TONE FEATURES AND TONE RULES

Victoria A. Fromkin
Department of Linguistics
University of California, Los Angeles

1. Introduction

Linguists, working within the Generative Phonology paradigm, are attempting to establish a theory which defines formally and substantively the phonological components of grammars which will specify all and only the set of possible sound systems of human languages. Such a theory must provide a set of features for characterizing tone and other prosodic phenomena of language. It must also include constraints on tonal representation, the formal properties of tonal rules, and conventions which govern the interactions of tone rules with other rules of the grammar. This paper attempts to deal with some of these important questions. In particular it will discuss the distinctive features of tones (level and contour tones), tonal representation, and rules in a generative grammar.

2. The Distinctive Features of Level Tones

In one of the earliest proposals concerning "The distinctive features of tone" [Gruber, 1964] two tonal features, High and High 2 were proposed, which features provided the means for distinguishing between two, three, or four level tones. Implicit in this proposal is the claim that the basic distinction in any tone language is between high tones and non-high tones, with all other tonal contrasts being made within these two disjunctive sets. Thus, a language with two contrastive tones, would utilize only the feature High, a language with three or four contrastive tones would, in addition, specify tones using the feature High 2 as shown under (1).

(1) a. Two tone contrast: [+high] and [-high]

b. Three tone contrast: [+high] [+high] [-high] [-high]

or [+high] [-high] [-high] [-high]

[+high 2] [-high 2] [-high 2] [-high 2]

c. Four tone contrast: [+high] [+high] [-high] [-high]

[+high 2] [-high 2] [-high 2] [-high 2]
Wang [1967] accepts Gruber's primary division between + and - High, but proposes two additional tonal features to replace High 2: Central and Mid. (In addition, he adds four 'contour' tone features, which will be discussed below.) Assuming a number of universal redundancies which he proposes, by this system, five level tones can be distinguished, as shown in (2).

\[
\begin{array}{c}
\text{High} & + & - & + & - & - \\
\text{Central} & - & - & + & + & + \\
\text{Mid} & - & - & - & - & + \\
\end{array}
\]

Wang's features seem immediately superior to Gruber's in that the ad hoc feature High 2 is replaced by features with greater phonetic plausibility, and, in addition, five rather than four tones may be contrasted.

Sampson [1969] finds certain difficulties with Wang's system, particularly in relation to the feature Mid, which according to Wang would be needed only for those languages with five level tones. He proposes, therefore that the three tone features be High, Central, and Low providing the contrasts as shown in (3).

\[
\begin{array}{c}
\text{High} & + & + & - & - & - \\
\text{Central} & - & + & + & - \\
\text{Low} & - & - & - & + & + \\
\end{array}
\]

Although in this proposal, both the features Central and Low are used only when there are more than three contrastive tones, it is of course possible that the particular tone system of a language could (unless constrained by the theory) utilize the features High and Central or High and Low. This position is argued for below.

The features suggested by Sampson are essentially those adopted by Woo [1967, 1969].

At a conference on 'Tone in Generative Phonology', held in Ibadan in 1970, following a suggestion by Maddieson [1971], the participants concluded that the features Raised and Lowered should be substituted for High and Low respectively "to avoid possible confusion with 'High' and 'Low' as features of tongue height." [p. 76] They further concluded that
for a four or five level system a feature Extreme be utilized. Two, three, four, or five tone systems would thus be represented as given in (4).

(4) a. **two tone system**

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Mid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raised</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

or ii.  

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Mid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowered</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

b. **three tone system**

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Mid</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raised</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lowered</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

c. **four tone system or five tone system**

(in a four tone system, the fourth tone is either 'extra high' or 'extra low')

<table>
<thead>
<tr>
<th></th>
<th>Mid</th>
<th>High</th>
<th>Low</th>
<th>Extra high</th>
<th>Extra low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raised</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Lowered</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Extreme</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

[Cf. Maddieson, 1971]

Maddieson attempts to present both phonetic and phonological reasons for the four level tone features exemplified in (4). The discussion does not seem to be over as yet. In a paper presented to the Seventh International Congress of Phonetics, Halle [1971] continues the debate. He agrees that "it is clearly necessary that the universal phonetic framework provide for a distinction of at least three pitch levels: high, mid, and low." Relating these pitch distinctions to vocal cord stiffness, he proposes two binary features which can distinguish these three tones: [stiff vocal cords] and [slack vocal cords]. (This proposal is earlier made in Halle and Stevens [1971]. The earliest proposal concerning these features was put forth by Maran [1968] and adopted by Bird [1971].)
The earlier discussions on tonal features do not attempt to provide physiological correlates for them but relate them more specifically to auditory correlates. Thus Maddieson notes that "the phonetic correlate of [+Raised] is 'higher than a notional median pitch' and of [-Raised] 'not higher than a notional median pitch'." [p. 8] The phonetic correlates of Lowered are described similarly. Halle's proposal on the other hand ties the pitch levels directly to glottal states as is shown in (5).

(5) \begin{tabular}{ccc}
Stiff & Mid - \( \hat{V} \) & Low - \( \check{V} \) & High - \( \check{\hat{V}} \) \\
Slack & & & \\
\end{tabular}

(The following symbols will be used throughout:
\( \hat{V} \) = Mid Tone; \( \check{V} \) = Lower Mid Tone; \( \check{\hat{V}} \) = High Tone;
\( \check{\check{V}} \) = Low Tone; \( \hat{\check{V}} \) = Rising Tone; \( \hat{\hat{V}} \) = Falling Tone;
\( \check{\hat{\check{V}}} \) = Down Step)

In a language with three contrastive tones, such as Yoruba or Nupe, these two features are needed to specify the phonological vowels as to tone. For languages in which there are only contrasts between two tones, such as Akan, Igbo or Hausa, it would seem, then, that either the feature Slack or the feature Stiff would be redundant. In languages such as English with no lexical tonal contrast, vowels would all be [-stiff, -slack], neither feature being distinctive. This follows the suggestion that the vowel specified as [-stiff, -slack] is the 'neutral' vowel [Halle, 1971; Halle and Stevens, 1971]. Marking conventions can then specify a 'mid tone' vowel as the unmarked vowel as shown in (6).

