ENGLISH AS A TONE LANGUAGE*

John GOLDSMITH
(Indiana University)

It is commonly acknowledged that what generative and traditional grammarians 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 of 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 orignally 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 Leben (1973), who discussed it at length.

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 MHL. Thus, we set the two strings parallel to one another on the page :

(1)   ma nong ga he* la

                M   H* 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)   V   V   (or)   V*
                → | *
             T*   T*   T*

(3)   ma nong ga he* la
                      /
                M   H* L

By some general principles, the L tone must be realized too, and obviously it must come to the right of the syllable "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 he* la   ma nong ga he la

                \\ |/ / |
                M   H* L

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 - 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].
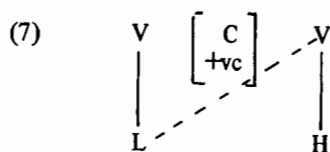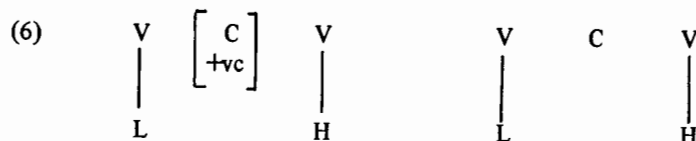
The domain of the toneme can be more, and can be less than a single vowel or syllable. And syllables can be underlyingly toneless. These perplexing facts, and other equally anomalous characteristics of tonal behavior, led to the autosegmental system. Put simply, this theory suggests that when we hear 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)

$$\underset{L \qquad H}{\overset{V}{\diagdown}}$$
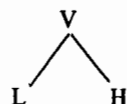
Or we may consider a slightly more complex example. In Nupe, for example, we find a very common type of "tonal flop" rule; see George (1970) for details. A High-toned vowel preceded by a Low tone becomes a Rising tone if the intervening consonant is voiced - clearly an "assimilatory" process (we find 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 convention that a dotted line represents a structural change, as in (7).

(6)

$$\underset{L}{\overset{V}{|}} \begin{bmatrix} C \\ +vc \end{bmatrix} \underset{H}{\overset{V}{|}} \qquad \underset{L}{\overset{V}{|}} \quad C \quad \underset{H}{\overset{V}{|}}$$

(7)

$$\underset{L}{\overset{V}{|}} \begin{bmatrix} C \\ +vc \end{bmatrix} \underset{H}{\overset{V}{|}}$$

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 princple, 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 of 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, May and Goldsmith, and 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 of 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 *Sound Pattern,* 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 *stress.* 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] :

(8)

	+ 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 :

(9)   tense /i/   unstressed: various   [i]   no vowel shift

                        stressed: variety   [ay]   vowel shift

     lax /æ/   unstressed : atomic   [ə]   reduced to ə

                        stressed : atom   [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 "Istress". 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

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

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

(11)

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

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

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

"anticipate"

4. Deriving the Melody

Now the question can finally arise : what do we do with the star one 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 : MH̊L. These three letters stand for Mid, High, and Low. 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) applicá̇tion

    |

  M H̊ L

What this means - so far - is that the High tone is being realized on the third syllable. Noting 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 (See (13)).

(13) applicá̇tion

       FORBIDDEN

  M H̊ L

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 "å" and the H with "tion", although, we could do *one* of these :

(14) applicá̇tion

       FORBIDDEN

  M H̊ L

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) applicá̇tion

    / |

  M H̊ L

By now both L and the final syllable, "tion", meet the Well-formedness Condition : that is, they are both associated with someting. 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 *minimal* way to meet the Well-formedness Condition. To get (15) requires only one additional association; (16) requires two. Therefore (15) is derived by the Well-formedness Condition.

(16)   applicátion

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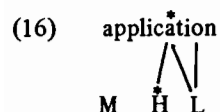
The principles we have specified so far leave it an open question whether (17) or (18) is the ultimately correct well-formed result of structure (12). Both (17) and (18) represent the addition of three association lines, and on purely numerical grounds, neither is preferred.

(17)   application

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

(18)   application

$$\text{M} \quad \overset{*}{\text{H}} \quad \text{L}$$
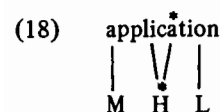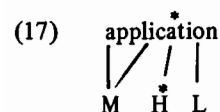
The same indeterminacy arises with the derived form of a word like "America". Is (19) to be derived, or (20) ? - both represent three added association lines, but are empirically different.

