are defective in the sense of restricting tones by number of contrasts or by position: “A pitch-accent system is one in which pitch is the primary correlate of prominence and there are significant constraints on the pitch patterns for words . . .” (Bybee *et al.* 1998: 277). Among such “significant constraints” are those enumerated in (64), where the tone in question is most commonly /H/:

- (64) A tone may be . . .
- a. obligatory: “at least one” must occur per domain (e.g. word)
 - b. culminative: “at most one” can occur per domain
 - c. privative: the underlying contrast is between presence vs. absence of the tone
 - d. predictable: assigned to positions by rule
 - e. restricted: occurring in only in certain positions (stressed syllable, first two syllables)
 - f. reducible: subject to reduction, subordination (e.g. in compounding, defocusing)

However, most or all of the above properties can be found in unambiguous tone systems. For example, consider OBLIGATORINESS as reflected in the attested tone patterns of Chuave (Papua New Guinea) (Swick 1966; Donohue 1997: 355) in (65).

(65) scheme:	/H/	/HL/	/LH/	/HLH/	/LHL/
1 σ = 1	H				
2 σ = 3	H-H	H-L	L-H		
3 σ = 7	H-H-H	H-H-L H-L-L	L-H-H L-L-H	H-L-H	L-H-L
4 σ = 15	H-H-H-H	H-H-H-L H-H-L-L H-L-L-L	L-H-H-H L-L-H-H L-L-L-H	H-H-L-H H-L-H-H H-L-L-H	L-H-H-L L-H-L-L L-L-H-L

As seen, all combinations of H and L tone are found on words of one to four syllables, except an all L pattern. H tone is therefore “obligatory,” but hardly accentual, given, for example, the 15 patterns possible on four-syllable words. Numerous other tone systems have two, three, or four word-tone patterns requiring an H, for example, /H, LH/ in Hup (Colombia) (Epps 2005), /H, HL, LH/ in Dom (Papua New Guinea) (Chida 2001), /H, HL, LH, LHL/ in Dogon (Jamsay) (Mali) (Heath 2008). It is not clear that there is anything special about such obligatory-H systems vs. those which allow a /L/ pattern. In fact, the distributions in (66) suggest that the obligatoriness of H in Tanimuca (Colombia) (Keller 1999) may be accidental:

(66) 2 σ = 3	H-H	H-L	L-H				*L-L
3 σ = 6	H-H-H	H-H-L	H-L-H	L-L-H	L-H-H	L-H-L	*L-L-L, *H-L-L

All patterns occur on bisyllabic words except L-L, suggesting obligatory H. However, three-syllable words show only six patterns. Obligatoriness predicts that L-L-L should not be possible, but has nothing to say about the absence of H-L-L. The generalization, therefore, may be that a word cannot end in two L tones. Unfortunately there are no monosyllabic words in the language, which could disambiguate between the two interpretations. In any case, the patterns which do contrast on three syllables again argue for a tonal rather than accentual interpretation.

While obligatory H is much more common, there are also cases of obligatory L: in Hakha Lai (Burma), the three underlying tones are /HL/, /LH/, and /L/ (Section 2). Not only does /H/ not exist, but all of the tonal alternations conspire to preserve input L features (Hyman and VanBik 2004). A number of languages have a tonal contrast on only one (possibly stressed) syllable per word. Thus, Dadibi (Macdonald and Macdonald 1974) and Fasu (May and Loeweke 1964), both spoken in the Southern Highlands Province of Papua New Guinea, require one syllable per word to be /H/ or /L/, remaining syllables being toneless, while Pame (Central, Mexico) (Gibson 1956), has one obligatory /H/, /HL/, or /L/ syllable per word.

It appears that tone is the only phonological feature which can have this obligatory property. (In the case of stress, it is the metrical structure that is obligatory.) Although all phonological systems have at least two contrastive vowel heights (Maddieson 1997: 636), no language requires every word to have at least one high vowel. Similarly, all languages have oral stops, but unless Rotokas is analyzed as /p, t, k, b, d, g/ rather than Firchow and Firchow's (1969) /p, t, k, β, r, g/, no language requires every word to contain at least one stop. Although tone is once again different from segmental phonology, obligatory H or L does not mean that a tone is an accent.