(6)

a. \( u \) Stiff vocal cords \( \rightarrow \) - / [\( \hat{V} \) +syllabic \[\hat{\check{\hat{V}}} \) -consonantal]

b. \( u \) Slack vocal cords \( \rightarrow \) - / [\( \check{\hat{V}} \) +syllabic \[\hat{\check{\hat{V}}} \) -consonantal]

(But see below for marking conventions concerning these features when specified for consonants.)
For languages with two contrasting tones only one vowel would be marked. For Akan or Igbo (following the suggestion of Maddieson [1971]) the high tone vowel would be marked in contrast with the neutral vowel, while in Hausa, the low tone would be marked. A non-tone language such as English would then have all unmarked vowels for these features.

One problem is immediately apparent using the Halle/Stevens tonal features. Halle states that the theory must distinguish "at least three pitch levels." 'At least' is not 'only'. There is no way given the features Slack and Stiff to distinguish more than three level tones, since, as they themselves point out, [+stiff, +slack] is physiologically impossible and must therefore be disallowed by marking conventions. Yet, "systems with four contrasting levels...are clearly established for several languages" as shown by Welmers [forthcoming] in the following examples:

(7)

<table>
<thead>
<tr>
<th>Tones</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tigong:</td>
<td></td>
</tr>
<tr>
<td>a. efíŋpí</td>
<td>mid-high-high</td>
</tr>
<tr>
<td>b. esya</td>
<td>mid-mid</td>
</tr>
<tr>
<td>c. enwá</td>
<td>mid-lowered mid</td>
</tr>
<tr>
<td>d. eki</td>
<td>mid-low</td>
</tr>
<tr>
<td>Ndoro:</td>
<td></td>
</tr>
<tr>
<td>e. síná</td>
<td>high-high</td>
</tr>
<tr>
<td>f. čarú</td>
<td>mid-mid</td>
</tr>
<tr>
<td>g. čsílā</td>
<td>low mid-low mid</td>
</tr>
<tr>
<td>h. šóřá</td>
<td>low-low</td>
</tr>
<tr>
<td>'axes'</td>
<td>'holes'</td>
</tr>
<tr>
<td>'men'</td>
<td>'canoes'</td>
</tr>
<tr>
<td>'adze'</td>
<td>'axe'</td>
</tr>
<tr>
<td>'stone'</td>
<td>'chicken'</td>
</tr>
</tbody>
</table>

Longacre [1952] has presented evidence for five distinctive pitch levels in the Mixtecan language Trique, and Fang Kusi Li reported that Black Miao has five level tones and two rising and falling tones (as reported by Voegelin [1965]).

It is of course true that phonetically many more levels of pitch must be distinguished. There are a number of ways to specify these, the easiest probably being to rewrite tonal features into relative pitch levels using integers. But the question as to whether we need more than three phonological tone contrasts is, to quote a linguistic cliche,
'an empirical one'. If Welmers and Longacre and Fang Kusi Li are correct, the Halle tonal features would be descriptively inadequate.

There are additional problems which arise even if it is shown that all the four and five level systems are lexically or basically three tone systems.

One of the main motivations for using Stiff and Slack as features for tonal specification was because "the same set of features governs both pitch levels in vowels and voicing in obstruents" [Halle, 1971]. This proposal then is based on the claim that one finds "three types of obstruents: voiceless, voiced, and intermediate: the first corresponding to the high pitch vowels, the second to the low pitch vowels, and the third to vowels with mid pitch." It is further argued that "the appearance of high pitch in a vowel adjacent to a voiceless consonant, and of low pitch in a vowel adjacent to a voiced consonant is not fortuitous but rather a case of assimilation (supporting)...the claim...that the same set of features governs both pitch levels in vowels and voicing in obstruents." Halle therefore proposes to substitute the two features discussed above for the feature Voiced. (Two additional features, Spread Glottis and Constricted Glottis are also substituted for the earlier features Tensity, Glottal Constriction, and Heightened Subglottal Pressure [Halle and Stevens, 1971]. A criticism of the phonetic adequacy of these features for consonants is beyond the scope of this paper but cf. Lisker and Abramson [1971], Ladefoged [1971a].)

Given these two features, the contrast between voiced and voiceless stops in languages with a two-way distinction would then be represented as:

(8)    Voiceless   Voiced
       Stiff       +     -
       Slack      -     +

Since the unmarked obstruent is presumably voiceless, we can reveal this by changing the Marking conventions given under (6) above to:
In other words, the unmarked value for Stiff for vowels is minus, but plus for consonants; and the unmarked value for Slack is minus for true vowels and true consonants, and glides. (Glides are also specified as [-stiff, -slack].) No mention is made of liquids [Halle and Stevens, 1971].

The highly marked nature of 'slackness' and the asymmetry of these two features is not the basic difficulty. Using the former feature Voiced, a phonological rule which voices obstruents intervocally reveals this assimilation as a change from an unmarked feature value to a marked [Schachter, 1969]. Using the new features the same process occurs.

(10) [+stiff] \(\rightarrow\) [-stiff] / [-stiff] ___ [-stiff]

But this assimilatory rule also occurs in tone languages. For the assimilatory nature of the process to be revealed, the intervocalic voicing would have to be restricted to occurring between non-high vowels, since high tone vowels, like voiceless obstruents are [+stiff, -slack]. Where intervocalic voicing occurs before all vowels, regardless of the tones, the rule will have to be written as in (11).

(11) [+stiff] \(\rightarrow\) [-stiff] / [+syl] ___ [+syl].