(19)   américa

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

(20)   américa

$$\text{M} \quad \overset{*}{\text{H}} \quad \text{L}$$
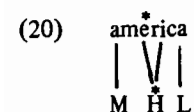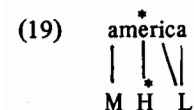
In the case of an ambiguity like this - that is, a competition between a starred toneme and an unstarred toneme for spreading through the Well-formedness Condition - it is always the unstarred toneme that dominates. In these cases, that means that (17) and (19) are derived, an (18) and (20) are not[8].

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 gives us (21).

(21)   Ka   la   ma   zóo

$$\text{M} \quad \overset{*}{\text{H}} \quad \text{L}$$
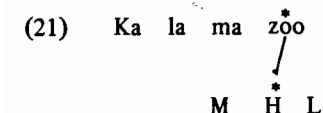
The Well-formedness Condition corrects this to (22).

(22)   Ka   la   ma   zóo

$$\text{M} \quad \overset{*}{\text{H}} \quad \text{L}$$
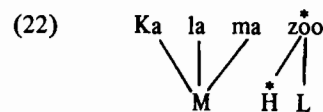
(22) contains a representation with two tonemes associated with the same syllable, which is, of course, perfectly permissible. It signifies a falling tone from High to Low. And that is precisely the fact in English : words with "final stress" realize a marked falling tone on the ultima, whereas words 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)   Wittgenstein

$$\text{M} \quad \overset{*}{\text{H}} \quad \text{L}$$
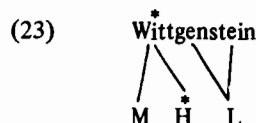
There is, however, no particular rising tone on the first syllable of such words, as this predicts there should be. Two possibiliteis present themselves : either the tone melody for English neutral intonation is really $\overset{*}{\text{H}}$L, not M$\overset{*}{\text{H}}$L; or there is a rule deleting M when it occurs on the same syllable as $\overset{*}{\text{H}}$.

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) or (25), where (24) is that predicted by the melody $\overset{*}{\text{H}}$L. Such a proponent might further propose an optional late rule raising the pitch of the starred syllable, giving the optional variant (25).

(24)   archipélago

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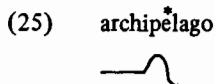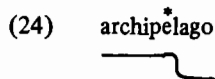
The defender of the melody M$\overset{*}{\text{H}}$L would be forced to posit a MH-Simplification rule, as in (26).

(26)

$$\begin{array}{ccc} V & & H \\ \wedge & \rightarrow & | \\ M\ \ H & & H \end{array} \quad \text{or} \quad \begin{array}{c} V \\ /\!\!\wedge \\ M\ \ H \\ 1\ \ 2 \\ \emptyset\ \ 2 \end{array}$$
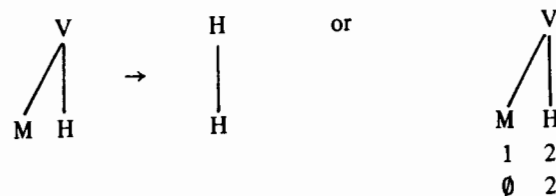
(Tonological transformations factor only the tonological string; the non-tonological elements that enter into such a rule, as in (26), are, strictly, conditions on the rule rather than part of the structural description of the rule). I shall leave the question open regarding which of the two solutions 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 in any reference to "stress subordination" rules which by convention would lower to 2 stress, and then 3 stress, the other syllables marked [+stress] in the word (see Chomsky and Halle, Halle and Keyser, Halle).

This absence of stress subordination is, of course, no oversight; the claim being made here is that the segmental feature "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 1 stress was associated with a rise in pitch, one might attempt to extend the interpretation so that 2 stress meant a slightly lower rise in pitch, 3 stress slightly lower again, and so on. Such an interpretation - admittedly not found in *Sound Pattern*[9] - would predict a rise in pitch on the final syllable of *Wittgenstein*, which we observe is false. The 3 stress there is Low, as Low as the second syllable "gen", which is unstressed.

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

6. Nouns 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 has been said in isolation, in accordance with the principles sketched above. Every word has a star (*) placed on it somewhere, and that would distribute the tone melody over it.

The test of our hypothesis - that non-binary segmental stress features may be eliminated within an autosegmental framework - will surely come when compound forms are confronted, where the *nstress* system of *Sound Pattern* made remarkable results. Up to this point, we have considered, in effect, [-stress], [1 stress](*), and [3 stress]; now [2 stress] and [4 stress] come in.

