Version II
March 1975

English as a Tone Language*

John Goldsmith
Massachusetts Institute of Technology

It is commonly acknowledged that what generative and traditional grammaricans have called "stress" is realized by a combination of pitch, length, and something else we might call stress-in-a-strict-sense. Nonetheless, Chomsky and Halle in Sound Pattern of English (pp 24-26) defend the position that there is a linguistic reality to a more abstract notion of stress, one with three or four levels realized in English, and on this score there has been little disagreement within the post-Sound Pattern literature.[1]

The purpose of this paper is to present an alternative to the Sound Pattern English stress system, an alternative based on two ideas: first, there are separate systems of rules that explain stress in the strict sense, pitch, and length; these factors, furthermore, cannot be deduced from a single multi-valued factor, as suggested in Sound Pattern. Second, the appropriate formal device to deal with pitch in English is autosegmental phonology. This framework was orig inally developed to handle various problems in the phonology of African tone languages, but has subsequently been applied to such diverse "prosodic" phenomena as Sanskrit intonation, Japanese accentuation, nasalization, and vowel harmony.[2]

1. Tone Melodies. We begin with the assumption that the tone melody of a phrase is composed of tonal segments, and following Nancy Woo (1969),

we restrict the inventory of tonal segments to level or static tones, such as High, Mid, or Low, to the exclusion of contour or dynamic tones, such as Rising or Falling. All occurrences of Rising tones are to be analyzed as sequences of Low and High, or an appropriate configuration with Mid.[3]

Given a melody, we may ask how it is to be organized with respect to the words and syllables in the utterance. A language has recourse to basically two principles, I would suggest: the tone-language principle and the accentuation principle.[4]

The tone-language principle was suggested in various forms by Edmundson and Bendor-Samuel, Williams, and discussed at length by Leben (1973). This principle says that within a word -- a unit formally defined by the language on universal principles -- the syllables of the word and the tones of the tone melody get connected up one to one, where for "connect" we may think "is realized on". If there are three tones -- e.g., Mid High Low, in that order -- in the tone melody and three syllables in a word, then the application of that melody to the word will create a word with syllables Mid, High, and Low, in that order. Should there be four syllables, then -- by principles we will look at in detail below -- the word will be realized as Mid High Low Low. That is, the tones and the syllables get assigned to each other one-to-one starting on the left; there was an extra syllable in the second case, so it was assigned to the final Low tone, too. This phenomenon of "spreading" we shall see is not restricted to tone languages proper, but is universal.

This one-to-one correspondence I shall call "the tone-language principle". There is another principle used in assigning a tone-contour to a word or phrase which I shall call "the accentuation principle", which says that when the language has a string of syllables (say, "Ma nong ga he la") and a tone melody (say, Mid High Low), then the language assigns an accent, written with an asterisk "*", on one of the syllables and one of the tones. In this case, for reasons we will get to below, the syllable "he" receives the accent, and so does the

H in the tone melody $M\overset{*}{H}L$. Thus, we set the two strings parallel to one another on the page:

(1) ma nong ga $\overset{*}{\text{he}}$ la

$$M \quad \overset{*}{H} \quad L$$

The accentuation principle says that the two starred items are to be associated with one another, making the syllable "he" High in tone, in this case. The accentuation principle is a rule expressed in (2); applying (2) to (1) gives us (3).

(2)
$$\begin{array}{c} V \\ \\ \overset{*}{T} \end{array} \longrightarrow \begin{array}{c} V \\ | \\ \overset{*}{T} \end{array} \qquad \text{(or)} \qquad \begin{array}{c} \overset{*}{V} \\ \vdots \\ \overset{*}{T} \end{array}$$

(3) ma nong ga $\overset{*}{\text{he}}$ la

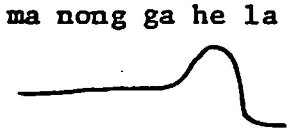
$$M \quad \overset{*}{H} \quad L$$

By some general principles, the L tone must be realized too, and obviously it must come to the right of the syllable "$\overset{*}{\text{he}}$" -- which leaves it with one possibility: to be realized on "la". The M tone likewise spreads over the first three syllables, giving a pitch contour as in (4).[5]

(4) ma nong ga $\overset{*}{\text{he}}$ la

$$M \quad \overset{*}{H} \quad L$$

ma nong ga he la



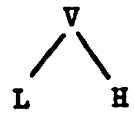
In this short paper, I shall concentrate on the accentuation principle, and on English in particular. Nonetheless, the principles discussed are general. There have been quite a few attempts in the past to assimilate tone languages to what were called accentual systems. One conclusion from the present suggestions, if they be held correct, is that the conventional ideas about accentual systems are off the mark, and that if anything, our ideas about the less exotic languages which we call accentual can be enriched by viewing them in the context of tone languages ~~in general~~ -- thus the title of this paper.

2. The autosegmental system.

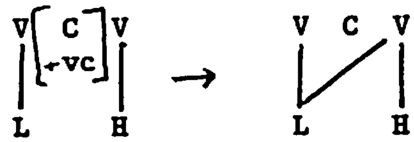
It has been recognized in the past that for tone languages in general, it is incorrect to view the domain of the toneme as the syllable. That is, on the one hand, it was clear that the tonal melody should be viewed as a sequence of level tones. When contour tones appear on single vowels, then, the domain of each level tone making up that contour is less than a single vowel or syllable. Still, there are many examples where the proper analysis shows that tones which are realized on one syllable can spread over neighboring syllables, sometimes when that neighboring syllable has no tone of its own, sometimes even when it does. In this second case, then, the proper domain of the tone is larger than the syllable.[6]

The domain of the toneme can be more, and can be less than a single vowel or syllable. And syllables can be underlyingly toneless. These perplexing facts, and other equally anamolous characteristics of tonal behavior, led to the autosegmental system. Put simply, this theory suggests that when we hear an utterance with a tonal melody, we are not witness to a single string of segments with features for tone; rather, we are observing two simultaneous strings of segments, one consisting of tonal segments, the other consisting of phonological segments in the normal sense. A formal analysis of such a thing would consist of two parallel strings of segments, organized with respect to each other in time by "association lines." A vowel uttered at the same time as the High part of the melody is realized, of course, as a High vowel (a vowel high in tone).

