1. Introduction

This article surveys developments and achievements in research on tone in the past thirty years, primarily in the generative framework, and outstanding issues that pertain to current research. I begin with a discussion of what tone languages are and point out some similarities between tone and intonation.

1.1. Tone languages and non-tone languages

It is not hard to identify some languages that are clearly tone languages. A good example is Standard Chinese (also called Mandarin), which has four tones on full syllables, shown in (1).

(1) Contrastive tones in Standard Chinese (Mandarin)

<table>
<thead>
<tr>
<th>[ma1]</th>
<th>[ma2]</th>
<th>[ma3]</th>
<th>[ma4]</th>
</tr>
</thead>
<tbody>
<tr>
<td>'mother'</td>
<td>'hemp'</td>
<td>'horse'</td>
<td>'to scold'</td>
</tr>
</tbody>
</table>

The numbers 1-4 refer to the four tones: tone 1 is a high, tone 2 is a rise, tone 3 is a low (or low-rise in final position), and tone 4 is a fall. The four words in (1) are phonologically identical in all respects except tone (I forgo the issue of duration for the moment). In other words, tone (or pitch contour) is lexically contrastive in Standard Chinese (for other functions of tone, see section 3.14). Most tone languages are found in Asia and Africa. In addition, tone had been reported in some European and Native American languages.

It is not hard, either, to identify some languages that are non-tone languages. For example, English is typically considered a non-tone language. When the word [kæt] 'cat' is said with different tones (or pitch contours), additional meanings are added, such as surprise, doubt, disbelief, and so on; however, the word still refers to the same animal. Thus, tone is not lexically contrastive in English.

However, difficulties arise when one tries to classify languages as either tonal or non-tonal. A well-known example is Japanese, in which every word has a fixed tone pattern. This can be seen in (2), taken from McCawley (1978, 113), where syllables are separated by a hyphen, and where H is a high tone and L is a low tone.

(2) Pitch-accent in Japanese

| ka-ki-ga | H-L-L | 'oyster' |
| ka-ki-ga | L-H-L | 'fence' |
| ka-ki-ga | L-H-H | 'persimmon' |

The examples show that tone is lexically contrastive in Japanese, as it is in Chinese. On the other hand, as noted by McCawley (1965) (see also Block 1946, Haraguchi 1977), the tone pattern of a Japanese word is predictable if we posit an abstract “accent” position for each word, shown by italics in (2). This analysis accounts for the fact that, for a given word form, there are only as many possible tonal patterns as there are syllables (ignoring “unaccented” words). For example, the form in (2) has three syllables, so it has three possible tone patterns. It should be noted that, unlike stress, accent is not necessarily accompanied by greater duration or amplitude. Apart from its
effect on pitch, accent is hardly felt by native Japanese speakers; I will return to this point below. Since tone (or pitch) is predictable from accent in Japanese, it is often called a “pitch-accent” language.

But English also has pitch contours. In addition, the contours are quite predictable (e.g. Goldsmith 1974 [published in 1981], Liberman 1975, Pierrehumbert 1980). For this reason, English has been called a tone language, too. For example, Goldsmith suggests that in neutral intonation English words have the tone pattern (M)HL, where M is a mid tone and H is linked to the stressed syllable. Some examples are shown in (3).

(3) Tone in English

<table>
<thead>
<tr>
<th>English</th>
<th>Cantonese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-ca-go</td>
<td>M-H-L</td>
</tr>
<tr>
<td>Bos-ton</td>
<td>H-L</td>
</tr>
<tr>
<td>A-me-ri-ca</td>
<td>M-H-L-L</td>
</tr>
<tr>
<td>Ca-na-da</td>
<td>H-L-L</td>
</tr>
</tbody>
</table>

Analyses like (2) and (3) can be extended to African languages (e.g., Goldsmith 1982, Kenstowicz 1987, Sietsema 1989, Kenstowicz and Kisseberth 1990), American Indian languages (e.g. Hinton 1991), and European tone or pitch-accent languages, such as Lithuanian (Halle and Vergnaud 1987) and Serbo-Croatian (Inkelas and Zec 1988). Such analyses raise serious questions for a typology of tone languages, as noted by Hyman (1978) and McCawley (1978). In particular, the distinction between tone and non-tone languages is no longer obvious. In addition, if English and Japanese are tone languages, one wonders what language is not.

### 1.2. Tone and intonation

The pitch contour in English is often called intonation. Indeed, all “non-tone languages” have intonation (I will return to the question of whether tone languages can have intonation). Although both tone and intonation are characterized by pitch (for a discussion on the difference between pitch and the fundamental frequency, see Beckman 1986), one may still wonder whether they are different. For example, are tone and intonation produced by the same articulatory mechanism? Are tone and intonation perceptually the same? Should tone and intonation be represented with the same phonological features? Phonetically, there is no evidence that tone and intonation are made with different articulatory mechanisms (cf. Zemlin 1981). In addition, as far as listeners are concerned, tone (in Chinese) and intonation (in English) are interchangeable. I will cite two examples. First, when English words are borrowed into Cantonese Chinese, a stressed English syllable usually becomes a high toned Cantonese syllable. This is shown in (4), from Chao (1980, 42), where stress in English is indicated by underline.

(4) Tone in borrowing

<table>
<thead>
<tr>
<th>English</th>
<th>Cantonese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oakland</td>
<td>[ok-løn]</td>
</tr>
<tr>
<td>Pacific</td>
<td>[pʰa-si-wik]</td>
</tr>
</tbody>
</table>

Obviously, Cantonese speakers hear the high pitch on a stressed English syllable as a high tone. Second, Standard Chinese has a rule that changes tone 3 (low) to tone 2 (rise) before tone 3, or 3 --> 2/3. Cheng (1968) noted that in code switching, tone 3 changes to tone 2 before an unstressed English syllable as well, but not before a stressed one. This is shown in (5), where Chinese is transcribed in Pinyin.

(5) Tone in code switching between Chinese and English

a. xiao2 professor (*xiao3 professor)
   'small professor'
b. xiao3 lecturer (*xiao2 lecturer)
'small lecturer'
The word xiao3 'small' is underlyingly tone 3. It changes to tone 2 in (5a) but not in (5b). The reason is that in (5a) the following syllable pro- is unstressed and has a low tone (although Goldsmith marks it as M, see section 3.1); this is perceived by Standard Chinese speakers as tone 3 (low) and as a result the rule 3→2 is triggered. In (5b), the following syllable lec- is stressed and has a high tone, which does not trigger tone change. Finally, let us see whether there are reasons to posit different phonological features for tone and intonation. Traditionally, tone refers to the pitch contour of a syllable or a word, and intonation refers to the pitch contour of a phrase or an utterance. But as Goldsmith (1981) pointed out, intonation in English consists of a sequence of word tones, similar to what one finds in a “tone” language. In addition, tonal behavior in English is similar to that in other tone languages. For example, when the word tone (M)HL falls on 'Japan', we get M-HL, where HL on the second syllable is realized as a fall; this is what one finds in tone languages (see below). We have seen, then, that tone and intonation display certain similarities. This point will recur in the rest of this article.