What must be shown, if our proposal is to be maintained, is that all differences between *nstress* and *mstress* are either illusory or due to some factor other than stress. Under this latter rubric fall *length* differences, which depend on a number or factors - but none of them are stress.

We shall argue that (1) the perceived difference between the 3 stress in (28) and the 2 stress in (29) is due to the difference in *length* .

(28) black-board   (29) black board
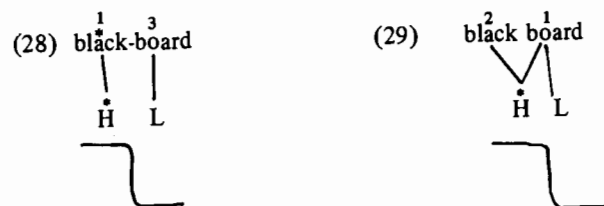
(ii)   that this difference is not a general fact about those vowels that the *Sound Pattern* system would designate as 2 stress and 3 stress, and not even a general fact distinguishing compounds from adjective-noun pairs (as in (28) - (29));
(iii)   that therefore the difference in length is not *attributable* to a difference in an abstract *nstress* specification, but to certain straightforward lengthening processes sensitive to word-boundaries; we shall see that such contrasts as those between *black-board* (that one writes on) and *Black board* (a committee presided over by Justice Black) show certain actual differences which would not be describable if length were predictable from an *nstress* feature; and (iv) that some other differences - between "elevator operator" and "elevator boy" - were awkwardly dealt with when an *nstress* system was used to describe them, but fall out naturally from this revision.

Consider first the contrast in (28) / (29), the well-known contrast between (28) a black-board that one writes on, and (29), a board that is black. The "2 stress" on *black* in (29) is supposed to indicate a prominence that *board* does not have in (28), being "3 stress" there. Although the representations in (28) and (29) (disregarding stress) do reveal a difference between the two syllables - *bláck* is H-toned, *boárd* is L-toned - a glance at a form like (30) shows that this could not be the relevant difference, since what are called "3 stressed" syllables can be H-toned as well as L-toned.

(30) $\overset{3}{m}\overset{1}{a}$gazine

            H   L

An analysis of the duration of the syllables shows, however, a clear difference which is not hard to hear. We can make the pair in (28) / (29) minimally different by giving them the same tonal pattern as in the context, "But that's *my* - ", which is roughly as in (31).

(31)   But that's *my* black board.....

For some reason - which we shall return to - the *black* in *black board* is considerably longer than the *black* of *black-board*.

Now we shall see that this difference of length is not due to a difference of stress patterns, if stress pattern is taken as in the sense of Sound Pattern. Other 2 stresses, for example, do not lengthen as the *black* does in (29). Thus (32), when the cyclic application of stress in the *Sound Pattern* system applies to give a 1-2-4 stress pattern, displays no lengthening of *black* in *kiddy black-board*[10].

(32)   Kiddy   black-board

	1	1	first cycle
	1	3	second cycle
1	1	3	
1	2	4	third cycle

Conversely, an adjective-noun sequence as in (33), where the adjective is not monosyllabic, displays no lengthening of the first syllable corresponding to *black* in (29).

(33) $\overset{2}{h}$acking $\overset{1}{c}$ough

In general, then, being marked as 2 stress is not a sufficient condition to undergo the lengthening displayed in (29).

Nor is it a necessary condition. Monosyllabic 1 stress words undergo this lengthening as in *the board was black*. Yet this occurs, again, only if the word is monosyllabic. In short, the difference in length observed in (28) / (29) arose from something quite other than a 2 stress/3 stress distinction. All the facts observed to this point concerning length follow from the hypothesis that lengthening occurs in monosyllables surrounded by double word-boundaries, if we assume that lexical compounds like *black-board* have the structure ## black#board## .

(34) Monosyllabic lengthening rule

     V becomes lengthened $/\#\# C_0 - C_0 \#\#$.

The assumption that lexical compounds have just the single word-boundary, rather than the double word-boundary proposed in *Sound Pattern*, is a natural consequence of the lexicalist hypothesis, not yet adopted in *Sound Pattern*. Given that the lexicon can construct words and place morphological boundaries between morphemes, it would be surprising if there were no constructions that did not place single word-boundaries inside a word; in effect, we are suggesting that the difference between a compound like *black#board* and words like *govern#ment* is that both halves of the former "could" be words, something not true of the latter.

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 *tone-deaf* and *stone deaf*.