Viewed in this light, there is nothing paradoxical about finding two tones in the course of uttering one vowel -- that is, finding a contour tone. This is represented as in (5); it states, in effect, that the transition from the L toneme to the H toneme occurs "during" the vowel.

(5) 

Or we may consider a slightly more complex example. In Nupe, for example, we find a very common type of "tonal flop" rule; see George (1970) for details. A High-toned vowel preceded by a Low tone becomes a Rising tone if the intervening consonant is voiced -- clearly an "assimilatory" process (we find virtually the same rule in classical Sanskrit without the condition on the intervening consonant; see May and Goldsmith for details). This rule would be written as in (6), or using the notational oonvention that a dotted line represents a structural change, as in (7).

(6) 

(7) 

In general, the strength of a theory like this one is to be judged on three types of grounds: its ability to describe simply whole realms of phenomena that the standard theory cannot; constraints on possible representations within the theory; and constraints on possible tonological rules -- that is, on rules which adjust the relation between the two lines of segments.

We shall see some evidence of the first sort in this paper when we observe contour tones on short vowels in English, a possibility not countenanced within standard theory. More interestingly, we shall dispense with stress subordination as a general principle, and eliminate the non-binary specifications of stress in English.

The second requirement -- constraints on possible representations -- is provided by a well-formedness condition, discussed in detail below. For better or worse, the tonological rules in English appear at this point to be very simple, so we will see no opportunities to develop interesting constraints on rules here. For more interesting results along these lines, see Goldsmith (1974,1975a,b), May and Goldsmith, Haraguchi.

3. English: Placement of Stress and the Star.

Now we are in a position to reanalyze the stress patterns of English words. Our first task is to review the system that assigns the tone melody to English words said in isolation. This point must be emphasized: the tonal contours which we derive, just like the stress subordination facts that Chomsky and Halle produce, are realized only when the word or phrase is said in isolation with neutral intonation, as when one is reading the word at the beginning of a dictionary entry.

The role of a binary stress distinction in determining vowel quality is firmly established in <u>Sound Pattern</u>, and I shall briefly review the relevant results.

I am not concerned here with the specific proposals for dealing with the distribution of the binary feature <u>stress</u>. A small number of rules assign +stress, and a small number of rules may change these +stress to -stress.

The stress, or lack of it, on a vowel is clearly indicated by whether it undergoes vowel shift (if the vowel is tense) or whether it reduces to schwa (if the vowel is lax). In short,[#7]

	stress	-stress
tense	undergoes vowel shift	no change
lax	no change	reduces to schwa

A pair of words with the same vowel stressed in one case and unstressed in the other, but retaining the same specification for tense/lax may be found:

tense /i /	unstressed	var*i*ous	[i]	no vowel shift
	stressed	var*i*ety	[ay]	vowel shift
lax /æ/	unstressed	*a*tomic	[ə]	reduced to ə
	stressed	*a*tom	[æ]	unreduced æ

For any word in English, a particular syllable may be located and called the accented syllable. In the Sound Pattern framework, this is the sole syllable marked "1stress". Because of the notation we use here, I shall often refer to this syllable as the "starred" syllable, for we shall assign an abstract element, the "star", to that syllable.

We use essentially a suggestion made in Halle (1973) and attributed there to Schane.

(10) "Detail" Rule -- Star Assignment

$$V \rightarrow \overset{*}{V} \;\; / \;\; \_\_\; Q \; (\; VC_0 \; (+y))\#\# \qquad \text{Condition: Q contains no } [+\text{stress}]$$

This rule picks out what was called the "1stress" by putting a * on it.

(11) $\begin{bmatrix} \text{æ} \\ +\text{str} \end{bmatrix} \text{nt} \begin{bmatrix} \text{i} \\ +\text{str} \end{bmatrix} \text{s} \begin{bmatrix} \text{i} \\ -\text{str} \end{bmatrix} \text{p} \begin{bmatrix} \text{æ} \\ +\text{str} \end{bmatrix} \text{t}$ Output of stress rules

$\begin{bmatrix} \text{æ} \\ +\text{str} \end{bmatrix} \text{nt} \begin{bmatrix} \overset{*}{\text{i}} \\ +\text{str} \end{bmatrix} \text{s} \begin{bmatrix} \text{i} \\ -\text{str} \end{bmatrix} \text{p} \begin{bmatrix} \text{æ} \\ +\text{str} \end{bmatrix} \text{t}$ Rule (10)

$* \; - \quad Q \quad - \; V \quad C_0$

"ant$\overset{*}{\text{i}}$cipate"

4. Deriving the Melody.

Now the question can finally arise: what do we do with the star once we have assigned it?

The answer is simple. The isolation tone melody for a phrase said with neutral intonation happens to be MHL with the * on the H: $M\overset{*}{H}L$. These three letters stand for Mid, High, and Low.[9] At this point we have a phonological string with a * assigned somewhere on it, but no tones; similarly, we have a tone melody but we have done nothing with it, so to speak. So we put the two together.