2. Developments in the past 30 years
In early works of generative phonology (1950s and 1960s), there was very little mention of tone. This is probably because a main issue then was distinctive features, and tone did not seem to fit in easily. In particular, distinctive features work well for the representation of segments, but tone has traditionally been thought to be a “suprasegmental” property, a property that belongs to units larger than the segment (e.g. Pike 1948, Firth 1957). Another difficulty is that some tones seem to require “trajectory” features, such as “rise” or “fall”, whereas regular segmental features are all “static”, such as [round] or [voice], but not [increasing round] or [diminishing voice]. The lack of discussion on tone or intonation left a striking gap in the monumental work of SPE (see Chomsky and Halle 1968, 329).

Wang (1967) made the first serious attempt to represent tone with distinctive features. He posited seven binary features to represent all tones. However, out of the seven features, four were trajectory features: [contour], [rise], [fall], and [convex]. In addition, with seven features many possible tones are predicted, of which only 13 were actually found. Perhaps because of this, Wang's work was not quickly adopted by other generative phonologists.

A year after SPE appeared, Woo (1969) proposed a theory of tone features that had great influence on later works. Woo’s theory has two parts. First, all contour tones are combinations of level tones. Second, tone bearing units (TBUs) are segments in the rime and each TBU can carry just one tone. Woo’s theory is illustrated in (6). Following Chao (1930) and Wang (1967), Woo posited five levels, but for illustration I will focus on two, H and L.

(6) Woo’s (1969) theory of tone features

<table>
<thead>
<tr>
<th>Traditional term</th>
<th>Tone features</th>
<th>TBUs needed</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>H</td>
<td>1</td>
</tr>
<tr>
<td>Low</td>
<td>L</td>
<td>1</td>
</tr>
<tr>
<td>Rise</td>
<td>LH</td>
<td>2</td>
</tr>
<tr>
<td>Fall</td>
<td>HL</td>
<td>2</td>
</tr>
<tr>
<td>Fall-rise</td>
<td>HLH</td>
<td>3</td>
</tr>
<tr>
<td>Rise-fall</td>
<td>LHL</td>
<td>3</td>
</tr>
</tbody>
</table>

In Woo’s analysis, a fall is HL, a rise is LH, and a fall-rise is HLH. This is similar to treating the diphthong [ai] as a combination of [a] and [i] and the triphthong [auə] as a combination of [a], [u], and [ə]. In addition, a short/light (monomoraic) syllable has one rime segment, so it can only carry
a level tone. A simple contour tone (rise or fall) consists of two level tones and must fall on a long/ heavy (bimoraic) syllable, which has two rime segments. A complex contour tone (fall-rise or rise-fall) consists of three level tones and must fall on a syllable that has three rime segments (trimoraic). For example, in Standard Chinese, all full syllables are long. In particular, tones 1, 2, 4, and the low version of tone 3 have two rime segments and are HH, LH, LL, and HL respectively; thus, [ma] in (1) should be [maa]. In addition, the low-rise version of tone 3 has three rime segments and is LLH.

Woo's work was significant in two ways. First, it proposed that tone could be represented with "static" features after all. Second, Woo's work was based on Chinese, where most contour tones are found. If Chinese tones do not require trajectory features, tones in other languages are unlikely to either. The idea that contour tones are made of level tones was essentially adopted by subsequent generative phonologists. However, the idea that there is a direct relation between the length of a syllable and the number of tones it can carry was often ignored, although it is, in my view, essentially correct. I will return to it below.

The next development came with the works of Leben and Williams. Leben (1971, 1973) suggested that tones should be represented on an independent tier, parallel to the segment tier. Williams (1971 [published in 1976]) further suggested that the two tiers are coordinated by "tone mapping rules", which later became known as "association lines" and "association conventions".

This line of research culminated in the influential work of Goldsmith (1976), through which the term “autosegmental phonology” became known (other equivalent terms are three-dimensional phonology, nonlinear phonology, and multi-tiered phonology). In the following decade, a large number of autosegmental studies ensued. Among works on tone, most dealt with African languages (e.g. Clements 1978, Clements and Ford 1979, Odden 1981, Hyman 1981, Pulleyblank 1983, Clements and Goldsmith 1984, Cassimjee 1987, Myers 1987, and many others). Main issues of interest included the association conventions, the representation of multiple tone levels, default tones, floating tones, downdrift and downstep, prelinked tones, accent marking, the OCP (Obligatory Contour Principle) effect, and the interaction between tone and morphosyntactic structure. Downdrift (a H after a L is lower in pitch than the preceding H) and downstep (a H is lower in pitch than a H immediately preceding it) were later found to be the same phenomenon after it was realized that a downstepped H always had a floating L right in front of it. Generative works have also been done on European tone (or pitch-accent) languages, such as Swedish, Norwegian, Serbo-Croatian, and Lithuanian (e.g. Bruce and Garding 1978, Garding 1979, Freltheim 1981, Withgott and Halvorsen 1988, Inkelas and Zec 1988, Halle and Vergnaud 1987). There were fewer generative studies on Asian tones in the 1970s and 1980s, with most of these focusing on Chinese languages. Main works included those of Yip (1980), Wright (1983), Chan (1985), Shih (1986), and Chen (1987). Primary issues of interest were the representation of contour tones, the representation of tone sandhi, and domains of tone sandhi. The Chinese dialects that received the most attention were Standard Chinese, in particular its tone 3 sandhi rule, Tianjin, Southern Min dialects, and Wu dialects.

Besides “real” tone languages, tone-related work has also been done on other languages, such as English, Japanese, Dutch, and German (Liberman 1975, Bing 1979, Ladd 1980, Pierrehumbert 1980, Poser 1984, Beckman and Pierrehumbert 1986, Gussenhoven 1988, Féry 1989, among others). In particular, Liberman (1975) has argued that English intonation can be analyzed in terms of a stress system and a tone system, where the tone system consists of a single sequence of Hs and Ls. This approach, now a standard in generative analyses of “intonation” languages, suggests that tone and intonation are essentially the same phenomenon.

Parallel to the generative study on tone, four other developments should be mentioned. First, there was an effort to find universals of tone through cross-linguistic survey (e.g. Fromkin
1972, Hyman and Schuh 1974, and Maddieson 1978). Second, there were works on tonogenesis in Asian languages, i.e., the emergence of tone as the result of the loss of consonant features (e.g. Haudricourt 1954, Maran 1973, and Matisoff 1973). Third, some phonetic models on tone were proposed (e.g. Halle and Stevens 1971, Hombert 1978). Finally, there is a continuous flow of descriptive works by linguists in China, mostly written in Chinese. Of special interest are experimental studies on Standard Chinese (e.g. Lin, Yan and Sun 1984, Lin 1985, Yan and Lin 1988, He and Jin 1992, Wang and Wang 1993), a series of works on Tibetan (e.g. Hu 1980, Hu, Qu and Lin 1982, Qu 1981), and a large number of studies on the Wu dialects, especially the work of Xu, Tang, You, Qian, Shi and Shen (1988) on Shanghai.

3. Current issues relating to tone

Despite the tremendous progress in tone research in the past three decades, many issues remain quite open. Consequently, in discussing current issues, I will often speculate on what I think is likely to be the right solution.

