Autosegmental Phonology
PhD dissertation
MIT

John Goldsmith

1976

ABSTRACT

A modification of the theory of generative phonology is
suggested in this thesis in the introduction of parallel tiers
of segments (or "autosegments"). This is shown, in the first
chapter, to resolve certain formal and substantive problems in
the current theory. A detailed analysis of Igbo, a tone lang-
uage of Nigeria, is presented in autosegmental terms in Chapter
Two, as well as a new analysis of the phenomenon of "downstep"
found in most African languages. In Chapter Three, these no-
tions are developed to account for stress and intonation patterns
in English, and various accentual and non-accentual systems are
dealt with as ways of co-ordinating the tonal and syllabic tiers
of autosegments. Work by G. N. Clements on vowel harmony is
cited to support a more general account of autosegmental phon-
ology. Chapter Four presents a hypothesis for the origin of
autosegmental phonology, suggesting that the inherent geometry
at the phonetic level is "autosegmental," but that language
acquisition will include the task of "de-autosegmentalization,"
which tends to collapse the multilinear autosegmental geometry
to a linear one at the lexical level.

EPIGRAPH

From the continuing saga of Guru Rug, the 34-year-old perfect Rug Salesman:

Student: O Master, I have had a revelation!

Guru Rug: And what is this revelation, my son?

Student: O Master, I have seen that you can spell "Guru Rug"--backwards!

Guru Rug: Yes, my son, that is true. You can spell anything backwards.

--WCAS Radio 1974

"For in the science of harmony, as you probably know, the same thing happens. The teachers of harmony compare the sounds and consonances which are heard only, and their labor, like that of the astronomers, is in vain."

"Yes, by heaven!" he said; "and 'tis good as a play to hear them talking about their condensed notes, as they call them; they put their ears close alongside of the strings like persons catching a sound from their neighbor's wall--one set of them declaring that they distinguish an intermediate note and have found the least interval which should be the unit of measurement; the others insisting that the two sounds have passed into the same--either party setting their ears before their understanding."

--Plato: *The Republic* Book VII

TABLE OF CONTENTS

Preface

"It all started about two Thanksgivings ago....
It was two years ago on Thanksgiving."

The idea behind this thesis began with a reading of Will Leben's thesis on suprasegmental phonology, in which he argued that in some languages, even short vowels could bear two successive tones.

Impossible!

....Because inexpressible, in segmental terms. And yet the facts persisted; Leben's case was clear. How <u>could</u> we express two successive tones on a single segment: In a little paper at the end of 1973 called "Tonemic Structure", I suggested a notation for falling toned vowels: V   . The observation was made then -- by Paul

H L

Kiparsky, I think -- that this would account for what, in this thesis, is called "stability" of tone melodies.

As a notation, this was simple enough; as a theory of anything, it was quite inadequate. Igbo, a tone language with an excellent descriptive grammar, Green and Igwe's, seemed a good testing ground for the theory. Edwin Williams had looked at a floating tone in the Egbo system, and his analysis was insightful, though it was formulated in the same theoretical framework as Will Leben's theses.
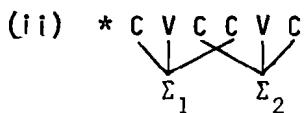
It proved fruitful. The work in Chapters 1 - 3 of this thesis is essentially the product of 1974. The last chapter, Chapter Four, is of more recent vintage, and remains quite a bit more speculative, but submits a hypothesis which accounts for the origin of auto-segmental tiers, and also explains why the notion of distinctive feature and the notion of an underlying inventory of atomic units -- call them phonemes -- why these two notions are not incompatible, but are rather two aspects of a more general strategy in language acquisition, what I call "de-autosegmentalization." It is, in any case, a <u>step</u> towards such an explanation.

There are several topics more or less directly connected with autosegmental phonology that I shall, more or less, eschew.