This is done with the mediation of the star. Set the two strings, phonological and tonological, parallel to each other, and then draw a line connecting the starred elements in either string -- that is, apply rule (2), deriving (12).

(12)

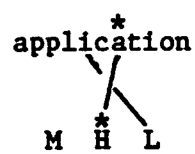
```
     *
applic ation
      |
      *
    M H L
```

What this means -- so far -- is that the High tone is being realized on the third syllable. Nothing has been said either about the tone of the other syllables nor about what happens to the Mid and Low tonemes, which have not yet been associated with any syllables.

At this point two universal principles step in. (1) All syllabics must be associated with at least one tone, it says, and all tonemes must be assigned at least one syllabic segment. A further condition -- an obvious one, perhaps -- is that (2) in setting up these association lines, no lines can ever cross. Crossing of lines is absolutely forbidden. I shall refer to these two principles together as the Well-formedness Condition.

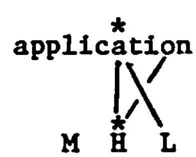
At this point, the structure (12) does not meet the Well-formedness Condition. However, violation of this condition is not to be taken as throwing the derivation out. Rather, violation means that the structure is adjusted automatically to maximally meet the Well-formedness Condition by addition or deletion of a minimal number of association lines. Such a declaration may lead to ambiguity, in that several different ways of adjusting the structure, each requiring the same number of deletions or additions, may be possible. Further principles must then be found to uniquely determine the resulting structure.

Let us see what the Well-formedness Condition tells us to do with the structure we have so far for a word like "application". The toneme L must be associated with at least one syllable. We have only a small number of choices as to which syllable the L will attach to: (a) only the last syllable, "tion"; or (b) the last and the penult syllable. We see that connecting the L with the first or second syllable is prohibited, because to do this, the association lines connecting the H with the third syllable would be crossed. Cf (13).

(13)


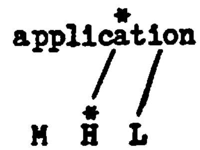
FORBIDDEN

We can see similarly that the final syllable can be associated with only the L or the H, or both, without the association lines crossing. Furthermore, we can't both associate L with "a" and the H with "tion", although, we could do one of these:

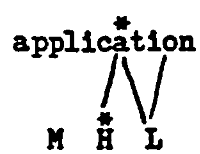
(14)


FORBIDDEN

From this it can be concluded that L and "tion" must be associated with each other, or else one of the two won't be associated with anything. So from (12) we can certainly see that the Well-formedness Condition produces (15).

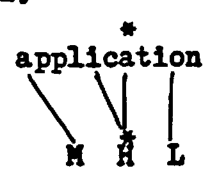
(15) 

But now both L and the final syllable, "tion", meet the Well-formedness Condition: that is, they are both associated with something. So, while it is logically possible to associate the L with both the last and the penultimate syllable, as in (16), this is not the <u>minimal</u> way to meet the Well-formedness Condition. To get (15) from (12) requires only one additional association; (16) requires two. Therefore (15) is derived by the Well-formedness Condition.

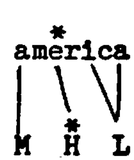
(16) 

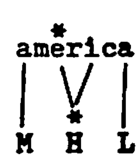
The principles we have specified so far leave it an open question whether (17) or (18) is the ultimately correct well-formed result of structure (12). Both (17) and (18) represent the addition of three association lines, and on purely numerical grounds, neither is preferred.

(17) 

(18) 

The same indeterminacy arises with the derived form of a word like "America". Is (19) to be derived, or (20)? -- both represent three added association lines, but are empirically different.

(19) 

(20) 

In the case of an ambiguity like this -- that is, a competition between a starred toneme and an unstarred toneme for spreading through the Well-formedness Condition -- it is always the unstarred toneme that dominates. In these cases, that means that (17) and (19) are derived, and (18) and (20) are not. 48

Let us consider the more complex case of words with final or initial "main stress" -- i.e., with final or initial star.

When a word has final star, as in Kalamazoo, balloon, Tennessee, and so on, the initial star-to-star accentuation association givus us (14).

(14) (14) Ka la ma $\overset{*}{\text{zoo}}$

M $\overset{*}{\text{H}}$ L

The Well-formedness Condition corrects this to (15).

(15) Ka la ma $\overset{*}{\text{zoo}}$

M $\overset{*}{\text{H}}$ L

(15) contains a representation with two tonemes associated with the same syllable, which is, of course, perfectly permissible. It signifies a falling tone from High to Low. And that is precisely the fact in English: words with "final stress" realize a marked falling tone on the ultima, whereas words with non-final main-stress realize that stress as a High pitch, an observation not captured within standard generative treatments of stress.

Words with initial main-stress, like Wittgenstein, similar, syllable, and so on, will have both the M and the H toneme associated with the first syllable, as in (23).

(23) $\overset{*}{\text{W}}$ittgenstein

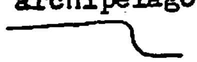
M $\overset{*}{\text{H}}$ L

There is, however, no particular rising tone on the first syllable of such words, as this predicts. Two possibilities present themselves:

either the tone melody for English neutral intonation is really $\overset{*}{H}L$, not $M\overset{*}{H}L$; or there is a rule deleting M when it occurs on the same syllable as H.

A proponent of the first position would point out that, in fact, the intonation of a word like archipelago can be either as in (~~24~~ 24a) or (~~25~~ 24b), where (~~24~~ 24a) is that predicted by the melody $\overset{*}{H}L$. Such a proponent might further propose an optional late rule raising the pitch of the starred syllable, giving the optional variant (~~25~~ 24b).