3.1. Levels of tone

There have been many discussions on how many tone levels are needed to describe all languages (e.g. Chao 1930, Wang 1967, Halle and Stevens 1971, Anderson 1978, Yip 1980, Clements 1983, Huang 1985, Hyman 1986, Hulst and Snider 1993), but there is still no agreement. Phonetically, pitch is the primary perceptual correlate of tone, and in real speech there can be many pitch levels. But if we focus on phonemic levels, i.e., those that are distinctive, the number of levels is quite small. In particular, if downstep or downdrift belongs to phonetic implementation, which need not be represented with separate levels, far fewer phonemic levels are needed. In African languages, two phonemic levels, H and L, are often sufficient, although many more have been posited (I return to “depressed” tones later). The same is true for English. For example, Goldsmith (1981) posited (M)HL as the neutral English word tone, where syllables before stress have a higher pitch level than syllables after stress. But if we assume downstep (or what Beckman and Pierrehumbert 1986 call “catathesis”), by which both H and L descend gradually, there is no need to posit M, and the English word tone can simply be (L)HL. For this reason, Liberman (1975) and Pierrehumbert (1980) only assumed two levels for English, H and L.

In Asian languages, three or four contrastive levels are quite common; five contrastive levels have also been reported (Chang 1953, Shi, Shi and Liao 1987). The widely used system of Chao (1930), adopted by the International Phonetic Association (1989), posit five levels. Anderson (1978) also posits five levels, primarily because of Chang's (1953) report that in certain Miao-Yao languages five levels are found on isolated monosyllables.

If tones are to be represented by distinctive features, one needs to convert multiple levels into binary features. It is possible to translate any multi-level system into a binary system; the challenge is to find justification for the translation. For example, English vowels have three heights, for which SPE uses two binary features, [high] and [low]. The analysis is justified by the fact that there are rules that refer to high and mid vowels, captured by [-low], and rules that refer to mid and low vowels, captured by [-high]. However, many Asian tone language have few tone sandhi rules, especially those that involve tone split and tone spreading, which might provide insight into the composition of the tones.

An important tone model was suggested by Yip (1980, 1993), in which four levels are posited. The four levels are divided into two “Registers”, [+upper] and [-upper] (not to be confused with a more loosely used notion of “register”, which simply means a tone level). Each Register is divided into two levels, [+raised] and [-raised], which I will write as [+H] and [-H] and...
call them the “Pitch” feature (in contrast to “pitch”, which refers to phonetic pitch height). Thus, the four tone levels are represented by two binary features, as shown in (7).

Yip’s (1980) model of tone features

<table>
<thead>
<tr>
<th>Register</th>
<th>Pitch</th>
</tr>
</thead>
<tbody>
<tr>
<td>+upper</td>
<td>+H</td>
</tr>
<tr>
<td>-H</td>
<td></td>
</tr>
<tr>
<td>-upper</td>
<td>+H</td>
</tr>
<tr>
<td>-H</td>
<td></td>
</tr>
</tbody>
</table>

The strongest argument for Yip's model, in my view, is that Register is independently related to the voicing of the onset consonant. In particular, in those Chinese languages that have retained onset voicing in obstruents, [+upper] Register invariably goes with a voiceless onset and [-upper] Register invariably goes with a voiced onset. The same effect is found in many African languages, where a voiced obstruent consonant lowers the pitch of a neighboring tone and makes it [-upper] Register. I will return to this point below.

An apparent shortcoming of Yip's model is that it cannot represent more than four levels. But reports of five (or more) levels are few and sketchy. There are, in addition, ways to accommodate five levels in Yip's system; for example, if one level is toneless, then the rest can be represented by the features in (7). If, however, more levels are needed, one can perhaps posit three tones in each Register, giving a total of six levels. Clear data that bear on this issue are still wanting.

Multiple tone levels also raise a question in perception. Since speakers differ both in pitch height and in pitch range, tone levels must be interpreted in relative terms. Now given (7), when one hears a higher tone followed by a lower tone, there are six possibilities: III-IV, II-III, II-IV, I-II, I-III, and I-IV. How does the listener decide which is correct? This question is addressed in the next section.

3.2. Pitch and Register

For all tones, the primary phonetic correlate is pitch. However, tones in [-upper] Register are usually accompanied by a special voice quality, sometimes referred to as breathy or murmured. This is true not only of Chinese languages (e.g. Cao and Maddieson 1992), but also of African languages, such as Zulu (e.g. Laughren 1984 and references therein). Although the role of voice quality is often ignored in tonal literature, it is possible that voice quality is the main perceptual cue for distinguishing [+upper] and [-upper] Registers. If this is correct, the task for tonal perception will be much simpler than previously thought: a binary distinction for Register based on voice quality, and a binary distinction for Pitch based on pitch level.

3.3. Articulation of tone

There have been many studies on the articulatory characteristics of tone, but the precise relation between the articulation of tone and tone features remains unclear. Zemlin (1981) surveyed previous works on articulatory pitch control and identified two major mechanisms: the cricothyroid muscles and the vocalis muscles. The former controls the elongation and the thickness of the vocal cords; the latter controls the “isometric tension” of the vocal cords. Duanmu (1990) speculated that these two mechanisms correspond to the two features of tone: the cricothyroid activity relates to Pitch, and the vocalis activity relates to Register, with a secondary effect on pitch (for a slightly different proposal, see Bao 1990). In this view, [+upper] is related to stiffer vocalis muscles, and [-upper] is related to slacker vocalis muscles. It is worth noting that [+upper] is related to a voiceless consonant and [-upper] to a voiced one, and that in the model of
Halle and Stevens (1971) voiceless consonants have [+stiff vocal cords] and voiced consonants have [+slack vocal cords]. In other words, the effect of consonant voicing on Register can be seen as the spreading of the feature [stiff/slack vocal cords] from the consonant to the vowel (independent of the Pitch feature the vowel already has). Such proposals make specific claims about the connection between phonological features and phonetics and remain to be verified by experimental work.

3.4. Contour tones

Woo's (1969) proposal that contour tones are clusters of level tones has been widely adopted by African phonologists, but it has met considerable skepticism from Chinese phonologists (e.g. Yip 1989, Bao 1990, Chan 1991). Contour tones pose two problems for distinctive feature theory. First, if contour tones are basic units, they require trajectory features such as [rise] and [fall], or a modified version of it, as shown in (8), after Yip (1989), where TBU is the tone bearing unit.

(8) Yip’s (1989) model of contour tone units
\[\begin{array}{c}
\text{rise} \\
\text{TBU} \\
L \\
\text{H} \\
\text{fall} \\
\text{TBU} \\
H \\
\text{L}
\end{array}\]

Duanmu (1994) argued that there is no compelling evidence for contour tone units. Instead, all contour tones are clusters of level tones. But even if this is true, there remains another problem. Since Leben (1971) and Williams (1976), it has been commonly assumed that two (or more) tones can be linked to a single short vowel, which creates a “short contour tone”. Now if a short rise has the feature [-H, +H], then the vowel carrying it has two values of the same feature [H]. A segment that contains two (or more) values of the same feature is a “contour segment”. If phonological theory admits contour segments, many possible segments are predicted, most of which are not found. Anderson (1976) noted this problem and suggested that perhaps contour features can be allowed for suprasegmental features but not for segmental features. However, the distinction between segmental and suprasegmental features is unclear. In addition, as will be discussed below, there are reasons to consider tone a segmental feature, too.