The first is the question of the syllable. In a paper on tone in Sanskrit by Robert May and myself, I suggested that the syllable might well be considered an autosegmental level. In this sense, the string of familiar C and V segments could be broken up into an autosegmental representation where the second tier was composed of syllables:

(i)   C V C V C V C V C
         Σ_1 Σ_2 Σ_3 Σ_4
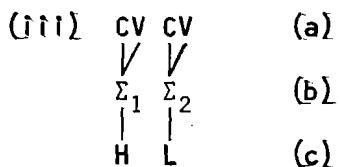
The Well-formedness Condition would permit structures like (i), where a single consonant occurred in two syllables, but none like (ii):

(ii) * C V C C V C
          Σ_1 Σ_2
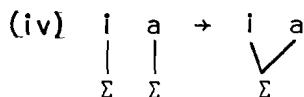
Furthermore, all segments would occur in at least one syllable.

This proposal has certain immediate merits, beyond merely formalizing the traditional notion, and beyond the advantage gained in talking about syllables without placing syllable boundaries in various places between the Cs and the Vs, a by now familiar approach that is always tripping over its own feet. Most importantly, this approach would remove the glaring restriction in the theory of autosegmental tonology that says that only syllabic segments can associate with tonemes. Given syllabic structures like (i), CV̆ CV̆ would become (iii).

(iii) CV CV   (a)
       Σ_1 Σ_2 (b)
 | |
       H   L   (c)
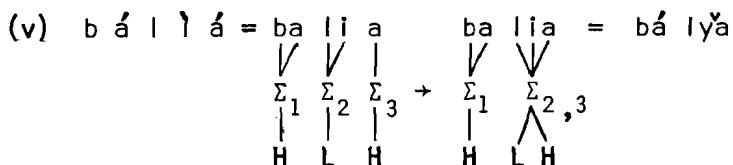
Here there is no asymmetry among levels; all segments in levels (a) and (b) enter into a Well-formed autosegmental relation, as do levels (b) and (c).
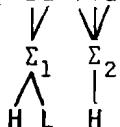
This notation suggests further that the familiar process sending e.g., la → ya is a syllable <u>merger</u> process, as illustrated in (iv).

(iv) i a → i a
      | | \/
      Σ Σ Σ
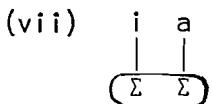
If this is true, then an immediate consequence of this is that i's tone would have to reappear on a, i.e., as in (v), but not as in (vi).

(v) b á l l á = ba li a   ba lia = bá lyă
                 Σ_1 Σ_2 Σ_3 → Σ_1 $\Sigma_{2,3}$
                 |¹ |² |³ |
                 H   L   H   H   L H
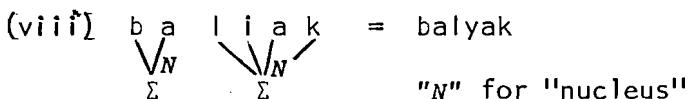
(vi)*ba lia = bâ ↓yá
        Σ_1 Σ_2
        H L   H

Such a process in (v) could be expressed as in (vii), where the circling of the two (necessarily neighboring) syllables would indicate an "identification", as mathematicians say, or joining of two syllables into one.

(vii)   i   a

$$\boxed{\Sigma \quad \Sigma}$$

Such a threory would then remove the feature "syllabic" from the highest tier in (iii) or (v), instead develop a more articulated notion of "being in a syllable", indicating how a sonorant could be the nucleus of a syllable with which it is associated. Various possibilities and notations come to mind here -- like (viii) -- but this is not a simple matter by any means, and as I say, I will not be pursuing it here.

(viii)   b a   l i a k   =   balyak

$$\bigvee_{N} \qquad \bigvee_{N}$$
$$\Sigma \qquad\quad \Sigma \qquad\qquad "N" \text{ for "nucleus"}$$

We may ask, further, what the elements on the syllabic tier are. I have indicated them with subscripted sigma's; but on all other tiers, the units were segments with feature specifications. What are these syllables?

Such a question is interesting, even important, within the context of the discussion of Chapter Four below, but I shall leave the question untouched. In a forthcoming thesis, Dan Kahn presents a theory of syllable phonology as sketched in (i) and (ii) and obeying the Well-formedness Condition.