(35)   tone-deaf   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 or she just can't hear music well, or perhaps the difference between the two words in (35). Consequently, "tone-deaf" is a compound, and the compound tone-assignment is expected. Thus the two forms of (35) are predicted, correctly, parallel to (28) and (29) above.

The Monosyllabic Lengthening Rule mentions nothing about stress, and so we would expect it to apply equally to 1-stressed ("starred") vowels and to the 2-stresses we have seen. This is, as we observed above, true (*The board was black*); the distinction is equally manifested in the pair in (36) :

(36) a. *The lion turned on his trainer* (no lengthening)
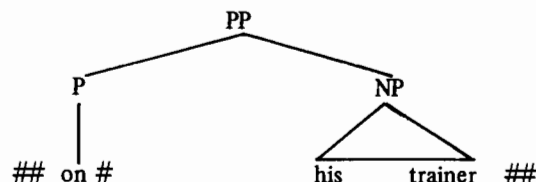
b. *The teacher turned on his students* (lengthening)

Furthermore, there are compounds that would be described in the n-stress system as having 1-3 patterns which display a different lengthening. If there were a committee chaired by Justice Black, such a Black board, though starred on the first syllable, thus displaying a tone pattern as in (37), is clearly pronounced differently from *black-board*.

(37)
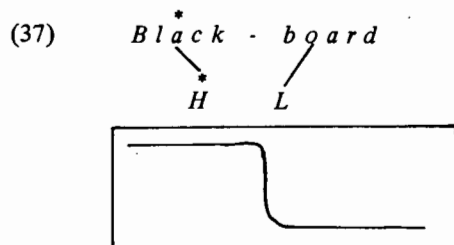
The difference here has nothing to do with tone or stress in the binary sense : *Black* is simply lengthened. Whether it is by the Monosyllabic Lenghthening Rule, suggesting two word-boundaries in (36b), which I suspect, or whether this is due to some other process, is irrelevant : the point is that describing both as 1-3 is inadequate, and attention to syllable lengthening that the autosegmental analysis forces is necessary here.

It would follow as a consequence of the strong assumption that stress is inherently binary that the difference between 4-stress and 3-stress that has been cited word-internally must be due to some other factor. In an attempt to measure this difference, I made recordings of the two pronunciations of the word "Ticonderoga", one with a "34010" stress pattern, and one with a "43010" pattern. The difference refers to whether the first syllable is more prominent in some sense than the second, or conversely. I was rather liberal in trying to emphasize either the first or the second syllable at the expense of the other, but the recordings sounded natural.

Below in Table 1 are the results of the analysis of the recordings for intensity, pitch, and syllable length. I could not measure the length of the final syllable, and its pitch measurement is probably inaccurate, but it is not relevant to our present question, fortunately.

Table 1

	3 4 o 1 o Ticonderoga	4 3 o 1 o Ticonderoga
Pitch (hz)	104/100/88/88/62?	92/108/92/96/72?
Amplitude	6.8/6.5/2.1/4.8/1.8	6.8/6.3/1.8/6.7/0.8
Approx. syll length (msec)	160/280/100/180/?	130/260/110/150/?

In short, what we find as a difference between the two utterances, such as it is, is that the second syllable is perceived as more prominent when the pitch goes up rather than gradually down. There is no significant difference in amplitude of the syllables in the two pronunciations, and our hypothesis is that the difference between these two pronunciations is one that is not linguistically significant. It could not be used to differentiate two morphemes, for example; the two are in free variation. Why ? Simply because pitch may go up and down before the main accent of a word; it is this freely varying pitch that has been interpreted in the *Ticonderoga* case as the reflex of stress. But we must emphasize that the discovery of pitch differences between the two ways of uttering the word does not support the n-stress analysis; there was no reason in that system to expect the 3/4 stress difference was purely a matter of a slight rise in pitch for the 3-stress, and no other prosodic effect. On the rather more simple-minded "only binary stress" view, these minor pitch variations are to be expected, but they are also expected to play no significant linguistic role. Which, in fact, seems to be the case.

All this does not disprove the validity of the n-stress system; but it does call into question what the empirical base is that it is supposed to rest upon. Pitch, however, not stress, is what is our concern here.

Another problem that has remained with various systems of stress-subordination is the apparent difference between compounds like *elevator-operator,* which has been held to have a "1-2" stress pattern, and *elevator boy,* which has been held to have a "1-3" pattern (see SPE, pp. 93, 109). That is, the second element in *elevator-operator* is thought to have more stress than the second element of *elevator-boy.* 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 *elevator-operator* and *elevator-boy* is most easily understood and explained in the light of the present discussion. The tone pattern assigned to *elevator-operator* is as in (38); this corresponds to a pitch pattern as in (40). Now, in fact, many speakers of English have sluggish transitions from High to Low tone; the pitch of this compound then is more like (42) than (40). By contrast, these speakers would say *elevator-boy* as in (43), derived from (41).