Duanmu (1994) argued that in Chinese languages a contour tone always falls on a long syllable, as suggested by Woo (1969). Each rime segment, therefore, can take just one tone, either H or L (ignoring Register). Duanmu (1994) also argued that, based on available evidence, “short contour tones” in African languages always occur on lengthened syllables. For example, the widely cited short contour tones in Mende (Leben 1973, Goldsmith 1976, and subsequent works) occur on lengthened vowels (Aginsky 1935, Ward 1944). The same is true in English. For example, in neutral intonation, 'black' is HL (fall) and 'blackbird' is H-L (H on the first syllable and L on the second). It appears that the former has a short contour tone. But as Goldsmith (1981, 299) pointed out, and as is well known phonetically, [æ] is lengthened in 'black' but not in 'blackbird'. In other words, real cases of short contour tones (contour tones on vowels that are not lengthened) are yet to be seen. This is not surprising: in a perception experiment by Greenberg and Zee (1979), it was found that a rising pitch contour will be heard as a level tone if the vowel is short (80 ms or less). There is, therefore, no compelling reason to posit contour tone features in distinctive feature theory (for arguments against other contour features, see references cited in Duanmu 1994).

3.5. The tone bearing unit (TBU)

Four kinds of entities have been proposed as the tone bearing unit (TBU). They are described in (9) and illustrated with the syllables [man] and [ma] in (10), where {} delimits a TBU.
(9) Proposals of TBUs
   a. The entire syllable (or the voiced part of it)
   b. The rime portion of the syllable (but not the onset portion)
   c. The mora (including the onset)
   d. The moraic segment (the segment in the rime)

(10) Illustration of different TBU proposals

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>man</td>
<td>ma</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>{man}</td>
<td>{ma}</td>
</tr>
<tr>
<td>{an}</td>
<td>{a}</td>
</tr>
<tr>
<td>{ma}, {n}</td>
<td>{ma}</td>
</tr>
<tr>
<td>{a}, {n}</td>
<td>{a}</td>
</tr>
</tbody>
</table>

(9a) was assumed by Chao (1930), Firth and Rogers (1937), and Wang (1967), among others. (9b) was proposed by Howie (1974). (9c) was proposed by Hyman (1985). (9d) was proposed by Woo (1969) and is consistent with the moraic theory of Hayes (1989), where the onset consonant is not dominated by a mora. There is no evidence that the onset can be a TBU by itself. Rather, the question is whether it can carry part of the tone when it is voiced, such as [m] in [man]. (9a) and (9c) assume that it does. (9b) and (9d) assume that it does not. The first assumption is often made but never argued for, as far as I am aware. In contrast, Howie (1974) has argued for the second assumption and shown that the pitch contour of a voiced onset (sonorant consonants or glides) is often irregular, and that the expected pitch contour does not start until the rime portion. If Howie is correct, the remaining choice is between (9b) and (9d).

According to (9b), both [man] and [ma] have one TBU. According to (9d), [man] has two TBUs and [ma] has one. There is no question that (9d) is true in some languages, such as Luganda and Japanese. (9b) is often thought to be true in Chinese languages (e.g. Chao 1930, Firth and Rogers 1937, Wang 1967, Yip 1989, Chan 1991). It may appear that the choice between (9b) and (9d) is a matter of language variation. But as Duanmu (1994) argued, the TBU in Chinese is also (9d), as suggested by Woo (1969) (see also Shih 1997, who argued that the TBU in Standard Chinese is not the syllable but the moraic unit). If this is correct, there may be no cross-language variation in the TBU. Instead, the TBU is a moraic segment in all languages.

3.6. Tonogenesis and consonant-tone interaction

Tonogenesis refers to the emergence of tone in a previously non-tone language based on (earlier) consonantal contrasts. The term is first introduced by Matisoff (1970), although the phenomenon has been studied in earlier works. Tonogenesis is commonly found in Asian languages. A well-known example is Vietnamese (Haudricourt 1954, Matisoff 1973), in which the loss of coda consonants created an initial set of three tonal contrasts, and a subsequent loss of onset voicing split the contrasts into six. In other languages, such as Lhasa Tibetan (Hu 1980, Duanmu 1992), the loss of onset voicing created a two-way tonal contrast. Some researchers believe that Chinese also evolved from a non-tone language through a similar process (e.g. Baxter 1992, 7). Tonogenesis in Native American languages is somewhat controversial (see de Jong and McDonough 1993 for a recent discussion). I am not aware of studies on tonogenesis in African languages (but see depressor consonants below).

The effect of coda consonants on tone is not well understood (see Maran 1973 for some discussion). Some linguists believe that coda consonants gave rise to four tones in Chinese (cf. Baxter 1992 and references therein). A final fricative (e.g. [s] or [h]) gave rise to a falling tone. A final glottal stop gave rise to a rising tone. A syllable with no obstruent coda gave rise to a level tone. A syllable that ended in [p], [t], or [k] gave rise to a fourth category, which is characterized
by its shortness. However, since all cases of tonogenesis from coda consonants happened long ago, and since there is no language that is undergoing or has recently undergone such a process, its exact nature is yet to be determined.

The effect of onset voicing on tone is better understood and can be generalized as “voiceless-high” and “voiced-low”: a voiceless obstruent consonant leads to a higher tone and a voiced obstruent consonant leads to a lower tone. The featural link the between voicing and tone is briefly discussed in section 3.3 and will be further illustrated below. Sonorant onsets may behave either like voiceless obstruents or like voiced obstruents. In Underspecification Theory (cf. Steriade 1987, Archangeli 1988), this is expected, since sonorants can be underlyingly unspecified for voice. (11a) shows an example from Lhasa Tibetan, which acquired tone relatively recently (Hu 1980); the voiceless [k] is aspirated, which is not transcribed here. (11b) shows an example from Shanghai Chinese, which still retains onset voicing in addition to different tones (Xu, Tang, You, Qian, Shi and Shen 1988).

(11) Tonogenesis in Tibetan and Shanghai
a. Historical Tibetan Lhasa Tibetan
   [ko] --> [ko] high
      'he'
   [go] --> [ko] low
      'hear'

b. Shanghai Chinese
   [se] high rise
      'umbrella'
   [ze] low rise
      'wealth'

In (11a), a historically voiceless onset gave rise to a syllable with a high tone, and a historically voiced onset gave rise to a syllable with a low tone. In Shanghai Chinese, there are two rising tones on “smooth” syllables (those without a glottalized rime). As seen in (11b), the high rise occurs with a voiceless onset and the low rise occurs with a voiced onset.

Apparent exceptions to voiceless-high and voiced-low have been observed (e.g. Kingston and Solnit 1988). However, such exceptions often result from subsequent tone change. For example, the historical Chinese tone “Ping” has split into Yin Ping and Yang Ping in Northern dialects (also called Mandarin dialects), all of which have lost onset voicing. Yin Ping occurs with historically voiceless onsets. In Beijing Mandarin, Yin Ping is a high level tone, but in Tianjin Mandarin, Yin Ping is a low level tone. It is reasonable to assume that Yin Ping in Beijing Mandarin is closer to the value right after Ping split than Yin Ping in Tianjin Mandarin. In other words, at the time of tone split, voiceless-high and voiced-low seems to be a valid correlation.

The effect of onset voicing on tone is not purely a split between a higher pitch set and a lower pitch set. As far as facts are clear, the split is always accompanied by a difference in voice quality: clear in the upper set and murmured in the lower set (cf. sections 3.2 and 3.3). For this reason, voiceless-high and voiced-low can also be called Register-Genesis or Register-split.

The split of an existing tone by onset voicing can be viewed in various ways. A possible analysis of (11b) is shown in (12) (see Hombert, Ohala and Ewan 1979 for a different proposal and Maddieson 1984 for criticisms of Hombert, Ohala and Ewan 1979).