Two other recent theses bear directly on the material here, and I shall say little about them. One is Mark Liberman's account of English intonation, stress, and timing; the other is Shosuke Haraguchi's analysis of Japanese dialects. As I indicated above, the first three chapters below predate these works, and at my risk, perhaps, I have left them basically unchanged. Some of Mark's ideas and criticisms have made their way -- no doubt, unrecognizably -- into Chapters Three and Four. I don't doubt that a firmer and more adequate synthesis is yet to be achieved. Of Sho's work, its bredth and depth speaks for itself. For that reason, I have restricted my attention in Chapter Three to different types of systems from those in Japanese. The interested reader is urged to consult Haraguchi's analysis of autosegmental melodic association rules (Haraguchi (1975, which will be published soon); see also, regarding Japanese, Goldsmith (1975a)). Again, a deeper understanding of universally possible systems will require a synthesis of Haraguchi's work on Japanese dialects with an analysis of African accentual tonal systems.

This thesis is addressed to generative phonologists; my background and training has been in this tradition. Autosegmental theory is, I believe, an improvement to this system. A fair amount of what distinguishes autosegmental representation from, say, that found in *Sound Pattern of English* bears a resemblance to certain non-generative systems. This relationship is discussed in the introductory section, following immediately below.

Prelude: The Question of Suprasegmentals

Ever since there have been segments in phonology, there have
been phenomena that evaded segmental classification, and so there have
been suprasegmentals. As Rulon Wells noted in 1945,

> In effect, the term 'suprasegmental phoneme' as
> used in recent discussions simply denotes phonemes
> which are neither vowels nor consonants.

There have been attempts, of course, to give a definition of
"suprasegmental" that carries more information than this -- attempts
to characterize "suprasegmental" in a positive rather than a
negative way. This thesis deals with a certain subpart of what have
traditionally been called "suprasegmentals"; it attempts to give
a specific characterization of these types of prosodic phenomena.

The general framework of this "autosegmental" approach is
generative phonology, and we are, therefore, concerned with phono-
logical rules that relate phonological and phonetic levels as
much as with providing a formal characterization of these two levels.
Much of the literature concerning suprasegmentals, some of which
we shall look at below, has been formulated within a structuralist
framework; the concern there is primarily with methods of
analyzing initial phonetic data into elementary units which may then be
built up into higher level units, such as phonemes and morphemes.
Despite this basic difference, there are a number of similarities
between the questions that arise within a generative and a struc-
turalist treatment.

If we look at the American linguistic tradition in the decade
of the 1940's, we find a fair amount of discussion regarding supra-
segmentals, with contributions by, among others, Bloch, Harris,
Haugen, Smith, Trager, and Wells. All begin with the initial
assumption -- spelled out clearly, for example, in Harris (1944)
and Bloch (1948) -- that logically the first procedure in linguistic
analysis is to "slice" the linguistic signal vertically into a
number of pieces -- segments, we may call them. Care must be
taken, of course, to make sure that the signal is sliced finely
enough to find the truly minimal units of the utterance. Once
this process of "segmentation" has been completed, the classification
of these slices may proceed. The formal representation of a
signal sliced into segments is, of course, a linear sequence of
elements, but this is due to the nature of the initial process
of segmentation, rather than the inherent nature of the speech
signal.

We shall look at some of these structuralist analyses, but an
initial distinction must be borne in mind. We must separate the
issue, on the one hand, of how English stress, accentuation, or

intonation is to be dealt with from, on the other, the defining characteristics of suprasegmentals. There are two tendencies here in defining "suprasegmental": either one defines a process or feature as suprasegmental from the outset -- in effect, making it something independent of the linguistic analysis or any particular language; or else "suprasegmental" is a name we give to any process we find occurring in a particular language when it displays certain general properties, either of distribution of, in generative terms, of behavior with respect to rules and rule applications.

More specifically, someone taking the first approach might say, with Haugen (1949):

> The term prosodeme has been used in recent writings
> on phonemics as a stylistic variant of prosodic pho-
> neme to describe such modifications of the basic
> speech sounds as tone, stress, and duration. (278)

Juncture, as Haugen also notes, was frequently classified as suprasegmental, but again the grounds for the classification are not language-particular, but rather theoretical.