But (11) makes it appear that this 'unstiffening' process is non-assimilatory, or ad hoc. Furthermore, in languages where vowels are devoiced after certain voiceless obstruents (such as in Japanese) the rule must be stated as in (12).

(12) [+syl] \(\rightarrow\) [+spread glottis] / [+syl, +stiff] ___
Using these features we are forced to abandon our traditional collective linguistic intuition that the voicing and devoicing of consonants and vowels is a natural process, i.e. that the voicing of intervocalic consonants occurs because vowels are voiced, irrespective of the pitch of the vowels. This is not to say that there is no relationship between tension of the vocal cords and pitch. But a rise in pitch may result from either an increase in the tension of the vocal cords or an increase in the air pressure below them [Ladefoged, 1963, 1964, 1967, 1971a; Ohala, 1970]. In fact, in tone languages, "there is often an increase in subglottal pressure during high tones" [Ladefoged, 1971b]. The attempt, then, on the part of Halle to explain the relationship between voicing and tones, obscures the relationship between identical glottal strictures for consonants and vowels. Ladefoged suggests instead that a feature *glottal stricture* be posited to account for just such phenomena, and for other phonological oppositions Halle and Stevens specify by their features. This feature would have a number of possible values (although, according to Ladefoged, any one language will not utilize more than three contrasting values.) He states that the feature defines a continuum and that it is impossible (even meaningless) to state how many possible values there are. He further suggests that on the classificatory level we might specify this feature by binary values for languages using only two states of the glottis, and by the integers /0 1 2/ in the classificatory matrices for languages contrasting three states of the glottis. Then, by either universal 'interpretive rules' or by context-restricted phonological rules, these items would be given appropriate values.
While I am in basic agreement with this proposal, the need for language specific mapping rules would create an unnecessary problem. If, however, the continuum is divided into a given set of discrete values and if, for example, [5 glottal stricture] means voiced for all languages, one can easily write a natural rule of voicing assimilation as in (14).

(14) [+segment] \(\rightarrow\) [5 GS] / [5 GS] ___ [5 GS]

One may recall that Chomsky and Halle [1968] argued for the substitution of the features High, Back and Low for the earlier features Diffuse, Compact and Grave because "the former framework...did not bring out the fact that palatalization and velarization characteristically occur before front and back vowels, respectively; the connection between palatalization and front vowels and between velarization and back vowels was no more motivated than a connection between glottalization or voicing and front vowels." [p. 308] If we accept this reasoning, and I think we should, one cannot at the same time accept the new features proposed, since the intervocalic voicing of obstruents between high tone vowels or the unvoicing of vowels after voiceless consonants would then be "no more motivated than a connection between...voicing and front vowels."

The desire to select features in the universal set which will explain diachronic changes as well as synchronic phonology is admirable. Until we can find a more 'explanatory' set of features, it would be better to include in the universal theory a set of statements relating glottal strictures and vowel tones than to substitute, for features such as Voiced, features which obscure natural processes in synchronic grammars.

One more question must be considered prior to deciding on the most adequate set of tonal features. In both Halle's system and Maddieson's, the 'neutral' or 'unmarked' tone is the mid tone -- [-stiff, -slack] and [-raised, -lowered], respectively. This presents certain difficulties for 'terraced level' [Welmers, 1959] tone languages. In these languages, one finds three 'phonetic' tones -- high, lowered high ('downstep' or 'drop' represented as \(\breve{\V}\)), and low, in which the 'downstep' tone occurs only after a high tone. This 'downstep' is illustrated in the examples from Akan, shown in (15):
Following a proposal by Stewart [1964], Schachter and Fromkin [1968] derive all 'downstep' tones by a pitch assignment downdrift rule and vowel deletion rules. In such a case, one need not be concerned about the feature specification for the 'downstep' tone, since it would be specified as a high tone (i.e. [+stiff] or [+raised]) and its differentiation from other high tones would be solely based on a pitch value assigned to it. In (15b) and (15c) the underlying forms for 'stone' and 'chest' are $\tilde{b}6$ and $\tilde{b}0$ respectively. The $\tilde{3}$ represents a nominal prefix which is deleted in certain contexts. The following rules will take care of the 'drop' tone which appears on the surface.

(16) Pitch assignment (PA): a. [+ high] $\rightarrow$ p 1  
b. [- high] $\rightarrow$ p 3  
\(p = \text{relative pitch value.}  \)  
\([\pm \text{high}] \) is used here arbitrarily, i.e. for these rules one could also use \([\pm \text{stiff}] \) or \([\pm \text{raised}] \). I am following the suggestion of Johnson [1970] and Williamson [1971] in their comments on Schachter and Fromkin [1968] to designate the highest relative pitch by '1' and all lower pitches by larger integer values.)

(17) Downdrift (DD)  
\(\text{RL: } [\text{aH}] \rightarrow [\text{aH} p \leftrightarrow 1 ] / [\text{aH} p] < [-\text{aH}]_{1} > \)  
\(\text{(where H = high tone)} \)

The 'RL' at the beginning of the rule specifies it is a Right Linear Rule [Johnson, 1970] which applies to the left most segment which meets the rule specification first, then moves to the next segment from left to right, applying in each case when the structural description is met. The formulation with slight changes was provided by Grover Hudson. It abbreviates the following four rules:

\[1\]In (15) and in subsequent examples the Akan utterances are given in orthographic representation except for the tones.
(17) a. \([+H] \rightarrow [+H \ p+1] / [+H \ p] \ [-H]_1\)
b. \([+H] \rightarrow [+H \ p] / [+H \ p] \ __\)
c. \([-H] \rightarrow [-H \ p+1] / [-H \ p] \ [+H]_1\)
d. \([-H] \rightarrow [-H \ p] / [-H \ p] \ __\)

(18) Vowel Deletion (VD) \(V \rightarrow \emptyset\) in certain contexts.

The following derivations illustrate how these rules apply.