The next restrictive property on tone is CULMINATIVITY: in some languages tone is not obligatory, but is restricted to at most one occurrence per domain. The standard example here is Tokyo Japanese, which, as seen in (67), has been subject to both accentual and tonal analyses (McCawley 1978; Haraguchi 1979; Poser 1984, Pierrehumbert and Beckman 1988, etc.):

(67)

		'pillow' + nom.	'heart' + nom.	'head' + nom.	'fish' + nom.
a.	accentual	ma [↓] kura ga	koko [↓] ro ga	atama [↓] ga	sakana ga
b.	tonal	makura ga H L	kokoro ga H L	atama ga H L	sakana ga
c.	approx. pitch	mákùrà gà	kòkòrò gà	àtāmá gà	sàkànā gā

Tokyo Japanese has been characterized as having at most one abstract accent (*) per word, one HL pitch drop, or one prelinked H tone. Culminative /H/ occurs in Somali (Hyman 1981; Saeed 1999) and Ocaina (Peru) (Agnew and Pike 1957), and is particularly well-attested in Papua New Guinea, for example, in Bahinemo (Dye 1992), Baruya (Lloyd and Lloyd 1992), Tinputz (Hostetler and Hostetler 1975), and Wantoat (Davis 1969). As in the case of obligatoriness, there are variations on the identity of the culminative tone(s): /HL/ (vs. /H, L/) in Obukuitai (Indonesia, Papua) (Jenison and Jenison 1991), /H/ or /HL/ (vs. /Ø/) in Awad Bing (Papua New Guinea) (Cahill 2001), /H/, /HL/, or /LH/ (vs. /L/) in Puinave (Colombia) (Girón Higuaita and Wetzels 2007).

While it was pointed out that only stress and tone can be obligatory, a number of other features have been reported to be culminative, for example, aspiration and glottalization in Cuzco Quechua (Parker 1997: 2), vowel length in Mam (Willard (2004: 7), mid vowels in Punu (Kwenzi Mikala 1980: 8; Fontaney 1980b: 55), nasalized vowels in Karo (Gabas 1999: 42n). Since one does not speak of "mid vowel accent" or "nasal accent," and so on, we can assume that culminativity is simply a restriction that can be placed on tone, as it can be on other features. That this is the right move is seen from Donohue's (1997: 367–368) presentation of the tone patterns in Arigibi Kiwai (Papua New Guinea) in (68).

(68) schema	1σ	2σ	3σ	4σ
/L/	L	L-L	L-L-L	L-L-L-L
/HL/		H-L	H-L-L	H-L-L-L
/LH/		L-H	L-L-H	L-L-L-H
/LHL/			L-H-L	L-H-L-L L-L-H-L

As seen, /L/ is obligatory, but /H/ is culminative. If these are taken to be diagnostics for accent, then the question would be: Which is the accent? On the other hand, /H/ is both obligatory and culminative in languages such as Kinga (Schadeberg 1973), Una (Indonesia, Papua) (Donohue 1997), and Nubi (Uganda) (Gussenhoven 2006). In such cases, since only pitch is involved in marking the obligatory+culminative prominence, it is not clear whether the H tone should be viewed in metrical vs. strictly tonal terms. In Kinga, at least, where the /H/ is assigned by mora, thereby creating syllables with a [HL] vs. [LH] contrast, as in Somali in (3b), this would be an important difference with canonical metrical systems where the syllable is the stress-bearing unit.

The remaining properties in (64) are probably even less reliable indicators of an accentual system, although they have historically played into such analyses. Some of the studies in Clements and Goldsmith (1984), van der Hulst and Smith (1988), and Inkelas and Zec (1990) deal with Bantu systems which have a privative /H/ vs. /Ø/ opposition (cf. Odden 1988, 1999). One such system is Haya, where one might also be misled by the fact that nouns in isolation can bear at most one H or HL tone, for example *o-mu-zi* ‘root’, *o-mú-ti* ‘tree’, *o-mu-lími* ‘farmer’, *o-bu-gólo* ‘snuff’, *o-mu-kâzi* ‘woman’. However, as seen in (69), the word /ó-mu-tí/ ‘tree’ has two underlying H tones and can be realized with any of the eight possible combinations of H and L (< /Ø/) tones (Hyman and Byarushengo 1984: 56):