On the other hand, some treatments -- of components, for example, by Harris (1944) and by Bloch (1948) -- develop definitions or criteria which may be applied either language-specifically or even utterance-specifically to determine whether a certain feature is being used suprasegmentally. As I have implied, this is the approach of the present work: to develop a concept (here, autoseg-mental tier) whose particular application to a language must be determined by the facts of that language. Thus, for example, the diagnostics, or indications, of autosegmental behavior that are developed in Chapter One below will be indications to the linguist that a certain feature is, in fact, autosegmentalized in a particular language.

If we return to the structuralists' assumptions, we must rephrase this question, of course; we ask not what indications there are concerning what the language itself is doing, but we ask, rather, what the appropriate methods are for analyzing data from a particular language, and under what circumstances these methods will lead, in effect, to suprasegmentals.

One factor that led to consideration of stress as being non-seg-mental -- and thus, under one interpretation, "suprasegmental" -- was the belief (cf. Trager and Smith pp. 35-6) that the very meaning of stress was relational, and therefore could only be defined in relation to neighboring segments in the utterance -- a property, it was held, that distinguished stress from "normal" features like voicing and vowel height.

A more accurate way to say this would be that if we abstract away from stress levels, we can (it was assumed) identify a segment for what it is independent of the context it is found in; if, however, we do not make this abstraction, and the stress level of a vowel determines in part just what sort of segment that vowel is, then identifying that segment as a (say) á (an /a/ with 2stress) requires looking at the context which the segment is found in.

This motivation for calling something "suprasegmental" is not the only possible one by any means. As Haugen puts it,

> Some sounds are thought of as occurring one after another, like bricks in a wall, while others occur simultaneously with these and usually span a number of the individual bricks at a time. (278)

These longer-than-usual bricks, or suprasegments, may be either (i) homogeneous in time, or (ii) non-homogeneous (or rather, not necessarily homogeneous) in time. Both theoretical options can be found in the literature: the clearest statement of the second position is found in Harris (1944), which we shall return to; the closest to the first position is Bloch (1948).

A third possibility arises if one works within a framework which provides a fixed inventory universally of phonetic features. This third possibility is that the longer-than-usual bricks, the suprasegments, may be changing in time, or dynamic, just in case there is a single phonological feature that, by itself, specifies this dynamic phonetic character. The structuralists we are looking at did not provide themselves with such a universal phonetic system, and so this option was not available to them. It is the position that we shall take, however, in this thesis, beginning in Chapter One. If there are no dynamic features -- which is possible, but not yet known for sure -- then this position is equivalent to (i) in the previous paragraph.

Yet the availability of phonological features is of more consequence than this matter of dynamic suprasegmentals. For in what is perhaps the most successful treatment of suprasegmentals within the structuralist framework, Zellig Harris', the two affairs (on our account, distinct) of "features" and "suprasegmentals" are reduced to the same thing. In "Simultaneous Components in Phonology", Harris maintains that, in effect, phonemes -- or "components", rather -- may occur at the same moment in time. These "simultaneous components", if all of the same length in time, play the role of Trubetzkoy and Jakobson's "features"; if one lasts longer than another, the longer one may be viewed as a suprasegmental.

The distinction between features and suprasegmentals will be central in this thesis, and so the attempt to reduce the two to one is of some real importance to us.

In his review of Trubetzkoy's *Grundzüge* three years earlier (1941), also in *Language*, Harris had rejected phonetic (read: distinctive) features as irrelevant to phonemic analysis. The only relevant criteria in determining the bredth and measure of a linguistic system, he maintained, were distributional.

> The types and degree of phonetic contrast (e.g. whether all the consonants come in voiced and unvoiced pairs) have nothing to do with the classification of the phonemes....It may indeed be interesting to work out patterns of the phonetic relations between phonemes and see how they compare with the distributional pattern. But that would be a new correlation, interesting for diachronic linguistics and for linguistic psychology, e.g., for the question: How do the physical (phonetic) differences within the ranges of phonemes (events to which people conventionally react uniformly) compare with the difference between different phonemes (events to which they react differently)? In synchronic linguistics, it is only the distributional pattern that would show what work each phoneme can do, what operations can be performed upon each, i.e., what its place is in the structure. (348).