(19) / me ɔbɔ / 'my stone' / me ɔbɔ / 'my chest'

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>DD</th>
<th>VD</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1 3 1</td>
<td>1 3 3</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While the above may solve the problem for terraced level languages in which all downstep tones are derived from underlying high tones, it seems that this is not true in all synchronic grammars. One finds that at a certain stage in the history of such languages, the derived 'downstep' may become 'phonemic'. That is, even in Akan, where most 'downstep' tones are still derived from high tones, as in the examples above, there are formatives which now have 'downstep' tones which cannot be derived from high tones after the deletion of low tones without a great deal of ad-hoc-ery, as shown in (20).

(20) a. ɖɛbɛ́ntɛ́ 'young man'
b. ɛkʊ́ 'Akua' (name of a girl born on Wednesday)
c. ɛɛnɛ́ 'yes'

One can, of course, set up an underlying low tone which never appears on the surface and which is later deleted. The 'absolute neutralization' [Kiparsky, 1968] solution only obscures what has really occurred historically in the language, i.e. the derived mid or downstep tone has become phonemic. As Vennemann [personal communication] points out, a solution
which does not posit an underlying 'downstep' tone fails to reveal a complexity in the language which has arisen by historical processes, and such a solution should therefore be considered unacceptable, even if by such a method we appear to arrive at a 'simpler' two-tone language solution.

Assuming the correctness of this position, how would one use either the Halle or Maddieson features to specify the three way underlying tonal contrast? If the 'downstep' tone is considered a 'mid' tone (and in their systems it would have to be), then it is just this mid, neutral, unmarked tone which is highly restricted and very infrequent. This is certainly a counter-intuitive specification. In addition, using their features, the 'downstep' tone would be equally distant from the high and low tones. This of course is not the case, since all derived 'downsteps' emerge from underlying high tones. Neither Halle's nor Maddieson's tone features capture the relationship between High and Downstep. Wang's and Sampson's features do. Thus, given a 'terraced level' language with both underlying 'downstep' tones and derived ones, we can rewrite the tone rules given above. One must also include the Morpheme Structure Condition (21) which constrains the 'Downstep' to positions after high tones.

(21) MSC: If: \([+\text{Mid}]\) Then: \([+\text{High}]\)

The features specifications of the three tones would be:

(22)

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Downstep</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Mid</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

To derive the correct relative pitches in a language with three underlying tones, high, downstep, and low, the pitch assignment rule (16) need not be changed, but the Downdrift Rules must be changed as given in (23):
which expands to:

a. \([+H] \rightarrow [+H p +1] / [+H p] [+M] [-M] [+M] \]

b. \([+H] \rightarrow [+H p +1] / [+H p] [-H] \]

c. \([+H] \rightarrow [+H p] / [+H p] \]

d. \([-H] \rightarrow [-H p +1] / [-H p] [-M] [-M] [+M] \]

e. \([-H] \rightarrow [-H p +1] / [-H p] [+H] \]

f. \([-H] \rightarrow [-H p] / [-H p] \]

(24) provides a derivation which illustrates how the rules work:

(24) \(/ \d‘bèr’antëtë \ bëk’ô \ èhô / [\d‘bèr’antëtë \ bëk’ô \ hô] \)

'"The young man will go there'"

PA 3 3 1 1 1 1 3 1
DD 22 2 2 4 3
VD \(\emptyset\)

By the Pitch Assignment rule all tones designated as \([+H]\) (including the Downstep Tones which are also \([+M]\)) are assigned a pitch value of '1', and all \([-H]\) tones are assigned a pitch, of '3'.

We start applying the Downdrift Rule to the left most tone. None of the rules collapsed by the schema (23) applies since it is the initial tone. Moving to the next tone, a, b, and c are inapplicable since the tone is \([-H]\); d does not apply since the tone is not \([-H, +M]\) (in fact, it can never apply since no tone will be so specified); e cannot apply since the tone is not preceded by a \([+H]\); f applies vacuously since the assigned \(p = 3\) already matches the assigned value to the preceding Low tone. None of the rules apply to the first high tone so the pitch value is left at '1'. a applies to the next tone, since it is specified as \([+H, +M]\) and is preceded by a \([+H, -M]\) tone. We therefore add '1' to the value assigned to the previous High Tone pitch value, deriving a pitch
value of '2'. Only _c_ applies to the next three [\(+H, -M\)] tones, assigning the same pitch value, '2', as has been specified for the previous [\(+High, +Mid\)] tone. Note that the value for the feature Mid does not influence this pitch assignment. The _c_ in _chɔ_ is assigned a pitch value of '4' by _e_, and the final tone is assigned a value of '3' by _b_. After the Vowel Deletion Rule has applied, the final output has two downstep tones, one underlying and the other derived.

By specifying the 'downstep' tone as [\(+high\)] we are able to show the relationship between the High tone pitch values and the 'downstep' tone, and use the same rules to assign pitch values for underlying 'drop' tones as well as derived 'drop' tones. This, I believe, argues strongly against either the Halle features or the Maddieson features. One can of course write a complicated set of rules deriving the correct pitch values using features which specify the 'downstep' as 'equally related' to both the high and the low tones. The effort does not seem warranted however, since the Halle features have been shown to be deficient on other accounts, and the Maddieson features do not seem to add anything or explain any more than do Sampson's or Wang's features. Rather, less is explained, since using a _mid_ feature the historical development of three-tone languages from two-tone languages is more easily revealed.