(38)   elevator   operator

H   L

(39)   elevator-boy

H   L

(40)

(41)

(42)

(43)
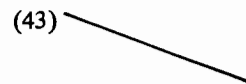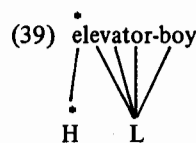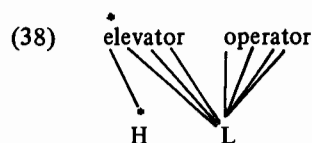
The difference is striking; the main-stress of *operator* in (43) is at a considerably higher pitch relative to the final syllables of *operator* ("perator"); in (43), however, "boy" is pronounced at the bottom of the pitch contour, giving the impression of lower relative prominence. In short, the *operator/boy* 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 MHL 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 sentence 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 star 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 dispose of this, for it is only an artifact of the limited kinds of examples considered up to this point.

(i)   Consider as indignant response (44).

(44) Elephantiasis 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 HL. This means that the L tone falls on the accented syllable - *ti* - and the preceding H tone spreads over the preceding syllables ("elephant-") through the well-formedness condition.

The complete analysis of a sentence like (44) 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 syggestion 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 number of interesting consequences of this approach, but I shall not pursue them here.

In short, the underlying form is as in (45), yielding (44) by the autosegmental principles duscussed above.

(45) $ Elephantiasis isn´t the most dread disease in the world ?

$ H L   ?
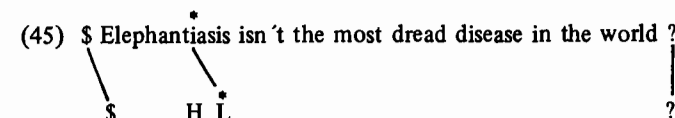
(ii)   Disjunctive questions are the sort of questions where the asker assumes that the person asked wants one of the choices : "Do you want coffee, tea, or milk ?" asked in such a way that the person asked feels it incumbent on him or her to take one of the three. This could be sketched as in (46).

(46) 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 (47).

(47) 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 (46).

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 (46), 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 (46)).

Precisely the same type of observation leads to the conclusion that the tone melody for the final disjunct is $H\overset{*}{L}$. To derive (46), the autosegmental representation is (48).

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

$$\$ \quad\quad \overset{}{L}\, H \quad \$ \quad L\, H \quad \$ \quad\quad H\, \overset{}{L} \quad \$$$

In short, the tone melody for the disjunctive question is (49).

(49) $(\overset{*}{L}\, H\, \$)^n_1 , \ \overset{*}{H}\, L\, \$$

If it were not apparant already, the existence of a single formula like (49)

makes it clear 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 (49), the same rule must at one point associate an L-tone and at another 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 (46), one which is slightly less polite, it seems to me; it could be sketched as in (50).

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

The formula for such an intonation is as in (51). I omit the derivation, for it is like the others we have seen.

(51) $L\ (\overset{*}{H}\,\$)^n_1 , \ \overset{*}{H}\, L\, \$$

If we compare (51) with the other formula for disjunctive questions, (49), 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 times to cease and desist, the intonation might be like in (52).

(52)   Will you $\overset{*}{\text{stop}}$ putting those empty $\overset{*}{\text{Coke}}$ bottles on my desk ?

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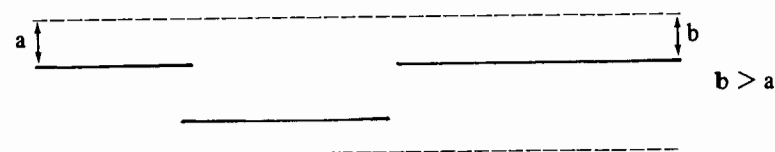
Experimenting with different possible sentences of this form, we find that the end of the initial High plateau is accentually-linked (that is, comes on a starred syllable), and likewise with the beginning of the final High plateau. The intervening syllables are Low. This is notated, as we have seen,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

The second High, however, is lower in pitch than the first, obligatorily. In fact, this lowering of pitch is just of the same degree as the "down drift" commonly noted in African tone languages 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 HL. (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 lengthening 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, one introduced, predicts pitch or intonational facts which whether they were explicitly observed or not - are paradoxical in a more orthodox generative system.