(~~24~~ 24a) archip*elago



(~~25~~ 24b) archip*elago

The defender of the melody $M\overset{*}{H}L$ would be forced to posit a MH-Simplification rule, as in (~~26~~ 25).

(~~26~~ 25)

```
  V          V                 V
 / \   →     |        or      / \
M   H        H               M   H
                             1   2
                             ∅   2
```

(Tonological transformations factor only the tonological string; the non-tonological elements that enter into such a rule, as in (~~26~~ 25), are, strictly, conditions on the rule rather than part of the structural description of the rule). I shall leave the question open regarding which of the two solutions sketched above is correct for the initial Mid, for it does not affect any of the other arguments in this paper.

5. Farewell to Stress Subordination.

What has been left out of this whole discussion is any reference to "stress subordination" rules -- those of the form (~~27~~ 26), which by convention would lower to 2stress, and then 3stress, the other syllables marked +stress in the word (see Chomsky and Halle , Halle and Keyser, Halle).

(26) [1stress] → [1stress] / X — Y

This absence of stress subordination is, of course, no oversight; the claim being made here is that the segmental features "stress" is everywhere only binary; and that the more variable facts observed in phenomena called "stress subordination" are mostly pitch, and should be treated as such in a system capable of accounting for suprasegmentals.

To my knowledge, the surface pitch facts were never explicitly predicted in the Sound Pattern system, but to the extent that 1stress was associated with a rise in pitch, one might attempt to extend the interpretation so that 2stress meant a slightly lower rise in pitch, 3stress slightly lower again, and so on. Such an interpretation -- admittedly not found in SPE -- would predict a rise in pitch on the final syllable of Wittgenstein, which we observe is false. The 3stress there is Low, as Low as the second syllable "gen", which is unstressed.

Worse yet for the stress-subordination system, a 3stress preceding the 1stress is uniformly High in pitch, as in the first syllable of archipelago in (24) or (25), but no higher or lower than the unstressed syllable "chi" which follows. All these facts, of course, proceed straightforwardly from the approach being developed here.

6. Noun compounds.

Let us now consider the stress and pitch contours of noun compounds, phenomena central to the Sound Pattern analysis of stress subordination. Compound nouns like the following -- as everyone knows -- show that the first of the pair of words forming a compound maintains the "main stress , or star.

(27) black-board, Pentagon-papers, Sound-pattern, stress-subordination, tone-melody

Each of these words going into making up a compound would receive its tonal pattern, if it had been said in isolation, in accordance with the principles sketbhed above. Every word has a star (*) placed on it somewhere, and that would distribute the tone melody over it.

Compounds, however, by nature have two stars, since each word going into making them up has a star assigned to it from the previous, individual word cycle:

(28) ## bl$\overset{*}{a}$ck # b$\overset{*}{o}$ard ##

Black-board, being a lexical item, is assigned the tonal melody M$\overset{*}{H}$L. The only question is how the morphological component reduces the number of stars in black-board to one. The answer is simple: the principle is that the star on the left wins out. Formally, we have rule (29):

$$(29) \quad * \rightarrow \emptyset \ / \ \#\# * X \text{ -- } \#\#$$

That is, when two stars occur in a lexical item, rule (29) says that the rightmost of the two is deleted. This leaves one star for the tonal association procedure to address when the M$\overset{*}{H}$L contour -- or whatever contour -- is assigned. In short, rule (29), is our "compound stress rule", though it has nothing to do with stress, strictly speaking. Compound words are just like any other word tonally, on our analysis; they are composed of two elements separated by a #, just like govern#ment; the only difference is the status of the two underlying parts. In compounds, by definition, each part could stand alone as a word.

In the Sound Pattern description of a compound like $\overset{1}{\text{black-}}\overset{3}{\text{board}}$, the two syllables were described as having 1-3 stress. But what does this mean? What is the significance of the primary stress assigned to "black" here? Does it have the same significance as the "1stress" assigned to "board" in "black board", the adjective-noun sequence? One would expect it should; and yet this is not the case, as we shall see. The distinction between the compound black-board and the sequence black board arises not out of a difference in stress assignment -- for both syllables are stressed in both expressions -- but rather a difference of pitch and of length.

Pitch: On the basis of (29) above, the star on blackboard when produced with neutral intonation should, according to the theory, produce a H-toned first syllable and a Low-toned second syllable, as in (30) -- which is correct. In the sequence black board (when uttered naturally as one intonational phrase) the accent rests on board[10]; thus the tone is correctly predicted to be Falling on the word board (a concatenation of High and Low). ~~See (31).~~

(30) bl*ack#board
H* L

(31) black##bo*ard
M H* L

Length: There is also a clear distinction in the length of the syllables in (30) and (31). Most noticeably, the syllable "black" is much longer in (31) than it is in (30). This difference is indirectly observed in the Sound Pattern notation that gives the longer form a "2stress" and the shorter form a "1stress"; but then, not all "1stressed" syllables are short in this way. A monosyllabic word -- "black", for example, uttered in isolation -- would be "1stressed" but not as short as it is in (30). The extent to which the Sound Pattern notation is even observationally adequate becomes unclear, for the very raw facts are inconsistently described in this notation. In sum, the notational system which expresses all these distinctions with 1,2,3,4 stress confuses the issues by mixing apples and oranges. Pitch and length form coherent linguistic subsystems but are not to be identified as levels ~~as~~ of stress. The development of the correct system for accounting for pitch patterns is the primary subject of this paper; let us digress a bit on the matter of the syllable lengthening process which helps distinguish (30) and (31).