For the reasons given above, I propose that the Universal Set of Distinctive Features include three tone features: High, Mid, Low. In addition, I suggest that it is impossible to specify which tone is the 'unmarked' tone for all languages, since each particular tone system evolves historically in different ways. While the mid tone may be the least marked tone in a language such as Yoruba, in Twi where we find a three tone system developing from a two tone system, it is the low tone which is 'unmarked'. One may, however, suggest that the historical changes which occur will move toward a more stable three tone system in which the 'downstep' will become the 'neutral' tone, or it may become a contour tone as in Gwari [Larry Hyman, personal communication]. More research is needed on the historical developments of tone languages before we can reach this conclusion. Finally, I suggest that there is a hierarchy for tonal features, i.e. that in tone languages the basic
division, following Gruber and Wang, is between high and non-high tones, but that the features Low and Mid are equally placed in this hierarchy. That is, for a three tone language, either Low or Mid will be the second tone utilized, depending upon the particular phonological system of rules in the synchronic grammar. Similarly, for a four or five tone system, the feature specifications for the mid tones should depend on the particular tonal rules present, i.e. a four tone system may include High, High-Mid, Mid, and Low tones, or, High, Mid, Low-Mid, Low, etc.

We may, of course, find, upon further investigation, that there are universal constraints governing the interaction of tones in a multi-tone system. Before such empirical evidence is presented, however, our theory must at the minimum be descriptively adequate. Placing constraints at this time may force us into Procrustean solutions which will obscure the intricacies of tonal phenomena.

3. The Need for 'Contour' Features

In Section 2, the features for the specification of level or 'register' tones were discussed. Phonetically, all linguists have observed the occurrence of 'contour' or 'non-stationary' tones. To provide for such occurrences, Wang [1967] proposed that a feature Contour be used to distinguish stationary from non-stationary tones. He further suggests that for tones specified as [+ contour], three additional features be available: Rising, Falling, and Convex. All [+ contour] tones would be redundantly [- central, - mid]. Using this set of features, one could then distinguish eight contour tones as shown in (25).

\[
\begin{array}{c}
\text{Contour} & + & + & + & + & + & + & + \\
\text{High} & + & - & + & - & + & - & - \\
\text{Rising} & + & + & - & - & + & + & + \\
\text{Falling} & - & - & + & + & + & + & + \\
\text{Convex} & - & - & + & + \\
\end{array}
\]

(Note that Convex is redundantly - for contour tones which have opposite values for Rising and Falling.)
There are two main questions which concern contour tones: (1) whether Contour tonal features are needed at all, and (2) if they are needed are Wang's features the set which should be included in the theory.

As to the first question, a suggestion by Woo [1967, 1968, 1969, 1970] that contour tones should be represented in the lexicon in all cases as sequences of level tones was adopted by the Ibadan Conference [1971], by Leben [1971] and by Halle [1971]. Halle extends the proposal to the phonetic level as well as the classificatory level. "...on the systematic level all tones are stationary. Non-stationary tones, such as 'rising', 'falling', or 'convex' are more or less surface phenomena; they have much the same status as the different formant transitions that are found in a given vowel when it is adjacent to different stop consonants." In other words, according to Halle, the universal set of features will not include contour features for tone.

The evidence for restricting underlying, phonological tones to stationary tones is based to a great extent on the existence of 'tone copy' rules. It is shown, for example, that there are no cases reported where a contour tone is copied; in the case of rising tones, a high tone is copied, and in the case of a falling tone, a low tone is copied, when it is the following tone which is changed [Leben, 1971]. Leben provides numerous examples from Hausa, Yala, and Mende in support of this position; Woo [1970] shows this to be the case in North Tepehuan; Halle [1971] presents further evidence from Serbo-Croatian and Slovenian.

In Nupe, George [1970] has shown that all rising and falling phonetic tones can be derived from underlying level tones. Rising tones are the result of a 'tone copy' rule which George writes as in (26)

\[(26) \quad [+H] \rightarrow [+G] / [+L] [+Cons, +Vcd] \quad \]

\([+H] = \text{High Tone}; \ [+L] = \text{Low Tone}; \ [+G] = \text{Gliding Tone}\]

This rule can also be written as (27) if it is stated that a sequence of two immediately following tones are realized phonetically as a glide.

\[(27) \quad [+H] \rightarrow [+L] [+H] / [+L] [+Vcd] \quad \]
In fact, unless one posits the rising tone (i.e. Low to High Tone Glide) as a sequence of two tones, the first being a copy of the tone which immediately preceded it, the rule seems very ad hoc.

Since no evidence has been put forth showing the need to consider underlying contour tones as intrinsic, can we then dispense with any contour tone features? Do contour tones indeed "have the same status as the different formant transitions"?

To exclude from the universal set of features any contour tonal feature would violate the very goals Halle accepts: "all grammatically determined facts about the production and perception...are embodied in the 'phonetic transcription'" [Chomsky and Halle, 1968]. Furthermore, as Chomsky [1967] points out: "it is important to note that the distinctive features postulated in universal phonetic theory are absolute in several senses but relative in others. They are absolute in the sense that they are fixed for all languages. If phonetic representation is to provide sufficient information for identification of a physical signal, the specification of feature values must also be absolute." [p. 404]

If we find that in every language, a succession of two tones is always realized as a contour or gliding tone, the absolute nature of the phonetic signal is fixed and one can, as Halle suggests, dispense with any contour tonal feature.

Unfortunately, the facts seem to contradict this assumption. In Nupe, as rule (27) shows, the sequence of a low tone followed by a high tone is realized as a low tone followed by a rising tone only if the two syllabic segments are separated by a voiced consonant. Thus one finds the following phonetic contrasts: (all examples from George [1970]).

(28) a. étú [ étú ] [ _ - ]  'parasite'
    b. édú [ édú ] [ _ ˛ ]  'taxes'

(29) a. ékpá [ ékpá ] [ _ - ]  'length'
    b. égbá [ égbá ] [ _ ˛ ]  'a garment border'
Leben [1971] provides other examples from Yala where phonetically one must differentiate between a high tone followed by a low tone, and a high tone followed by a falling tone:

(30) a. ò kà [− \_\_] 'he said'
   b. kò kà [− _ ] 'let him speak'

According to Leben, (30b) does not have the glide because of the deletion of an intervening low tone, which rule follows the glide formation rule. If the phonetic representation is to represent the grammatically determined facts of Yala, and "sufficient information for the identification of a physical signal", then the difference between a contour tone and a sequence of level tones must somehow be represented in the systematic phonetic output of the grammar.