This position changes somewhat by 1944, though by no means completely or always explicitly. In "Simultaneous Components", a method of analysis is introduced which "breaks all or most of the phonemes into new sub-elements (components). Each of the old phonemes will be a particular simultaneous combination of one or more of these new elements." (181). "It will be possible," Harris continues, "to select and symbolize the components in such a way as to show immediately the limitations of distribution, and in many cases the phonetic composition, of the phonemes in which they co-occur." (emphasis added). In this tack towards the phonetic are included "components which are precisely the length of a phoneme", i.e., "simultaneous breakdown of each phoneme by itself." This may "indicate the phonetic composition of each phoneme," as well as "eliminate phonemes of defective distribution." This technique, it is suggested, is of particular importance in separating off tone in tone languages. But if the technique is useful in distributional statements about tone-langauges (182), it may well be "profitable to continue this extraction [of components, i.e., simultaneous breakdown of segments -- JG] until all phonemes have been reduced to combination of components." (200). Such components may include "front, middle, back", the components which distinguish m, n, and ŋ (200); they may include that which /p,t,k/ have in common (Harris calls it "x", adding that it "may well be the combination of the unvoicing and the stop components") (200).

The crucial step comes next. "The physical movements of articulation may offer certain absolute factors common to various

phonemes: /p,t,k/ are generally voiceless, fortis, stopped. Since the components will in the last analysis have to identify articulatory as well as acoustic events [one wonders why? JG], it is desirable to reflect these as closely as possible." And it is on grounds of "phonetic simplicity" that tonal features are separated from vowels in tone languages (201). This remark might be contrasted with Harris' footnote, some pages earlier (p.183) that "the Prague Circle more closely approached the technique of dividing elements into simultaneous components, but purely on arbitrary phonetic grounds...."

On the whole, it would not be unfair, I think, to interpret Harris' attitude toward Trubetzkoy's "features" (and Jakobson's -- he is thanked in Harris' introductory footnote) as ambivalent and ambiguous. Harris' program is aimed at showing that the distributional techniques of traditional American phonology, if carried to what he suggests is their logical extensions, componential analysis, would lead the linguist to the useful, "non-metaphysical" part of the Prague conception of distinctive features. By resting with a familiar technique of classification by distributional restrictions, Harris could, he was suggesting, arrive at the useful part of the Prague school's feature-theory.

If finding simultaneous components is a discovery procedure for simplifying grammatical description (in Harris' basic sense of reducing the number of constraints that need be stated in a grammar on the distribution of basic elements), then the character of these components is essentially arbitrary, not phonetically constrained. These "long components" could then express relations between successive segments as well as internal composition of particular segments; these relations between successive segments, furthermore, can logically be most anything at all -- what might today be called morpheme-structure constraints, or the effects of phonological rule.

If this interpretation is right -- and I shall cite some of Harris' passages in support of this below -- then two quite different notions of simultaneous components are being presented at the same time. One conception, presented earlier in Harris' paper, sees them as phonetically meaningful; the second, as phonetically arbitrary.

Thus at the beginning of the article, Harris suggests that components have a "statable physical characteristic", and that they are "time stretches of sound (sound waves), or sets of simultaneous motions of the vocal organs [footnote: e.g,, vibrations of the vocal cords, giving 'voice'....)]"(186). Again, he says, "since components are to be physical elements, we must consider the phonetic value that they can have" (187).

This step toward phonetic, or feature, orientation is strikingly illustrated in the very first paragraph, in fact, where Harris gives an example of a componential analysis of /b/, consisting of "voicing

plus lip position plus stop closure, all occurring simultaneously."
And then, in a footnote to this apparently phonetic bent, Harris corrects
this impression: "This example of phonetic components is given
here only for introductory simplicity. The analysis presented
below is primarily distributional rather than phonetic."

Thus, by section 4.1 (187), "It is not required that components
have a constant phonetic value throughout their duration. A component
may have a phonetic value which changes in a fixed way in respect to
its end points: e.g., Falling tone, increase in nasality, voiceless.
beginning and voiced ending." Harris, parenthetically, makes it clear
that what have been called "intonational melodies" (see Chapter Three
below, for example) are, for him, single components, although he does
make reference to their status as lexical elements (in present termi-
nology). This is irrelevant, nonetheless, for the componential
analysis, according to Harris.