- (69) a. L-L-L : o-mu-ti gwaa káto ‘Kato’s tree’
 L-L-H : o-mu-tí gwange ‘my tree’
 L-H-L : o-mú-ti ‘tree’
 L-H-H : o-mú-tí káto ‘a tree, Kato!’
- b. H-L-L : a-gul’ ó-mu-ti gwaa káto ‘he buys Kato’s tree’
 H-L-H : a-gul’ ó-mu-tí gwange ‘he buys my tree’
 H-H-L : a-gul’ ó-mú-ti ‘he buys a tree’
 H-H-H : a-gul’ ó-mú-tí káto ‘he buys a tree, Kato!’

The rules involved are as follows: (i) An H tone vowel becomes L after pause, as in (69a); (ii) an H tone is deleted before a genitive noun phrase; (iii) a phrase-final H tone is anticipated onto the penultimate (e.g. before vocative ‘Kato!’); (iii) an H tone becomes L before pause (where it also will be phrase-final). (The apostrophe in the examples in (69b) indicates the elision of the final vowel of a-gul-a ‘he buys’.) Even though the underlying /H/ vs. /Ø/ contrast is privative, the system is clearly

tonal. The same is true of systems where Hs are attracted to metrically strong positions, for example, the penultimate in Shambala (50b) and Giryama (51b).

Finally, bona fide tone systems may also show phrasal reduction phenomena reminiscent of loss of stress when out of focus or stress subordination in compounding. For example, the East Sudanic language Nara (Eritrea) distinguishes at least the tone patterns in (70), taken from Hayward (2000: 255), where the mora (μ) is a vowel or sonorant:

(70) schema	μ	$\mu\mu$	$\mu\mu\mu$	$\mu-\mu$	$\mu\mu-\mu$	$\mu-\mu\mu$	$\mu-\mu-\mu$	$\mu\mu-\mu-\mu(\mu)$
/L/	(L)	LL	(LLL)	(L-L)	LL-L	(L-LL)	L-L-L	LL-L-L
/H/	H	HH	(HHH)	H-H	HH-H	(H-HH)	H-H-H	(HH-H-H)
/LH/		LH	LLH	L-H	LL-H	L-LH, L-HH		LL-L-H
/HL/	HL	HL	HLL	H-HL	HL-L, HH-L	H-LL	H-H-L	
/LHL/			LHL		LH-L	L-HL	L-H-L	LL-L-HL

As seen, the five schemas establish that there is a full tone system, although with the following constraint: “. . . attention is directed towards an obvious asymmetry with regard to the distribution of the two tones, for we do not find an HLH contour as a counterpart to the LHL contour . . .” (Hayward 2000: 256). Recall that /HLH/ was also missing from Kpelle in (18), and that this sequence served as the input to H tone plateauing in Luganda in (40b) and (53b). In fact, as summarized in the table in (71), whenever two words occur which produce an H-L-H sequence in a COMPLEMENT+HEAD construction, the second H becomes L, as indicated by L in the appropriate cells:

(71) $W_1\downarrow/W_2 \rightarrow$	/L/	/H/	/LH/	/HL/	/LHL/
/L/					
/H/			L		L
/LH/			L		L
/HL/		L	L	L	L
/LHL/		L	L	L	L

Thus, H-H # L-H will become H-H # L-L, H-L # H-H will become H-L # L-L, and so forth. Whereas Luganda changes H-L-H to H-H-H, Nara modifies H-L-H to H-L-L. Both are a response to the constraint against *HLH (Cahill 2007) or *TROUGH (Yip 2002: 137). Since the relevant Luganda constructions have the reverse

HEAD+COMPLEMENT structure, the two languages reveal that it is the tones of the head that are modified to avoid a tonal trough. While this corresponds with the observation that the focus position is often post-verbal in a VO language, but pre-verbal in an OV language (Harries-Delisle 1978: 464), we do not need to see the phenomenon as accentual. Rather, prosodic features such as stress, tone, and also vowel length (Kisseberth and Abasheikh 1974) can be sensitive to specific syntactic configurations.