Another example of the need for a contour/non-contour contrast is given by Stewart [1962]. He points out that in some dialects of Fante, depending on the syntactic and lexical features of a string, a pre-pause high tone may be realized as a high-rising gliding tone rather than a high level pitch, as illustrated in (31).

(31) a.  ámbékà [ _ − ] 'he will remain'
    b. ámbékà [ _ − − ] 'he will bite it'

Furthermore, in Nupe, there is a phonetic falling tone which derives from a sequence of either high-low or mid-low in rapid speech, which is realized as step tones in slow deliberate speech as shown in (32) [George, 1970].

(32) ebe ëtʃ 'monkey howling'
    a. slow speech: [ebeëtʃ] [− − − ]
    b. rapid speech: [ebeëtʃ] [− \_ − ]

Before discussing the implications of these facts for phonological theory, it is of interest to relate the question of these contour tones
to the previous question of tonal features. The Nupe examples in which
the contour tone occurs if and only if the intervening consonant is
voiced may seem to support the use of [-stiff, +slack] specification for
both voiced consonants and low tone vowels. The copy rule would be as
given in (33).

(33)  [+stiff]  \[
       \begin{array}{c}
       \text{[-stiff]} \\
       \text{[-slack]} \\
       \text{[+stiff]} \\
       \text{[+slack]}
       \end{array}
\]

The 'tone copy' would not occur after a [+stiff, -slack] (e.g.
voiceless) consonant. But since according to Halle's new features a
voiced consonant has the same 'tonal' specifications as a low tone vowel,
one might expect that a high tone vowel followed by a voiced consonant
would phonetically be realized as a high-low fall on the preceding vowel,
or a low to high or mid tone glide on the following vowel, no matter
what the tone of the preceding vowel is. This, of course, is not what occurs.
Furthermore, in Yoruba, Fresco [1970] and Courtenay [1966] show that the
gliding rule occurs when the intervening consonant is voiceless as well
as voiced, as shown in (34).

(34) Yoruba:  

a. ọta \[
       \begin{array}{c}
       \text{[-]} \\
       \text{[\_\_]} \\
       \text{'enemy'}
       \end{array}
\]

b. ọwọ \[
       \begin{array}{c}
       \text{[-]} \\
       \text{[\_\_]} \\
       \text{'friend'}
       \end{array}
\]

c. ọgbá \[
       \begin{array}{c}
       \text{[-]} \\
       \text{[\_\_]} \\
       \text{'senior sibling'}
       \end{array}
\]

d. ọṣọ \[
       \begin{array}{c}
       \text{[-]} \\
       \text{[\_\_]} \\
       \text{'nine'}
       \end{array}
\]

On the systematic phonetic level, without a contour tonal feature,
how is ọtú \[
       \begin{array}{c}
       \text{[-]} \\
       \text{[-]} \\
       \end{array}
\] in Nupe to be distinguished from ọtá \[
       \begin{array}{c}
       \text{[-]} \\
       \text{[\_\_]} \\
       \end{array}
\] in Yoruba?

It seems quite obvious that the gliding pitch which occurs in some
tonal sequences but not others is not predictable in the way that formant
transitions are. The example given in (32) shows that even when two
different tones occur on adjacent vowels level step tones are possible.
Furthermore, even if some gliding occurs between two tones on adjacent
vowels, on the phonetic level we must be able to distinguish between a
contour tone on a single short vowel and a level tone.
If we do require some contour tonal feature for phonetic specification one may then question the hypothesis which would constrain all phonemic representations of contour tones to sequences of level tones. It seems highly unlikely that a phonetic contrast will never occur as a phonological contrast. As pointed out by Margaret Langdon [personal communication] given a phonetic contrast one can assume that historically such a contrast will or can become restructured as an underlying phonemic contrast. This is not too different from the suggestion that historically all nasal vowels derive from a sequence of vowel + nasal. To suggest that nasal vowels never occur at the systematic phonemic level contradicts synchronic facts [cf. Hyman, forthcoming].

What is being suggested here is that the universal set of features must include a feature or feature combination which will distinguish between contour tones and level tones. The particular feature(s) are discussed below, as is the question of deriving all contour tones from sequences of level tones, since the answer to the first question is dependent on the latter.

4. Segmental vs. Suprasegmental Representation of Tone

Woo's proposal [1970] that "prosodic features are segmental rather than suprasegmental" was based on the claim that all contour tones occur only when there are long vowels in the underlying forms. Maddieson [1971] points out that "both Longacre [1952] and Spears [1968] talk of syllables on which three tone phonemes or pitches appear without any mention of lengthening of the vowel. These cases include syllables of the structure CV in which C is voiceless." [p. 15] The Ibadan conference concluded "that a contrast must be made between a sequence of two (or more) 'tones' and a sequence of two (or more) pitches which form a single 'tone'. This may be done by marking one or more of the segments bearing pitch as [-syllabic]." [p. 81] They further support this conclusion by stating that this "implies that one expands the class of 'glides' and forgets the myth that non-syllabic vowels necessarily have closer tongue positions than syllabic ones." At the conference Williamson pointed out that
"in Igbo, in a sequence, close vowel + open vowel which are on different tones, both remain syllabic, whereas in the same sequence on identical tones the close vowel becomes non-syllabic." [p. 81]