We suggested a structure for compounds above in (28). In order to justify this structure (## X # Y ##) rather than the alternative structure (## X ## Y ##) suggested in <u>Sound Pattern</u>, we will pursue the process that lengthens the duration of monosyllabic words in English.

Consider three sets of facts from English.

(A) Prepositions can either be clitics or non-clitics; in the second case, they are generally called "particles", and do not take a complement or object. A minimal pair[11]:

(32) (a) This is the hat John sat on.

(b) This is the hat John tried on.

In (32a), where the preposition is a clitic, it is much shorter in duration than in (32b), where it is not a clitic. In (32b), it is a particle, with the structure $[_{PP} [_P \text{ on}]]$.

Or compare (33a) with (33b).

(33) (a) His constant enthusiasm turned on his students. ('got them excited').

(b) The ferocious lion turned on his trainer. ('became disobedient')

In (33a), the preposition is not a clitic with the object "his students"; it consequently is longer in duration than in (33b).

(B) Compare the duration of the following syllabes when they are independent words, and when they are found in a polysyllabic words:

(34) tie, tyrant; calm, comic; limb, limber.

When the monosyllable forms its own word, it is significantly longer than when it forms part of another word.

(C) We find the same difference in length of the word <u>black</u> in the pair <u>black-board</u> (to write on) and <u>black board</u> (board which is black). Its length in the compound is markedly shorter than its form in the adjective-noun sequence.

(D) This same long form ofthe adjective is found when the adjective is alone in the predicate, as in "The board is black".

There appear to be basically two ways we could try to account for the conditions of Monosyllabic Lengthening, a process applying to non-clitic prepositions, and to monosyllabic nouns and adjectives found outside of compounds. Either we could see the rule as phonologically conditioned -- that is, conditioned by phonological boundaries of some sort around the form -- or we could try to find some category or set of aategories which encompassed the set of forms undergoing the rule of lengthening. I shall argue that the second hypothesis ("the categorical hypothesis") is not correct, but rather that the first is, and that the rule is essentially (35).

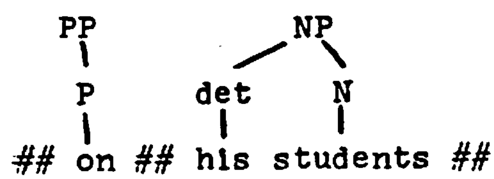
(35) Monosyllabic Lengthening

$V \rightarrow \hat{V}$ (i.e., is lengthened) / ## C_0 -- C_0 ##

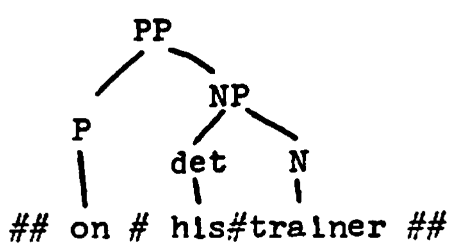
Now, the fact that the same category of words -- prepositionxs-- both does and doesn't undergo the lengthening, depending on whether it takes an object, is suggestive. On the categorical hypothesis, lengthening would not be said to apply to the category P = preposition, since prepositions with an object, as in in 33b, do not undergo lengthening. Rather, the rule would have to be tha monosyllabic PP = prepositional phrase becomes lengthened. Otherwise, "on" would undergo lengthening in both 33a and 33b, whereas it actually should only in 33b.

The relevant structure of 33a and 33b, we are assuming, in line with recent work, is as in 36a and 36b, respectively.

(36a)



(36b)



We see, now, that (35), according to which Monosyllabic Lengthening is conditioned by phonological boundaries, makes the correct predictions for (36a) and (36b); it applies only in (36a). A categorical solution, (37), would work equally well.

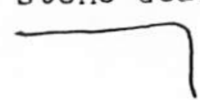
(37) $V \rightarrow \hat{V}$ (i.e., is lengthened) / $[_{PP}\ C_0$ -- $C_0\ ]$

However, (35) accounts immediately for the data in (B) as well, while (37) does not; (35), futthermore, accounts for the data in (D). We conclude that (35) is correct rather than (37). The facts in (C) will fall naturally under this generalization, too, if the structure of a compound is [redacted] ## X # Y ##, but not if there are two #'x between the elements. We conclude, therefore, that there is only one, as illustrated in (28) and (30) above.[12]

This conclusion fits well with the Lexicalist Hypothesis (see Chomsky (1967)). Compound nouns, being produced by the lexicon, will tend to have idiosyncratic meanings. Combining the Lexicalist Hypothesis with the present analysis, a principled explanation of the difference in tone is arrived at for a minimal pair such as <u>tone-deaf</u> and <u>stone deaf</u>.