To summarize this section, we have seen that tones can be relatively free or considerably restricted. At one end of the spectrum, tone specifications are “dense,” with a tone feature required for every TBU. At the other end, tone specification is “sparse,” with tone features subject to obligatoriness, culminativity, privativity, or restriction to specific positions. While systems which tend toward the latter characterization have sometimes been termed “accentual,” this is an analytical move which is hard to justify on independent grounds. As Gussenhoven (2004: 42) puts it, “‘Accent’ . . . is an analytical notion and cannot be measured. [It is] thus different from stress, which is typically an observable phenomenon, and different also from tone, whose existence is equally measurable.” Thus, while some languages must be analyzed with stress and others with tone, it is important to underscore that no language *MUST* be analyzed as “pitch accent.” A tonal interpretation is always possible.

The greatest challenge to this position comes from tone systems which have a metrical character. Consider, for example, the case of the Iroquoian language Seneca (Iroquoian) (Chafe 1977, 1996). Seneca and closely related Cayuga have received considerable attention in the metrical literature (see Hayes 1995a: 222–226 and references cited therein), which has generally analyzed Seneca as having iterative iambs constructed left-to-right. The properties of Seneca in (72) summarize the trochaic reanalysis in Melinger (2002):

- (72) a. mark the first syllable extrametrical
 b. build bisyllabic trochees left-to-right
 c. assign an H tone to the first syllable of a trochee iff either syllable is closed

As indicated, Seneca has a metrical system. However, several things set it apart from more usual stress systems. First, the trochees are not used to establish stress, but to assign H tones. Second, the H tones are assigned by open vs. closed syllable. Third, the required closed syllable can be in either position of the trochee. This produces the rather unusual tonal distributions seen in the schemas in (73).

- (73) a. <σ> (CáC.Ca) (CaCa)
 b. <σ> (Cá.CaC) (CaCa)
 c. <σ> (Cá.CaC) (CáC.Ca)
 d. <σ> (Ca.Ca) (Ca.Ca)

In each example, the first syllable has been marked off as extrametrical, and the following four syllables have been grouped into two trochees. In (73a) the head

of the first trochee consists of a closed syllable and thus receives an H tone (vs. the second trochee which consists of two open syllables). (73b) also receives an H on the head of the first trochee, although in this case it is the closed syllable of the non-head which licenses the H. In (73c) both trochees contain a closed syllable, and so an H tone is assigned to the first syllable of both, in violation of culminativity. Finally, since there are no closed syllables in (73d), no H tone is assigned, in violation of obligatoriness. The question here is whether Seneca has a stress system, a tone system, both, or something else. Following Prince (1983), Hayes suggests that “Lake Iroquoian accent is partly metrical, partly tonal” (Hayes (1995a: 225). In other words, Seneca has a stress system, which happens to be realized in terms of non-obligatory, non-culminative H tone (as well as by other means, for example, lengthening of a strong open-syllable penultimate vowel). Seneca thus offers a case of where tone is superimposed on stress (Hyman 1978: 5), something which might be designated as “a metrical tone system” for convenience. With time, the subsequent segmental changes, which have already begun to obscure the transparency of the metrical system, may make the placement of H tones less predictable. In any case, we are far from the phenomena which are called “accent” because of the sparseness of their H tones. The general conclusion of this section is that so-called pitch accent has no independent properties, rather represents a “pick-and-choose” between a number of properties which are normally associated with stress vs. tone systems (Hyman 2001b, c; 2009a). The languages cited in this section are thus better viewed as “restricted tone systems” (Voorhoeve 1973) rather than belonging to a coherent third category.

6 Conclusion

In the preceding sections I have covered some of the conceptual and analytical problems one faces in approaching the study of tone. As we have seen, tone systems are both complex and diverse. The impression one gets is that Hs, Ls, Ms etc., may exhibit as much variation as all of segmental phonology combined: *Tone can do everything that segmental or metrical phonology can do, but the reverse is not true.* Although we rejected the misconceptions about tone in Section 2, the conclusion nonetheless is that tone is both quantitatively and qualitatively different from segmental features and from the other two traditional suprasegmentals (length and stress). Compared with other phonological properties, we can say the following about tone:

(i) *Tone is the most syntagmatic:* as we saw especially in the phrasal examples in Section 4.2, tone shows the greatest tendency either to wander from its original TBU (Yip’s “mobility”) or to have effects with other tones at great distance. Such properties are particularly observed in systems which contrast tonally marked vs. / \emptyset / TBUs. As few if any other features show such postlexical action at a distance, it is not clear what tone’s closest competitor is.