(12) Featural analysis of Register-split in Shanghai
a. \[ \text{LH} \quad \text{LH} \]
   \[ \text{see} \quad \text{--->} \quad \text{see} \]
A monosyllable is long in isolation (Duanmu 1994). Before split, the vowel carries LH (a rise). Following the theory of underspecification, let us assume that voicelessness and [+upper] are unspecified, as in (12a). In (12b), the voiced onset has the feature [+slack (vocal cords)] (after Halle and Stevens 1971). Next [+slack] spreads to the vowel, causing a split in tone: (12a) has [+upper] by default, and (12b) has [-upper] = [+slack], spread from the onset. The latter is characterized by a murmured voice quality, and accompanied by a lowered pitch level owing to the slackness of the vocal cords. If [+slack] is further delinked from the onset, which often happens in Shanghai, (12b) will become (12c), where the Register on the vowel solely carries the contrast that used to be carried by the voicing of the onset.

The Tibetan case can be analyzed in a similar way. One possibility is shown in (13). As in Shanghai, monosyllables are likely to be long in Tibetan, but for illustration I have ignored vowel length. In addition, on monosyllables H is usually realized as HL and L is usually realized as LH, which need not concern us here (see Duanmu 1992).

(13) Featural analysis of tonogenesis in Tibetan

a. 
\[ \text{H} \]
\[ \text{ko} \] --- > \[ \text{ko} \]

b. 
\[ \text{L} \]
\[ \text{go} \] --- > \[ \text{ko} \] --- > \[ \text{ko} \]

[+slack] [+slack] [+slack]

The default Register is [+upper], which is found in (13a). The spread of [+slack] creates [-upper] Register in (13b). The Register difference between the two words further induces a Pitch difference: [+upper] induces H and [-upper] induces L. The fact that certain features tend to co-occur is not unusual and has been attributed to “enhancement” (Stevens and Keyser 1989). For example, voiceless consonants tend to be aspirated and voiced consonants tend to be unaspirated, and back vowels tend to be rounded and front vowels tend to be unrounded. In fact, as in English, [k] is aspirated in Tibetan (Hu 1980), apparently because aspiration and voicelessness enhance each other.

It has been reported that in some Mon-Khmer languages, the loss of onset voicing has lead to a split in the voice quality of the vowel, without an obvious split in pitch levels (cf. Jenner et al 1976). Voiceless onsets has lead to clear voice (the so-called “first register”) and voiced onsets have lead to murmured voice (the so-called “second register”). Such cases can be analyzed in (14), using [ko] and [go] as hypothetical examples.
Featural analysis of Register-split in Mon-Khmer

a. ko

b. \( \sigma \rightarrow ko \)

\[ [+slack] [+slack] \]

Again, the default Register is [+upper], as in (14a). The spread of [+slack] creates [-upper] Register in (14b). The difference between Tibetan and Mon-Khmer is that in Tibetan the Register-split has further induced a Pitch difference, whereas in Mon-Khmer it has not.

The effect of consonants on tone has also been observed in African languages, such as Zulu, Xhosa, Ewe, Nupe, Ngizim, Mijikenda, and Nguni (e.g. Hyman and Schuh 1974, Laughren 1984, Cassimjee and Kisseberth 1992). For example, a voiced obstruent, known as a “depressor” consonant, is found to lower the pitch level of a neighboring tone. The lowered tone can be seen to be in the [-upper] Register and analyzed in the same way as in (12), where [+slack] spreads from the consonant to the vowel.

3.7. Tone and stress

Since McCawley's (1965) pitch-accent analysis of Japanese, many other languages have been analyzed in the same way. However, since “accent” lacks phonetic correlates, it cannot be checked directly with the native speaker, and as Pulleyblank (1983) pointed out, sometimes accent is used merely as an idiosyncratic diacritic marker. Pulleyblank further pointed out that in many cases, accent is simply a placeholder for a lexical tone, and if we assume lexically linked tones, there is no need to assume accent. For example, the accent locations in Japanese can simply be seen as locations to which H is linked in the lexicon (Poser 1984). Similarly, Blevins (1993) argues that Lithuanian tone can be analyzed by assuming lexically linked H instead of accent.

Pulleyblank's proposal raises the question of whether stress (or accent) and tone can both occur in the same language. This issue has generated considerable discussion. There is little doubt that in languages like English, both stress and tonal patterns are determinable (at least by trained linguists). But perhaps English is not a “real” tone language. Indeed, I am not aware of any “real” tone language in which native speakers or trained linguists have a clear judgment for phonetic stress. For example, in Chinese languages, native speakers do not agree on stress locations on full syllables, nor do Chinese linguists agree (cf. Chao 1968, 38 for Mandarin and Selkirk and Shen 1990, 315 for Shanghai). Similarly, in African languages, Luganda has been at the center of the debate, yet there is no agreement on whether it has accent. Some linguists consider it to have tone only, some consider it to have accent (from which tone is predicted), and some consider it to require both tone and accent (Hyman and Katamba 1993, and references therein).

There is, in fact, a good reason why stress is hard to feel in tone languages. The perceptual study of Fry (1958) shows that pitch is the most important cue for stress. Since pitch variation is not freely available in tone languages, it is natural that their speakers find it hard to tell stress.

In metrical phonology, both stress and accent refer to the head of the prosodic unit “foot” (I ignore higher levels of stress). It is likely that prosodic structures are present in all languages (e.g. Selkirk 1981, McCarthy and Prince 1986). For example, Poser (1990) shows that despite the lack of phonetic stress, Japanese has foot structure. Similarly, Shih (1986) and Yip (1992, 1994) have argued that, despite the lack of agreement on stress, Chinese languages also have foot structure. Nevertheless, as Pulleyblank pointed out, it is rarely shown whether the purported accent in tone languages is motivated on independent metrical grounds. In addition, if accent can be justified independently, one would like to know how it interacts with tone.

In Chinese languages stress is determinable on independent metrical grounds (Ao 1992,
In addition, there is a specific relation between tone and stress, which I will call the Tone-Stress Principle and state in (15).

1993, Duanmu 1993, 1995). In addition, there is a specific relation between tone and stress, which I will call the Tone-Stress Principle and state in (15).

(15) **Tone-Stress Principle**

A stressed syllable is accompanied by an underlying tone pattern.

An unstressed syllable is not accompanied by an underlying tone pattern.

A stressed syllable need not carry the entire tone pattern by itself. For example, in Goldsmith's analysis of English, a stressed syllable in neutral intonation is accompanied by (M)HL. When it is the only syllable, it carries HL, and when it is surrounded by unstressed syllables, it carries just H. According to (15), when a syllable loses stress, it will lose its underlying tone pattern. This agrees with the fact that in fast speech, where more syllables are destressed, fewer tone domains occur.

I will illustrate (15) with some data from Shanghai Chinese, which has two tone patterns on monosyllables, LH and HL (ignoring Register). Polysyllabic domains also have two patterns, [L-H-L...L] if the initial syllable is underlyingly LH, and [H-L...L] if the initial syllable is underlyingly HL (I ignore a more restricted third pattern). Thus, [L-H-L...L] is an expansion of LH over the first two syllables, with additional syllables receiving a default L. Similarly, [H-L...L] is an expansion of HL. In normal careful speech, polysyllabic foreign words form disyllabic tone domains (with a trisyllabic final domain if there is an odd final syllable), indicated by parentheses in (16), where underlying tones are shown above surface tones (syllables are separated by a hyphen; [z] is syllabic in [sz]). The initial syllable of each domain is generally longer than the noninitial syllable(s), which is not indicated here (see Zhu 1995, Duanmu 1999).