(38) (a) tone-deaf

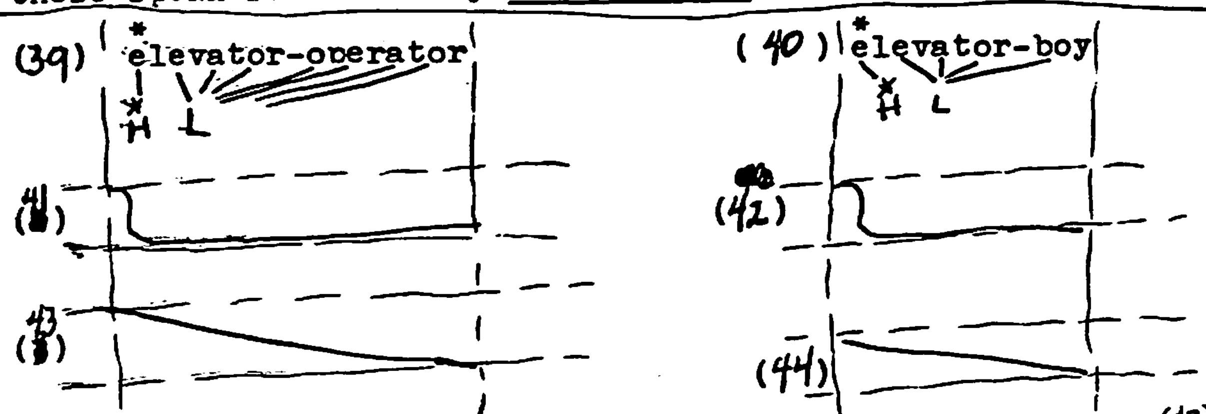
(b) stone deaf



Someone who is stone deaf is indeed deaf; "deaf" is here the familiar adjective, now modified by the word "stone". Someone who is tone-deaf, however, isn't deaf at all; he just can't hear music well, or perhaps the difference between (38a) and (38b). Consequently, "tone-deaf" is a compound, and the compound tone-assignment is expected. Thus (38a) and (38b) are predicted, correctly, to parallel (30) and (31), respectively.

Another problem that has remained with various systems of stress-subordination is the apparent difference between compounds like <u>elevator-operator</u>, which has been held to have a "1-2" stress pattern, and <u>elevator-boy</u>, which has been held to have a "1-3" pattern. (See *Sound Pattern* pp 43, 109) That is, the second element in <u>elevator-operator</u> is thought to have more stress than the second element of <u>elevator-boy</u>. The cause has been ascribed to the monosyllabicity of "boy", in contrast to the polysyllabicity of "operator".

Within the framework sketched here, there is no 2/3 stress distinction, and yet the difference between <u>elevator-operator</u> and <u>elevator-boy</u> is most easily understood and explained in the light of the present discussion. The tone pattern assigned to <u>elevator-operator</u> is as in (39); this corresponds to a pitch pattern as in (41). Now, in fact, many speakers of English have sluggish transitions from High to Low tone; the pitch of this compound then is more like (43) than (41). By contrast, these speakers would say <u>elevator-boy</u> as in (44), derived from (42).



The difference is striking: the main-stress of <u>operator</u> in (43) is at a considerably higher pitch relative to the final syllables of <u>operator</u> ("perator"); in (44), however, "boy" is pronounced at the bottom of the pitch contour, giving the ippression of lower relative prominence. In short, the operator/boy distinction is an optional, low-level pitch distinction, not a stress distinction.

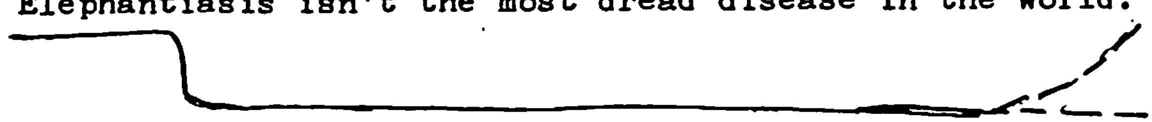
7 More Tone Melodies

The reader may have been wondering all along where the melody $\overset{*}{M}HL$ came from; it is, however, no deus ex machina. In fact, it is just an observation concerning the tonal contour of the neutral intonation contour in English, and applies not just to words or compounds, but to phrases in general. It is difficult, however, to say a meaningful sentences in a "neutral" way -- it is not even clear that such a notion makes sense.

However, we shall see that there are many tonal patterns in English; and in so doing, we shall set to rest the possibility that the star "*" is unnecessary on the tone formula level. Whereas the preceding discussion has had many examples of rules affecting the location of the staf on the syllabic level, the star rested tranquilly on the "H" in the tonal level. On the basis of the examples so far, one might suppose that the star in the upper level is just associated with the "H" on the lower level always. We may quickly dispose of this, for it is only an artifact of the limited kinds of examples considered up to this point.

(1) Consider an indignant response (45).

(45) Elephant$\overset{*}{\text{i}}$asis isn't the most dread disease in the world.



What we find here is that the steep drop in pitch occurs precisely on the accented syllable; this would be characterized by the principles defined here with the formula $H\overset{*}{L}$. This means that the L tone falls on the accented syllable -- <u>ti</u> -- and the preceding H tone spreads over the preceding syllables ("elephant--") through the well-formedness condition.

The complete analysis of a sentence like ( ) requires two further elements: an explanation of star-placement in full sentences; and an explanation of the optional rise at the end of such indignant responses. Both questions are interesting and important, but I shall not discuss them at any length here. The first requires a serious analysis of discourse semantics; and as for the second, I shall provisionally adopt a suggestion of Ivan Sag that such boundary-conditioned rises are triggered not by tone segments but by the phrase boundary itself; this special "rising" boundary I shall indicate by "?". There are a number of interesting consequences of this approach, but I shall not pursue them here.

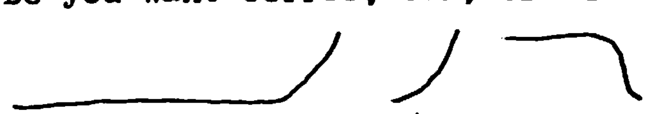
In short, the underlying form is as in (46), yielding (45) by the autosegmental principles discussed above.

(46) %Elephant*iasis isn't the most dread disease in the world.?

%   H $\overset{*}{L}$?