(ii) *Tone is the most paradigmatic*: despite the highly syntagmatic nature of tone, tone can also be highly paradigmatic, especially in languages in which monosyllabic languages contrast multiple tone levels and tonal contours on monosyllabic words (recall Trique (Itunyoso) in (9)). No other single phonetic dimension offers as many potential phonological contrasts along a single dimension as F0 does for tone – up to five tone heights, as was exemplified from Kam (Shidong) in (7). The only close competitor is vowel height. Among the 451 languages in the UPSID database (Maddieson and Precoda 1990) the only language which has the five-way contrast /i, ɪ, e, ε, æ/ in its inventory of short front unrounded oral vowels is Somali. However, it is clear that this is made possible by the fact that Somali has two sets of vowels which harmonize: peripheral /i, e, æ, u, o/ vs. centralized /ɪ, ε, a, ʊ, ö/ (Saeed 1993). This suggests that another phonetic dimension, perhaps [±ATR], is involved. While some tone systems show that the multiple levels naturally group in higher vs. lower registers (Yip 1980), sometimes with accompanying phonation differences (breathiness, creakiness), in many multilevel tone systems, the only differences are in F0. The possibility for multiple paradigmatic contrasts on a single syllable, including tonal contours, is the aspect of tone that makes it the most different from the use of pitch to mark stress and intonation.

(iii) *Tone is the most analytically open-ended*: tone appears to offer a wider range of reasonable analyses and interpretations than other phonological features. In different languages, a two-level tone system can be analyzed as /H, L/, /H, Ø/, /Ø, L/, or /H, L, Ø/, and there are more options if one succumbs to the temptation to treat tone as “accent” (Section 5). Similarly, depending on the language, the M of a three-level system can be a lowered H, a raised L, an independent third tone, or /Ø/. Using the tone features introduced by Yip (1980), M can be [+upper, –raised] or [–upper, +raised]. While this feature difference is quite useful when there are four contrastive tone levels, even a three-level tone system can have two phonetically M tones which contrast in certain contexts, suggesting that they are underlyingly different. While it has been occasionally claimed (e.g. by Goldsmith 1985) that other features can be binary in some languages but function privatively in others, or that one feature value can be marked in one language vs. unmarked in another (Hume 2003, 2007), tone offers especially compelling examples of both. Thus, while /H/ vs. /Ø/ is more common than /L/ vs. /Ø/, both are found in Bantu (Hyman 2001b) and in Athabaskan (Hargus and Rice 2005), and “marked L” is also found in Malinke (Kita) ((Creissels and Grégoire 1993), Galo (NE India) (Post 2007), and the closely related NW Amazonian languages Bora (Weber and Thiesen 2000) and Miraña (Seifart 2005).

(iv) *Tone is the most abstract*: in many cases the relation between input and output tones is a very abstract or indirect one. In the case of two-level tone systems, this is in part due to the considerable possibilities for spreading, shifting, deletion, and insertion, as when the tones of Haya /ó-mu-tí/ ‘tree’ appeared to become inverted to [ò-mú-tì] in (69) above. In addition, underlying systems can be realized with more levels in the output than they started with. Thus, although Ngamambo only has underlying /H, L/, after an H, there is a five-way surface

contrast between H, M, \downarrow M, L and L^o, the last being a level L tone distinct from an L which falls in pitch before pause (Hyman 1986). Floating tones add considerably to the possibilities for abstract underlying representations, and are often easier to motivate than their segmental counterparts.