Leben [1971] also shows that at least in Hausa and Mende short vowels must be specified phonologically as a High-Low sequence. The existence of 'contour' tones on short vowels (revealed by many linguists over the years, cf. e.g. Welmers, Pike, Wang, Longacre, etc.) convinced Leben that Woo's hypothesis constraining contour tones to long vowels must be abandoned. Because he and Halle [1971] wish to maintain Woo's other hypothesis, i.e. that there are no underlying contour tones, they propose that "the theory be modified so as to allow prosodic phenomena to be treated also as suprasegmental phenomena." In an attempt to justify this position Leben criticizes the hypothetical feature [+ [+High] followed by [-High] ] (to represent a sequence realized as a falling tone) suggesting that "this would render Woo's hypothesis vacuous: it would permit the representation on a single segment of any sequence of level tones, regardless of whether its syllable contained a short vowel, a long vowel, or whatever. In this case the claim that contour tones are underlying sequences of level tones would become nearly empty, since the representation [+ [+high] followed by [-high] ] on a segment is empirically equivalent to the representation [+falling]." [p. 14-15] If in the suprasegmental matrix posited by Leben and Halle, a sequence of [+high] [-high] is later to be mapped onto a single segment (how, is never made clear), it seems to me that this is also "empirically equivalent to the representation [+falling]." Leben further suggests that "it is not clear how the rule of tone copying could be prevented from incorrectly copying the feature [+ [+high] followed by [-high] ] instead of copying the feature [-high]." It seems to be the case that in progressive 'tone copy' the final feature value is copied, whereas in regressive 'tone copy' rules the initial tonal feature is copied. By a universal convention this can certainly be specified. The rule itself can make this clear. Leben cites some examples from Mende compounds in which the following Compound Rule is posited:
(35) [Leben, 1971, p. 188]

(a) Copy the last tone of the first member of the compound onto the first syllable of the second member.

(b) Assign a low tone to the remaining syllables of the second member.

This rule accounts for the following examples:

(36) a. pélɛ + hanl = pélɛ -hánl
    b. bɛlɛ + hanl = bɛlɛ -hánl
    c. mbû + hanl = mbû -hánl

Using the feature he rejects, one can write (35a) as:

(37) [tone] --> [atone] / [(β tone followed by) atone] ___

He states further that "the falling tone on mbû is converted into a high tone by tone deletion." This can be accomplished by a tone simplification rule as well as a tone deletion rule as in (38).

(38) [[+high] followed by [-high]] --> [+high]

There seems to be no formal reason why a feature such as [[+high] followed by [-high]] cannot be utilized to specify contour tones in underlying representation. It would create fewer problems than the proposal to include two matrices, one segmental and the other suprasegmental, for each surface structure, as I will attempt to show below.

One can however utilize a [-segmental] but tone bearing unit which could accomplish the same thing (as was proposed by Schachter and Fromkin [1968]) or adopt the Ibadan Conference proposal utilizing tone-bearing non-syllabic vowels. Thus, in Mende, the segmental matrices given under (39i) and (39ii) are equivalent to the suprasegmental matrices given under (39iii).

(39)       i.  ii.  iii.

a. mbû  'owl' / mbû' / / mbû' / / [+h] [-h]/
    b. mbû  'rice' / mbû' / / mbû' / / [-h] [+h]/
    c. mbû  'companion' / mbû' / / mbû' / / [-h] [+h] [-h]/

(The V' represents a non-syllabic vowel.)
The number of tonal units, without segmental features, or the number of possible syllabic vowels in a string, can be restricted by morpheme structure conditions, and universally, there can be a convention which transfers the tones of the non-segmental units or non-syllabic vowels to the preceding segment. Once this is done of course, the Contour feature will also have to be added.

Leben, in keeping with the basic goals of generative phonology, is desirous of including the strongest possible constraints in the theory, thereby limiting the class of possible grammars. If, then, one solution restricts the kind of rules which can be utilized in a grammar, and if this constraint is supported by empirical evidence, this constraint should be incorporated in the general theory. He suggests that this is the case if tone is represented suprasegmentally rather than in segmental matrices, and attributes his argument to Wang [1967]. Leben states: "If tone is a feature on some entity more abstract than the segment -- such as the syllable or the morpheme -- then it is impossible to state a rule which changes the tone in an environment determined by the segments below it. This shows that the assumption that tone is expressible phonologically can serve to limit the class of phonological grammars defined by the theory in a way which cannot be done with underlying segmental features of tone." [p. 198] But this is a specious argument since Leben himself points out that "the status which the theory must give to Mandarin and Thai is that the point in the derivation at which tone is initially expressed as a segmental feature is the beginning of the derivation -- i.e. that for such languages there is no stage in the derivation at which tone is a suprasegmental feature." [p. 199] Thus, tone both is and is not a suprasegmental feature. And nowhere does either he nor Halle explain how the sequence of tones which phonetically are realized as contour tones is to be specified once such sequences are mapped onto single segments. He is correct of course in stating that this would limit the class of languages given that he is correct in stating that "if a language does not meet the narrow constraints imposed by Woo's framework, then it is subject to the constraint imposed by suprasegmental
representation: rules like (a low-tone raising rule) must be ordered after all those rules in which tone is expressed as a suprasegmental feature." [p. 199] But in the Nupe examples given above, it is clear that the tone-copying rule must depend on segmental feature specification. Thus, according to Leben, it should occur after suprasegmental information is translated into segmental information. This requires a rule assigning two tones to a single segment early in the derivation. How can this be done without a 'contour' feature or without the use of [-segmental] or [-vocalic] tone bearing units? If such a feature or such a method of representation is needed at some point in the derivation why not use it for the lexical representation? The example Leben cites from Hausa similarly involves a tone change depending on vowel length (i.e. segmental information). He does not consider the Hausa case a counter example because in this case he says that the suprasegmentals must have already been applied to the segments.

If every such counter example involves a similar early mapping, then I fail to see how the class of languages is constrained. And to repeat myself, if two tones must be applied to one segment there must be some way of doing this which can conceivably be done from the start.

Not only are there tone rules which must include segmental feature information but tone rules must also include syntactic information present in the surface structures of the string feeding into the phonological component.

In the Akan examples given above, the Vowel Deletion rule deletes all the segmental features as well as tonal features. By Leben's proposal, then, the mapping of suprasegmental matrices onto segmental matrices must occur before the Vowel Deletion rule. The Vowel Deletion rule must also occur after the Pitch Assignment and Down Drift rule. One may conclude that the mapping of suprasegmental onto segmental matrices occurs at the beginning of the derivation. The problem is complicated, however, because there are certain cases where tone is the only realization of grammatical morphemes.
In (40) Ø represents the Habitual Low tone morpheme.