(16) **Tone domain formation in Shanghai Chinese (Duanmu 1995)**

LH  LH   HL LH LH LH  HL LH LH LH LH  
L  H  )H L )L H )H L )L H  
dge'-kha' sz- lu- va'-kha' ka- li- fo'-ü- ya
'Czechoslovakia' 'California'

The underlying tone pattern of a syllable in a foreign name is that of the character used to represent the syllable. In hyper-articulated speech, where every syllable is stressed, every syllable can surface with its underlying tone pattern. (16) shows binary foot formation. In addition, it suggests that stress is initial in each foot. This is confirmed by a comparison between [1 2] and [2 1] compounds (numbers indicate the number of syllables in a word). In normal careful speech, a [1 2] compound forms one domain, as in (17a), but a [2 1] compound forms two, as in (17b).

(17) **Tone domains in [1 2] vs. [2 1] compounds in Shanghai**

a. ( sã fe-ga) b. (fe-ga)(tʰa)
raw tomato  tomato soup 
r'aw tomato'  'tomato soup'

The difference between [1 2] and [2 1] is predicted if stress is left-headed, so that in [1 2] there is a stress clash, namely, two stresses occurring on adjacent syllables, as shown in (18a). In metrical phonology, stress clash should be avoided; this can be achieved by deleting stress from the second word, which gives rise to a trisyllabic foot. In contrast, there is no stress clash in [2 1], as shown in (18b), therefore both feet remain.

(18) **Left-headed metrical structures in [1 2] vs. [2 1]**

a. x x  b. x x  
(x) (x x) (x x) (x x)
Stress clash  No stress clash

In many other Asian tone languages, such as Standard Chinese, Cantonese Chinese, and Thai, consecutive full syllables do not form disyllabic domains. According to Duanmu (1993), the reason is that in those languages every full syllable is heavy and forms a bimoraic foot, so that it is
able to retain its underlying tones. In contrast, in languages like Shanghai all syllables are underlyingly light (which can become heavy under stress, as happens when a syllable occurs in isolation), therefore a foot usually contains two (or more) syllables. In other words, the Tone-Stress Principle applies to both kinds of Asian tone languages.

In English (e.g. Pierrehumbert 1980), Norwegian (e.g. Withgott and Halvorsen 1988), and Dutch (e.g. Gussenhoven 1988), a tone pattern accompanies each stressed syllable. This is similar to the Tone-Stress Principle. It remains to be seen whether the Tone-Stress Principle applies to African tone languages, once accent is motivated on independent metrical evidence (for some discussion, see Kenstowicz 1987, Hyman 1987, 1989, Sietsema 1989, Kisseberth and Cassimjee 1992, Hyman and Katamba 1993).

The Tone-Stress Principle is not the only source of tone assignment. An obvious exception is boundary tones, which occur at the edges of certain phonological units (Liberman 1975, Pierrehumbert 1980, Beckman and Pierrehumbert 1986). Another exception is the “associative tone” in some African languages, which is a floating tone that occurs between two nominals in a genitive relation (e.g. Odden 1980, Williamson 1986). It is unclear whether the associative tone is a boundary tone, and I will leave the issue open.

3.8. Typology of tone languages

I mentioned in section 1 that it is difficult to establish a typology of tone and non-tone languages, and that in some sense all languages are tonal. Beach (1924, 84) made a similar remark. However, there are obvious differences among languages like English, Japanese, and Chinese, which ought to be captured.

Pike (1948, 3) offered a narrower definition of tone languages, according to which every syllable in a tone language can carry a contrastive tone. By this definition, Japanese is not a tone language, because only the accented syllable carries a contrastive tone. Similarly, Swedish and Norwegian are not tone languages, because only the stressed syllable carries a contrastive tone (Pike 1948, 14). However, by the same definition, Chinese languages will not be tone languages, because, as discussed in section 3.7, only stressed Chinese syllables are accompanied by a contrastive tone pattern. Consequently, differences among languages like English, Japanese, and Chinese are again left unaccounted for. In fact, if the Tone-Stress Principle is correct, no language would qualify as a tone language by Pike’s definition.

Goldsmith (1981) suggested that all languages have tone, but there is a distinction between “tone” and “accentual” languages in terms of tone linking. In a tone language, such as Mende, tones are linked to syllables (or TBUs) in a word from left to right. In contrast, in an accentual language, such as Japanese and English, a designated tone is linked to an accented syllable first (such as the H in the English (M)HL pattern, shown in (3)), and other tones are linked afterwards. However, this proposal implies that there is no accent or stress in a tone language (if there is, the question will be why the accent does not attract a designated tone), which is incorrect. In addition, the difference between Japanese, whose word tones are fairly rigid, and English, whose word tones are more flexible, is still left unaccounted for.

Extending Goldsmith’s idea, I suggest that all languages have both stress (accent) and tone (see section 3.7). In addition, each stressed (or accented) word is associated with a tone pattern. Languages differ in how many tone patterns they have and whether the tone patterns are lexically contrastive. (19) shows how the three traditional typological categories are interpreted in the present system.
(19) Interpretation of tone language typology
   a. Tone language (e.g. Chinese)
      A stressed word can be linked to one of two or more tone patterns that are lexically
      contrastive.
   b. Pitch accent language (e.g. Japanese)
      Every stressed word is linked to the same tone pattern.
   c. Non-tone language (e.g. English)
      A stressed word can be linked to one of two or more tone patterns that are not
      lexically contrastive (depending on intonational meaning).

In Standard Chinese, a stressed syllable can take one of four contrastive tone patterns, as seen in
(1). In Japanese, there is one tone pattern, (L)H(L), for all stressed words, as seen in (2). In
English, a stressed word can take several tone patterns. Two examples are shown in (20), taken
from Bolinger (1986, 24), where stress is indicated by * and where the contours Bolinger drew
were converted into tone features.

(20) Two patterns of tone (intonation) in English
   a. L H-L H L HL
      She only wants to help.
      *   *   *
   b. H L-H L H LH
      She only wants to help.
      *   *   *

According to Bolinger, (20a) “is more like a reprimand or ‘telling’ rather than ‘persuading’”, and
(20b) “might be used to soothe someone who has misunderstood the woman's motives”. The first
two stressed syllables are linked to a H in (20a) and a L in (20b). The last stressed syllable is linked
to HL in (20a) and LH in (20b). According to Pierrehumbert (1980), English has seven tone
patterns for a word with a stress (boundary tones aside). However, whatever its tone pattern, the
lexical meaning of an English word remains unchanged.

The categories in (18) may not apply in all cases. For example, Chonnam Korean has two
word tone patterns, HHL and LHL, predictable from the underlying voicing of the initial segment
(Jun 1990). If voicing is taken to be contrastive in Chonnam Korean, then its tone patterns are not,
so Chonnam Korean is not (18a). And since there are two tone patterns, Chonnam Korean is not
(18b) either. Finally, the choice of the tone patterns does not depend on intonational meaning, so
Chonnam Korean is not (18c). Languages like Lithuanian present another kind of problem. In
Lithuanian, a long syllable can carry one of two contrastive tones, a rise or a fall, shown in (21).

(21) Tone in Lithuanian
    HL-L          LH-L
    [viiras] [viinas]

‘man’ ‘wine’

The data can be analyzed in three ways, shown in (22), where VV represents a long vowel.