(v) *Tone is the most arbitrary*: while much of tonology is phonetically grounded in a transparent way, many tonal alternations appear arbitrary from a synchronic point of view. This occurs both in morphology, for example, nominative vs. accusative tone marking in Maasai in (14), as well as at the phrase level, for example, the Wuxi case discussed in Section 4.2. In many cases one tone pattern simply corresponds arbitrarily to another. Perhaps the most famous case of this occurs involves Xiamen tone sandhi rule in (74a) (Chen 1987):

- (74) a. $T \rightarrow T' / \text{ ___ } T$ within a tone group (T = base tone; T' = sandhi tone)
 b. 24, 44 \rightarrow 22 \rightarrow 21 \rightarrow 53 \rightarrow 44
 c. # yi kiong-kiong kio gua ke k'uah puah tiam-tsing ku ts'eq #
 44 24 24 21 53 44 21 21 53 44 53 32 = T
 22 22 22 53 44 22 53 53 44 22 44 = T'
 he by force cause I more read half hour long book
 'he insisted that I read for another half an hour' (Chen 1987: 113)

While the chain shift in (74b) shows that each of the five non-stopped base tones is replaced by another tone, the actual featural changes cannot be motivated synchronically. It hard to think of an equally complex arbitrary input-output relation outside tone.

(vi) *Tone is the most autosegmental*: As was seen in Section 3, tone is THE autosegmental property par excellence. Compared to segmental features, tone is far more likely to float as a lexical or grammatical tone, to show stability effects, to undergo dislocation, or to interact with like features at a distance. Tone sequences are much more likely to be treated as contours which can be manipulated as units or as "melodies" which can be mapped over multiple TBUs. While there are analogues to many of the autosegmental properties of tone, tone provided – and still provides – the model.

To sum up the above observations in one sentence, tone is extraordinarily versatile, a lot of things at once. Over the past several decades there has been tremendous progress both in documenting this versatility and in developing models to deal with it. Autosegmental phonology has been successful largely, because, as Kenstowicz (1994: 312) puts it, "... tone behaves independently from other features and so can be discussed in relative isolation." However, it is when we approach the INTERDEPENDENCY of tone with other features that the picture becomes less clear. Although of great practicality, the use of informal Hs, Ls, Cs, and Vs disguises several interrelated questions which have not been settled: (i) What is or can be a TBU? The syllable? mora? segment within a rime? onset? (ii) Where does tone link up within the feature geometry? Exclusively to the TBU? In addition or instead of the laryngeal node, where Halle and Stevens' (1971 pitch-affecting glottal features [stiff], [slack], [constricted], and [spread] can naturally

interact with tone? (iii) What is the correct set of tonal features? Is a unique set of tonal features even possible, or does it vary from language to language? Possible solutions to such questions are considered in several monographs and dissertations, for example, Duanmu (1990), Bao (1999), Bradshaw (1999), Snider (1999), Yip (2002), and Pearce (2007b). The answers ultimately also bear on the question of how laryngeal gestures are implicated in the process of tonogenesis (Matisoff 1973; Hombert, Ohala, and Ewan 1979; Thurgood 2002). However, despite the large number of proposals, the field is far from consensus. Recently, Clements, Michaud, and Patin (2009), and Hyman (2009b) have argued that the kinds of arguments supporting a universal set of segmental distinctive features are largely absent or ambiguous when applied to tones. While the Chomsky and Halle (1968) feature specifications [+high, –low] were designed to capture both the intersecting class of high vowels, /i, ü, u, u/ and their relation to palatal and velar consonants, the same cannot be said for [+upper, +raised] or [+stiff, –slack]. As opposed to [+high, –low], such features specify only one tone, namely H, not a class of tones, and also fail to capture five contrasting tone heights. Where they do interact with laryngeal features, for example, voicing, aspiration, or glottalization, the diachronic and synchronic correlations are often contradictory (see Kingston 2003, 2005 and references cited therein). As a result, tone is less reliably “gridded in” than other features, thus again suggesting that it is “different.”

Given the independence and extreme range of tonal phenomena, perhaps there will not be a single, definitive answer to one or more of the questions raised in this chapter. In fact, the above questions (i)–(iii) may even contain a misunderstanding that we still have either about tone, or more likely, about phonology in general. Perhaps when these questions are answered, tone will once again contribute in a central way to phonological theory.

ACKNOWLEDGMENTS

I am grateful to various audiences who had heard and responded to different sections of this chapter under titles such as “Is tone different?” and “There is no pitch-accent prototype.” I would especially like to thank Nick Clements, Grev Corbett, Denis Creissels, John Goldsmith, Carlos Gussenhoven, Morris Halle, Will Leben, Doug Pulleyblank, and Donca Steriade for their stimulating reactions, which still keep me thinking.