NOTES

* This paper was written in late 1974, and revised in March 1975. My thanks for comments and criticisms of this paper go to Morris Halle, Mark Liberman, and Dan Kahn. At that time, I was unaware of Dwight Bolinger's important contributions to this subject, which parallel and in many ways go beyond what I propose. The framework here, however, is different from his. The "autosegmental" system sketched here is developed at some length in two later references, Goldsmith (1976a and b). A revised version of this paper appears in section 2.3 of Goldsmith (1976b). Another, subsequent development of related ideas is found in Liberman (1975). At this point, although I find the material presented here correct in its essentials, as far as it goes, it seems clear that the phenomena dealt with here will require ultimately a system which is quite a bit more complex. This work was supported by a grant from the National Institutes of Health, 5T01 HD00111-10, at the Massachusetts Institute of Technology.

1   The only serious challenge I am aware of is Ladefoged and Vanderslice (1972). Their points of disagreement with *Sound Pattern* are quite similar to mine; we diverge completely in our positive suggestions.

2   On Sanskrit, see May and Goldsmith; on Japanese, Goldsmith (1975a) and Haraguchi (1975); on nasalization, Goldsmith (1976a and b); on vowel harmony, Clements (1975, 1976).

3   Such a hypothesis is actually too strong, I believe. I would accept the suggestion made (and later rejected, although 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 the English yes/no question, for example. See also section 7 below.

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

5   The details of this process are discussed below.

6   When the tonal inventory includes a downstepping toneme, as in Ohuhu Igbo, for example, it is easily demonstrated that such spreading does not result from 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 the individual toneme, as the autosegmental system predicts. For further discussion, see Goldsmith (1975b), (1976(a,b)).

7   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 participate in vowel-shift. These matters are discussed at some length in *Sound Pattern* and are largely irrelevant to the present discussion.

8   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 system of Osaka Japanese (Goldsmith (1975a)), Samaveda Sanskrit (May and Goldsmith), and Tonga (Cohen (1974), cited in Goldsmith (1976b)). It is modified somewhat in Goldsmith (1976b), and Clements (1976).

9   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", and 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 Surely we won't hear the response now that the lengthening rule is cyclic, and that (32) is saved by "strict cyclicity". If strict-cyclicity would block its application in (32), it would equally block it, wrongly, in (29).

BIBLIOGRAPHY

CHOMSKY N., 1967, Remarks on Nominalizations, in N. Chomsky, *Studies on Semantics in Generative Grammar*, Mouton, 1972.

CHOMSKY N. and M. HALLE, 1968, *The Sound Pattern of English*, Harper and Row, New York.

CLEMENTS G.N., 1975, A Set of Universal Conventions for Vowel Harmony, MIT unpublished paper.

-, 1976, The Autosegmental Treatment of Vowel Harmony, Delivered to the Third International Phonology Meeting, Vienna, Sept. 4, 1976; to appear in *Phonologica 1976*, ed. W.U. Dressler.

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, Nupe Tonology, *Studies in African Linguistics*, 1, 1.

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, *Ohio State Working Papers in Linguistics*, vol. 20.

-, 1976a, An overview of Autosegmental Phonology, *Linguistic Analysis*, 2, 1.

-, 1976b, *Autosegmental Phonology*, doctoral dissertation, MIT : Indiana University Linguistics Club.

HALLE M., 1973, Stress Rules in English : A New Version, *Linguistic Inquiry*, 4, 1.

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 PhD dissertation, MIT.

LADEFOGED P. and R. VANDERSLICE, 1972, Binary Suprasegmental Features and Transformational Word-Accent, *Language* 48, 819-38.

LEBEN William, 1973, *Suprasegmental Phonology*, unpublished PhD dissertation, MIT.

LIBERMAN Mark, 1975, *The Intonational System of English*, unpublished PhD dissertation, MIT.

MAY R. and J. GOLDSMITH, 1975, Tone Sandhi in Sankrit, unpublished manuscript.

POPE E., 1971, Answers to Yes/No Questions, *Linguistic Inquiry* 2, 1.

SELKIRK E., 1972, *The Phrase Phonology of English and French*, unpublished PhD dissertation, MIT.

WILLIAMS E., 1976, Underlying Tone in Igbo and Margi, *Linguistic Inquiry*, 7 3 (written in 1971).

WOO N., 1969, *Prosody and Phonology*, MIT PhD dissertation; Indiana University Linguistics Club.