(40) ɔ bɔ Ø Ghana [ɔbɔ Ghana] 'He comes to Ghana (habitually)'

The 'drop' tone on the first syllable of Ghana may be compared to the tone on this syllable in other contexts, as shown in (41).

(41) a. ɔdwaa wɔ Ghana [ɔdwaa wɔ Ghana] 'Odwaa is (located in) Ghana'

b. ɔkrɔn nɛ Ghana aheŋkuro [ɔkrɔn nɛ Ghana aheŋkuro] 'Accra is the capital of Ghana'

To derive the 'downstep' in (40), the pitch assignment and downdrift rules must apply prior to the deletion of the 'non-segmental' low tone morpheme, as shown in (42).

(42) ɔbɔ Ø Ghana
PA 3 1 3 1 3
DD 3 1 4 2 5
Tone Del 3 1 2 5

Since the tone is deleted in Vowel Deletion rules or when a tone occurs without any segmental features it is necessary to map (if one wishes to take this road) the suprasegmental matrices onto the segmental prior to these deletion rules, which rules must occur after the other tonal rules (e.g. PA and DD). This being the case there must be some way to represent a non-segmental tone in the segmental matrices. If we utilize the proposal of Schachter and Fromkin, that is, represent such units as [-segmental, +tone], there is no difficulty. Contour tones on short vowels may be similarly represented. All units may then be specified as [+segmental], with [-segmental] further divided into [+boundary]. [-segmental, -boundary] units will be tone bearing units
with no segmental features specified. I believe this is a better solution than the proposal by the Ibadan Conference which would represent such units as [-syllabic, +tone] vowels, since in the case of grammatical morphemes such as the above, it would be stretching a point to specify this vowel as the least marked vowel /a/ which, as in the case of the Akan Habitual, is never realized on the surface. If it was possible to avoid the mapping of the suprasegmental matrices onto the segmental matrices then there might be some point in including two distinct matrices. But since the examples from Hausa, Nupe, Akan, etc. force us to treat the tonal features segmentally, interacting with segmental features, and requiring syntactic information for the correct application of tone rules, I cannot see that the new proposal has any merit. In fact, since the motivation for the new Halle/Stevens features was to use identical features for vowel tone and consonantal glottal stricture, the proposal to then separate these features into two matrices seems rather unwarranted.

A contour tone in an underlying matrix may then be represented as in (43):

(43) a. \[ \begin{array}{c|c} \text{+h} & \text{+seg} \\ \hline \text{-h} & \text{-seg} \end{array} \] = falling tone \(\wedge\)

b. \[ \begin{array}{c|c} \text{-h} & \text{+seg} \\ \hline \text{+h} & \text{-seg} \end{array} \] = rising tone \(\vee\)

c. \[ \begin{array}{c|c|c} \text{-h} & \text{+h} & \text{-h} \\ \hline \text{+seg} & \text{-seg} & \text{-seg} \end{array} \] = convex tone \(\wedge\wedge\)

By a universal convention any sequence of two or more tone bearing units, in which the first is [+segmental] and the second (and third) is [-segmental] will be further specified by adding the feature [+contour] to all such units. In other words this kind of representation can be the formal way of specifying contour tones. In a language such as Yoruba, the 'gliding' rule can be stated as (44).

(44) \[ \text{[tone]} \rightarrow \text{[atone]} \begin{array}{c} \text{[-atone]} \\ \hline \text{[-seg]} \end{array} / \text{[atone]} \]

where, if \(\alpha = [+H]\), - \(\alpha = [+L]\)

and, if \(\alpha = [+L]\), - \(\alpha = [+H]\)
Once the feature Contour is available, it will also be possible to utilize this feature in languages where tone copy rules do not occur. Schane [personal communication] has observed that since there are tone languages in which only register, level tones occur, but no tone languages in which only contour tones occur, register tones take precedence in the hierarchy; that is, the Contour feature can only be used if there are tones without the contour feature. Furthermore, a tone represented as [+high, +contour] will then specify a 'rising' tone, one represented as [-high, +contour] or [+low, +contour] will specify a 'falling' tone, and one represented as [+high, -high, +contour] will specify a 'convex' tone. By a universal convention, a progressive tone copy rule will copy the specified [+high] of a preceding contour tone, whereas a regressive tone copy rule will copy the opposite value of the following contour tone. The convention can be stated as in (45).

\[(45) \begin{align*}
\text{Tone} & \rightarrow \begin{bmatrix}
[\![\text{aH}]\!] & / & \begin{bmatrix}
[\![\text{aH}]\!] & \text{[+contour]} & -
\end{bmatrix}
\end{align*}
\end{align*}
\]

Using these conventions it is possible to represent all contour tones in the lexicon without resorting to the [-segmental] units. For non-segmental tonal units which represent grammatical morphemes in the surface structure however, the [-segmental, +tone] units may still be necessary.

5. Conclusion

The above discussion on tonal features and tonal rules has attempted to show (a) that the features Stiff Vocal Cords and Slack Vocal Cords for both vowel tones and glottal states are inadequate and should not be adopted; (b) that three features for tones High Mid Low prove to be the best yet proposed (whether High and Low are called Raised and Lowered is immaterial) since the use of the feature Mid permits specification of a mid tone either as 'closer' to High, or to Low or 'intermediate'; (c) that a contour tonal feature is necessary for descriptively adequate phonetic representations; and (d) that the inclusion of suprasegmental
matrices for tonal specification in surface structures does not resolve the problem of poly-tones (or contour tones) on one segment -- it merely places the problem in a different part of the grammar.

REFERENCES


Hyman, L. "Nasals and nasalization in Kwa." (forthcoming)