(22) Three analyses of Lithuanian tone
   a. Stress analysis
      *
      VV VV
      H H
   b. Tone analysis
      VV VV
      H H
c. Stress and tone analysis

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VV</td>
<td>VV</td>
<td></td>
</tr>
<tr>
<td>HL</td>
<td>LH</td>
<td></td>
</tr>
</tbody>
</table>

In (22a), suggested by Halle and Vergnaud (1987), an accent is assigned to either the first or the second slot of a long vowel. Then, H is assigned to the accented slot and L is assigned to other vowel slots. However, (22a) assumes that accent can occur on the second mora of a heavy syllable, which is metrically controversial (Kager 1993, Hayes 1995). In (22b), suggested by Blevins (1993), H is lexically linked to either the first or the second slot of a long vowel. A default L can then be assigned to other vowel slots. As Blevins pointed out, (22b) avoids the use of accent and so it appears to be more economical. On the other hand, (22b) implies that Lithuanian has no metrical system, which is doubtful. Besides, (22b) does not simplify the theory overall, since in languages like Chinese and English, both stress and tone must be posited anyway (see section 3.7). (22c) assumes that Lithuanian has accent and two lexical tone patterns, LH and HL. Which analysis is correct depends on (a) whether Lithuanian has accent and (b) whether accent can fall on the second mora of a heavy syllable. If (22a) is correct, Lithuanian is a pitch-accent language. If (22c) is correct, Lithuanian is a tone language. If (22b) is correct, Lithuanian belongs to none of the categories in (19). Whether all languages can be subsumed under a single tone-accent theory is an interesting question that invites further discussion.

3.9. Tone domains

The term “tone domain” has been used to refer to two different things. First, it refers to a stretch of sounds over which an expected contour occurs. For example, in Shanghai Chinese the pattern [HL] can occur on one or more syllables, as in (23). A monosyllabic expression is clearly lengthened, which is shown in (23). The initial syllable of a polysyllabic domain is also longer than noninitial syllables (cf. Zhu 1995), which is not shown in (23) (see Duanmu 1999 for a more detailed account of syllable length in Shanghai).

(23)  Tone domain in Shanghai

```
see se pe se du pe
'three' 'three cups' 'three big cups'
```

Domains like those in (23) have one stress, therefore they are a stress domain, or a foot. A second meaning of tone domain refers to a stretch of sounds that are affected by a tone rule. For example, in Xiamen (Chen 1987), every full syllable has two lexical tone patterns, one used in the final position (A) and the other in nonfinal positions (B). An example is given in (24), following Chen’s transcription, where the pitch values of A and B tones need not concern us.

(24)  Tone domain in Xiamen (Chen 1987)

```
pang hong- ts'e
'fly kite'
```

A domain like (24) has been called a tone domain because every nonfinal syllable takes its B tones. But a domain like (24) is not a single stress domain. In (24) all the three syllables are stressed and bear their own lexical tones, whether it is the A or the B copy (the syllable ts'e has a long vowel, which is not shown in the transcription). Both A and B copies are lexically specified and totally unrelated to what tones other syllables have.

Both kinds of domains interact with morphology and syntax in intricate ways. For

3.10. Is tone a prosodic feature?

Tone has traditionally been considered a prosodic feature, which belongs to a unit larger than a segment (e.g. Chao 1930, Pike 1948, Firth 1957, and Wang 1967). There are several reasons for this view. First, a tone pattern usually remains constant independent of its carrier (see (3) and (23)). Second, tone can survive vowel deletion and relocate to another vowel. Third, tone is quite free to move or spread from one syllable to another; in contrast, segmental features are usually not so mobile.

On the other hand, there are reasons to consider tone a segmental feature. First, although a tone pattern can extend over several syllables, each moraic segment carries just one tone, H or L. In this regard, a tone feature ([+H] or [-H]) is like a segmental feature, which ultimately resides in a segment and which cannot occur twice within a segment (see sections 3.4 and 3.5). Second, a suprasegmental feature is usually thought to be different from a segmental one. However, in many Asian languages, tones came from consonant features. For example, in (12) [+slack] is a voicing feature on the consonant and Register feature on the vowel. If we are to treat a feature consistently, we ought to consider [+slack] to be a segmental feature in both cases. Third, tone is not unique in surviving segment deletion. Features like [+nasal], [+back], and [+round] are also able to survive segment deletion. Finally, let us consider the mobility of tone. First, it will be noted that tone is not unique in moving or spreading. Vowel features [back], [high], [low], and [round], and consonant features like [nasal] and [retroflex], have all been found to spread in harmony processes. In the current view of feature theory (e.g. Sagey 1986, Halle 1995), every feature lies on an independent tier. A feature F can move or spread when there is nothing in its way, that is, when there is no specification of F in the intervening segments. Tone is more mobile only because unstressed syllables are not specified for tone (cf. the Tone-Stress Principle in (15)), as a result of which tone can spread onto and through them. It is worth noting that in tone spreading it is usually the Pitch feature that spreads, and not the Register feature. This follows if Register is the same as voicing (sections 3.1, 3.2, 3.3, and 3.6), which is usually specified in consonants. Since most syllables have a consonant, Register rarely spreads out of a syllable.

We have seen then that tone features are not fundamentally different from features like [nasal], [round], and [back], which are usually considered segmental features. Of course, one can define suprasegmental features by their behavior in a given context. For example, one can define a suprasegmental feature as one that spreads across syllables. Thus, [nasal] is a suprasegmental feature when it spreads, but a segmental feature when it does not. However, by this definition, tones are also segmental features when they do not spread, and since in most Asian languages tones do not spread, one would conclude that Asian tones are mostly segmental. But the usefulness of this definition is not obvious: as discussed in section 3.7 and 3.9, whether a tone pattern spreads or not depends on stress. When there are many stressed syllables, tones appear to be inactive, and when there are fewer stressed syllables, tones appear to be mobile. A definition of suprasegmental features based on their mobility would therefore be quite superfluous.

In the phonemic tradition, tone has sometimes been called a phoneme. This is primarily because lexical tones are contrastive and they did not appear to be part of the segments. In the present view, tones reside on segments and tone features are essentially similar to other segmental features. There is, therefore, no motivation to call tones phonemes, just as there is no motivation to call [nasal] or [round] phonemes.
3.11. Categorical and gradient values

There has been some debate about the categorical vs. gradient nature of linguistic features and tone bears on this issue. In real speech, pitch levels are infinitely variable. However, no language uses more than a small number of contrastive tone levels. In particular, as discussed in section 3.1, there may be only four contrastive tone levels, and if Register is cued by voice quality (sections 3.2 and 3.3), there may be just two levels that are cued by pitch, H and L. In any case, contrastive tone levels are strikingly small in number, especially in view of the fact that in singing any ordinary person can use at least a dozen or so music notes at ease. The paucity of contrastive tone levels supports the view that linguistic features are categorical in nature.

3.12. Tone and pitch values

In real speech the same tone feature can occur at different pitch levels. This is due to several factors. First, H has a higher pitch with greater stress than it does with lesser stress. Second, the same tone will gradually descend in pitch owing to the downdrift effect. Downdrift was first observed in African languages but was later found in other languages, too, such as English, Japanese, Korean, Chinese, and German (Pierrehumbert 1980, Beckman and Pierrehumbert 1986, Jun 1990, He and Jin 1992, Féry 1993). Third, segmental features, such as voicing and aspiration, can also affect pitch contour. To interpret the pitch contours of a language, therefore, it is necessary not only to know its tone patterns but also its prosodic structure, including various levels of stress that are sensitive to phrasal structure and pragmatic focus (e.g. Gussenhoven 1984, Bolinger 1985, McLemore 1991, Selkirk 1995, among others). This is a big challenge, not only because phrasal stress and pragmatic focus are less understood than word stress, but also because stress in tone languages is not as obvious as stress in non-tone languages.