(ii) Disjunctive questions are the sort of questions where the asker assumes that the person asked want one of the choices: "Do you want coffee, tea, or milk?" asked in such a way that the person asked feels it incumbent on him or her to take on of the three. This could be sketched as in (47).

(47) Do you want coffee, tea, or milk?



This contrasts sharply with the question form where the person is simply asked if he or she would like any of the choices offered, and where the answer "no" is perfectly all right, as in (48).

(48) Do you want coffee, tea, or milk?

This is the intonation of a normal, polite yes/no question, which I shall not consider here. I would like to discuss, instead, the pattern illustrated in (47).

The choices may number anywhere from two up. The reason is clear: to make a choice, you must have at least two options. If the question has N options, (in (47), N=3), the first N-1 options are spoken with a rising intonation on each option, and the last option has a falling intonation.

In particular, the rise and fall are star-linked. The starred-element in each disjunct -- each offer -- is the final low of that phrase. In short, the pitch pattern is exactly what would derive from the autosegmental formula $\overset{*}{L}H$ for each of the first N-1 disjuncts. When the disjunct has final star, this is realized, as we would naturally expect, as a rising tone ("tea" , e.g., in (47)).

Precisely the same type of observation leads to the conclusion that the tone melody for the final disjunct is $\overset{*}{H}L$. To derive (47), the autosegmental representation is (49).

```
(49) $Do you want  cóffee $   téa  $  or mílk  $
     |                    |        |           |
     $             L* H   $  L*H   $   H*L     $
```

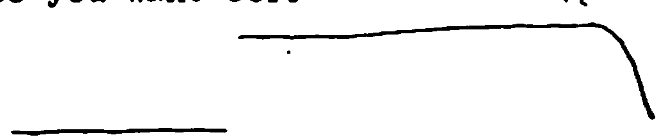
In short, the tone melody for the disjunctive question is (50).

(50) $\left(\overset{*}{L}H\$\right)_{1}^{n} \; \overset{*}{H}L\$$

If it were not apparent already, the existence of a single formula like (50) makes it abundantly so that the star-to-star association rule must be stated in terms of a starred element, not in terms of an H-tone or an L-tone, because in applying to formula ( ), the same rule must at one point associate an L-tone and at anbther associate an H-tone with the same mechanism. The only characteristic that the two share is that they possess the phrasal star.

(iii) There is another possible intonation for the disjunctive question in (47), one which is slightly less polite, it seems to me; it could be sketched as in (51).

(51) Do you want $\overset{*}{\text{coffee}}$ $\overset{*}{\text{tea}}$ or $\overset{*}{\text{milk}}$?



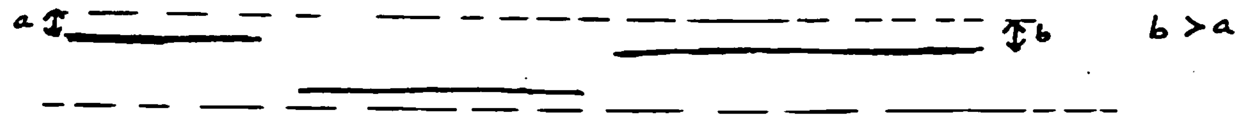
The formula for such an intonation is as in (52). I omit the derivation, for it is like the others we have seen.

(52) $L\left(\overset{*}{H}\$\right) \overset{*}{H}L\$$

If we compare (52) with the other formula for disjunctive questions, (50), we find that the two differ in this notation precisely in the parenthesization. That the asterisk has moved inside the parentheses is predictable rather than accidental, for each intonational phrase must be accentually linked with its melody.

Let us consider one final tonal contour, instructive in several respects. Imagine yourself exasperated by someone making rings on your desk with empty soda bottles. By the time you have asked him three timesto cease and desist, the intonation might be like in (53).

(53) Will you st*op putting those empty C*oke bottles on my desk?



Experimenting with different possible sentences of this form, we find that the end of the initial High plateau is accentually-linked (that is, comes on a starred syllable), and likewise with the beginning of the final High plateau. The intervening syllables are Low. This is notated, as we have seen, $\overset{*}{H}$ L $\overset{*}{H}$.

The second High is lower in pitch than the first, however, obligatorily. In fact, this lowering of pitch is just of the same degree as the "downdrift" commonly noted in African tone langues in the tonal sequence HLH. For a language to have such a down drift is the normal, unmarked case; English is no exception in this regard. The difference in surface pitch between the two Hs thus in no way suggests the introduction of a third toneme to the English lexicon.

The intonational pattern in (53) is interesting in that it is exceptional to a fairly general characteristic of English sentences: pitch either falls or rises at the end of a sentence. This particular observation, however, is not a real linguistic generalization: the rises arise from the boundary element "?" alluded to above, and the falls from the neutral pattern $\overset{*}{H}$L. (53) is a demonstration that, given the proper circumstances, final level pitch is quite possible in English.

8 Conclusion

I have attempted to show in this paper that when lenthening and intonation are taken into account, as they of course must be, there remains little justification for the claim that stress is not a binary feature, and that the questions of stress and intonation must not be separated in the way they have been traditionally. Most importantly, the aim of this paper has been to demonstrate the power of a generative system extended to autosegmental levels. In such a system, the existence of accentually-linked contour tones leads immediately and without choice to an autosegmental analysis of tone for the language in question (here, English). We have further observed how the autosegmental level, once introduced, predicts pitch or intonational facts which -- whether they were explicitly observed or not -- are paradoxical in a more orthodox generative system.