3.13. Tone and intonation in the same language

According to Liberman (1975) and Pierrehumbert (1980), among others, intonation patterns in English can be represented by a linear sequence of Hs and Ls. The sequence is made by a proper choice of tone patterns for each accentual phrase plus a proper choice of boundary tones and “phrase accents”. However, in languages like Chinese, word tones are lexically determined, and there is little flexibility in varying the sequence of Hs and Ls independently (I ignore tonal Register here). This raises the question of how languages like Chinese express intonational meaning, such as statement, doubt, surprise, query, command, etc.

Chao (1933) suggests that many functions of intonation in other languages are fulfilled in Chinese by the use of particles. When particles are not used, tone and intonation can be combined through “addition”. There are two cases. The first is “successive addition”, where a rise or a fall is added to the end of an utterance; this is similar to a boundary H or L in English. Two examples in Standard Chinese are shown in (25) (cf. Chao 1933, 131). Strictly speaking, the resulting syllable must be lengthened to three moras in order to carry the three tones (cf. section 3.5). For simplicity the lengthening is not shown.

(25) Successive addition of tone and intonation in Standard Chinese

<table>
<thead>
<tr>
<th>Tone</th>
<th>Intonation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>LH</td>
<td>L</td>
<td>LHL</td>
</tr>
<tr>
<td>nan</td>
<td>nan</td>
<td>nan</td>
</tr>
<tr>
<td>‘difficult’</td>
<td>‘affirmation’</td>
<td>‘Surely difficult!’</td>
</tr>
</tbody>
</table>
Tone + Intonation  \[ \text{HL} \rightarrow \text{HLH} \]

\text{mai} ‘sell’ \quad \text{‘question’} \quad \text{mai} ‘Sell?’

[nan] ‘difficult’ has a rising tone (LH). The boundary intonation L carries an affirmation meaning. When the two are added, the result is a rise-fall (LHL). Similarly, [mai] ‘sell’ has a falling tone (HL). The boundary intonation H carries a question meaning. When the two are added, the result is a fall-rise (HLH).

The second case is “simultaneous addition”, where intonation is superimposed on word tones, as a result of which the pitch range of an utterance is raised, lowered, expanded, or compressed, while the word tones remains distinctive. Simultaneous addition is supported by the phonetic studies of Shen (1989) and He and Jin (1992), who found that the question intonation raises the pitch height of the entire utterance, without changing the distinctiveness of word tones.

If simultaneous addition occurs, tone and intonation must be different entities. This raises a serious problem. First, to represent tone and intonation in Chinese, a single sequence of Hs and Ls (ignoring Register) is not enough. Second, there is the potential question of whether English pitch contours reflect tone, or intonation, or both. If they reflect tone, one wonders why English has no intonation. If they reflect intonation, it means that, like tone, intonation can be represented by a linear sequence of Hs and Ls, and then there is the question of what relation there is between the intonational Hs and Ls and the tonal Hs and Ls. If they reflect both, and if Liberman (1975) and Pierrehumbert (1980) are essentially correct, then one wonders why a single sequence of Hs and Ls can represent both tone and intonation in English but just tone in Chinese.

The solution to the problem is not obvious. But let us consider two proposals. First, it is possible that languages can use one of two ways to indicate an intonation: either by a boundary tone, or by changing the pitch level of an entire utterance. In the first way, intonation and word tones are in a linear sequence (Chao’s “successive addition”). English uses this method throughout, but Chinese uses it only sometimes. In the second way, intonation and tone are superimposed on each other (Chao’s “simultaneous addition”). English does not use this method, but Chinese sometimes does. As a result, English tone and intonation can be represented by a single sequence of Hs and Ls, whereas Chinese tone and intonation can be represented this way only sometimes. This proposal essentially admits simultaneous addition at least in some languages and calls for a representation that goes beyond a single sequence of Hs and Ls. For example, one may assume two levels of representation, a lower level of Hs and Ls for word tones, and a higher level of Hs, Ls, rises, falls, etc., for phrase and sentence intonation.

The second proposal attempts to preserve a single sequence of tones, and account for their undulations in other ways. A motivation for this proposal is that, if [+H] and [-H] are values of a distinctive feature, there ought to be just one sequence of them, since there is only one sequence of values of any other distinctive feature. It is worth noting that Chao (1933) mentioned only two cases of simultaneous addition in Chinese, “raised level of pitch” and “lowered level of pitch” (pitch range widening and narrowing were found to be correlated to pitch raising and lowering). In He and Jin (1992), six intonational meanings were designed (statement, expecting confirmation, question, simple request, command, and exclamation), but again just two cases of simultaneous addition were observed: raised or lowered. If these are the only cases of simultaneous addition (besides the neutral intonation), there is a possible solution. The raised level may have an accented final boundary H, which (whether it is actually produced or not) is targeted at a higher pitch level (as any H with a main accent does) and so may have prevented downstep in the preceding syllables. Similarly, the lowered level may have an accented final boundary L, which is targeted at a lower pitch level and so may have accelerated downstep in the preceding syllables. At this point,
this proposal is rather speculative. Whether it is a correct account of simultaneous addition will be left open.

3.14. Functions of tone

A basic function of tone is lexical distinctiveness. However, tone can serve other functions, such as syntactic distinction and attitudinal meaning. For example, in Igbo, a floating H can indicate the associative structure “noun of noun” (Williamsom 1986). In Chinese, a boundary H can indicate a question, and a boundary L can indicate an affirmation or a “protesting statement” (Chao 1933, 131-133). In English, attitudinal meanings can also be expressed by word tones (pitch-accents) and boundary tones (e.g. Liberman 1975, Pierrehumbert 1980). The fact that tone has multiple functions is again not unique among distinctive features. For example, palatalization ([back, +high]) is lexically distinctive but can also indicate the mimetic structure in Japanese (Mester and Ito 1989).

4. Conclusions

Research on tone in the past three decades has tremendously increased our understanding of the subject. It also triggered developments in many other areas of phonology. For example, the rise of autosegmental phonology, resulting from tone research, has brought about progress in such areas as harmony processes, templatic morphology, syllable structure, feature geometry, and metrical phonology. Works on tone domains have also propelled research on the syntax-phonology interface. For example, Clements's (1978) work on Ewe and Chen's (1987, written in 1985) work on Xiamen have lead to the influential end-based theory of Selkirk (1986) on the syntax to phonology mapping. Works on tonogenesis and tone features call for a reconsideration of the traditional distinction between segmental and suprasegmental features. Finally, works on contour tones have direct implications for the status of contour features in distinctive feature theory.

As seen in section 3, many issues remain unresolved and will continue to concern researchers for some time. In addition, although I have not discussed the large body of phonetic literature on tone, it is clear that tone is a principal area in which phonetic and phonological researches come together.

Acknowledgments

Thanks to Pam Beddor, Lesley Milroy, James Myers, Bonny Sands, and Rint Sybesma for their comments.
References
(Note: as requested by the editors, more items are included than cited in the text)


Reproduced by the Indiana University Linguistics Club, Bloomington, Indiana.


van der Hulst, H. & N. Smith (eds.) 1988. *Autosegmental studies on pitch accent*. [Linguistic...