Footnotes

(October 1974)

*My thanks for helpful comments on an earlier version of this paper must go to Dan Kahn, Mark Liberman, and Morris Halle. I should add that since writing this paper I have become aware of Bolinger's important contributions in this field; a number of the points made here can be read as formalizations of his earlier observations, though in point of fact they were not arrived at as such.

This work was supported by a grant from the National Institutes of Health, 5T01 HD00111-10.

1. The only serious challange I am aware of is Vanderslice and Ladefoged (197). Their points of disagreement with SPE are quite similar to mine; we diverge completely in our positive suggestions.

2. On Sanskrit, see May and Goldsmith (to appear); on Japanese, Goldsmith (1975a) and Haraguchi(1975); on nasalization, Goldsmith (1975c) and (to appear); on vowel harmony, Clements (1975a,b).

3. Such a hypothesis is actually too strong, I believe. I would accept the suggestion made (and later rejected, though I don't know why) by Ivan Sag, that certain boundary elements cause a raising of the pitch, or more accurately, of the pitch registers. Thus even H-tones are raised, relatively, by this non-accentually linked element. This occurs in the final rise of English yes/no questions, for example. See also section 7 below.

4. For a more detailed exposition and development of some points in section 1, see Goldsmith (1975a).

6. When the tonal inventory includes a downstepping toname, as in Uhũhũ Igbo, for example, it is easily demonstrated that such spreading does not result in the copying of the tonal feature specifications to each syllabic segment, for that would produce a downstep between each occurrence. Rather, the spreading is of the association-domain of an individual toneme, as the autosegmental system predicts. For further discussion, see Goldsmith (1975b,c) and especially (to appear).

7 ~~6~~. This is a slight over-simplification. The notions "tense" and "lax" are, first of all, highly problematic; secondly, the back, unrounded vowel /a/ does not undergo vowel-shift. These matters are discussed at some length in Sound Pattern ~~SPE~~ and are largely irrelevant to the present discussion.

8 ~~10~~. This property of the Well-formedness Condition is not put to any further use in the analysis of English tone, to my knowledge; it plays an important role, however, in the tonal systems for Osaka Japanese (Goldsmith (1975a)), Samaveda Sanskrit (May and Goldsmith (to appear)), and Tonga (Cohen(1974), although not stated as such in his analysis).

9 ~~11~~. Such a suggestion was in fact made by Emily Pope (1971), where she says, "In general, in declarative sentences, all other things being equal,

a syllable with greater stress receives higher intonation;" and she gives the example "acclimated," where the first syllable receives "primary stress", the third tertiary. She suggests then that the third syllable in fact is realized with a rise in intonational pitch, but this is not true.

10. Explaining this assignment is not our task here; this issue is one of the central topics of English intonational studies: the placement of sentential accent.

11. Pointed out to me by Jill Carrier.

12. This conclusion is not surprising. Margaret Allen, for example, in Linguistic Inquiry VI 2, has argued that compoudds in Welsh are of two sorts, those with a weak boundary, +, and those with a stronger boundary -- which she notes as ## -- between them. She presents no reason why the stronger boundary might not just be #, which would be the more neutral position. After all, if the lexicon (the morphological component) can insert boundary elements between morphemes, and chooses to place + in certain compounds, why should we not find a parallel situation where # was inserted? The answer, surely, is that we should expect that; and the cases where ## has been imputed are, more reasonably, to be taken as those where in fact only a single # appears. And recall that the position in Sound Pattern which seemed to lead to the conclusion that there were two #'s in compounds was based on a pre-Lexicalist Hypothesis view of lexical insertion.

References

Chomsky, N. (1967) "Remarks on Nominalizations," in N. Chomsky, Studies on Semantics in Generative Grammar, Mounton (1972).

Chomsky, N. and M. Halle (1968) The Sound Pattern of English, Harper and Row, New York.

Clements, G.N. (1975a)

--- (1975b)

Cohen, M. (1974) "Topics in Tonga Tone," unpublished paper, MIT.

Edmundson, T. and J.T.Bendor-Samuel (1966) "Tone Patterns of Etung," Journal of African Languages, 5, 1-6.

George, I. (1970)

Goldsmith, J. (1975a) "An Autosegmental Typology of Tone: and How Japanese fits in," in Papers from the 5th Annual Meeting of the North Eastern Linguistics Society, Harvard University, Cambridge, Mass.

--- (1975b) "Tone Melodies and the Autosegment", in the Proceedings of the 1975 African Linguistics Conference, ~~txm~~ Ohio State Working Papers in Linguistics, vol. 20.

--- (1975c) "An Overview of Autosegmental Phonology" unpublished MIT mimeo; presented to the Conference on Formalism in Phonology at MIT, June, 1975.

--- (to appear, 1976) Autosegmental Phonology, doctoral dissertation, MIT, Cambridge, Mass.

Halle, M. (1973)

Halle, M. and S.J.Keyser (1971) English Stress: Its Form, Its Growth, and its Role in Verse, Harper and Row, New York.

Haraguchi, Shosuke (1975) The Tone Pattern of Japanese, unpublished doctoral dissertation, MIT, Cambridge, Mass.

Leben, William (1973) Suprasegmental Phonology, unpublished doctoral dissertation, MIT, Cambridge, Mass.

May, ~~Robert~~. and J. Goldsmith (to appear) "Tone Sandhi in Sanskrit".

Pope, E. (1971)

Selkirk, E. (1972) The Phrase Phonology of English and French, unpublished doctoral dissertation, MIT, Cambridge, Mass.

Williams, E. (1971) "Underlying Tone in Igbo and Margi", unpublished paper, MIT.

Woo, N. (1969) <u>Prosody and Phonology</u>. Indiana University Linguistics Club mimeo.