In the most detailed part of the article (see sections 5.2 to 5.4,
pp. 191 - 201), Harris spells out the uses of long components as
restrictions on distribution. He considers several examples, including
Grassman's law in classical Greek (permitting only one aspirate, under
normal conditions, per word), devoicing of consonants in German before
juncture, and the presence of clusters like /rtr/ (*partridge*) in
English, but not */trt/.

In the last example, he posits a long component

> extending over the length of a cluster and having
> the following phonetic values:
> > in first position, general consonant
> value (serving incidentally to distinguish consonants
> from vowels; this because any consonant may occur here);
> > in second position, continuant
> or /t,d/ if it follows a stop, otherwise general
> consonant value;
> > in third position, continuant if it
> follows a stop (but if the stop is /t,d/, this
> value only if a continuant precedes it), otherwise
> vocalic value.
> > The value [Harris continues] of the
> consonant-component permits any cluster of two
> except stop plus /p, b, k, g/, and then permits
> the third place to have continuant value (and to
> remain a member of the cluster) only if the pre-
> ceding two are continuant and stop; otherwise the
> component has non-consonant value and thus changes
> the third position into a vowel. This statement
> does not allow for clusters with middle /s/, as in
> *sexton*, and omits several details which would be
> taken care of in the other components for the
> individual phonemes. However, it is included here
> to show that even fairly complicated clusterings can
> be described by single components. (197-198)

I have quoted at length from Harris' work to show the phonetic
or feature-like constraints on long components are, in reality, nil.
Harris was aware of the ambiguity of his position, apparently; in
a footnote (numbered 22a, perhaps added after the first version), Harris
points out that there are indeed differences between the kinds of long
components he has been dealing with. But nothing comes of this
observation.

A few years later, Bernard Bloch developed "A Set of Postulates
for Phonemic Analysis" (1948), in which he axiomatizes a method for
determining the phonological system of a language from a set of
utterances. His system is of some interest to us because he
proposes that "a feature that is common to all the segments of a span
(or to all the segments that are open to it) is a feature of the span
itself, not of any segment in the span." (52.4, p. 36). His particular
statement is, curiously, inconsistent; if we define a feature of a
span in terms of the features of the segments of the span, we cannot
then say that the segments in that span do not have those features.
Nonetheless, Bloch's intent is clear; it would be more clearly
stated if Bloch had had recourse to another level in the grammar,
for example, at which spans of segments "had" features; his definition
would then be a definition of the relationship between this new level
in the grammar and the phonetic segmental level.

But while this line of analysis is temptingly similar in some
ways to the autosegmental we will develop below, Bloch's treatment
is ultimately quite different from ours, for the fact that there may
be spans with features is a fact, in his system, determined completely
by surface phonetics, and does not correspond to his notion of
suprasegmentals (it should be noted, too, that his use of the term
"feature" does not correspond to present use; his "aspects" are
rather more like our "features," but the difference is not crucial
here.) With respect to suprasegmentals, Bloch says,

> 48.2 Definition. Every distinctive aspect of stress
> or of pitch is a prosodic feature of the segment in
> which it occurs....
>   (The term "prosodic feature" is intended
> merely to name one kind of feature, as defined in
> section 47.2).
>   This kind of definition (admittedly
> ad hoc in its phrasing) allows us to abstract the
> qualities of stress and pitch from the totality of
> distinctive aspects in a segment, without obliging
> us at the same time to abstract certain other qualities
> that we do not usually wish to treat in the same way --
> for instance voicing and vowel color. For a language
> like Hindustani, where nasalization acts very much like
> pitch or stress in the more familiar languages Postulate
> 48 can be rephrased or supplemented so as to allow this
> quality also to be abstracted as a prosodic feature.
> For still other languages, where still other qualities
> act in a similar way, still other modifications or
> extensions of our postulates will be required.

If we abstract away from Bloch's disclaimer (p.46) that his set of postulates is not intended to reflect any truth, but simply to axiomatize an arbitrary but useful procedure ("To ask whether this is a 'correct' or a 'true' transcription of the given utterance is meaningless..."), we come quickly to the conclusion that Bloch has not provided a universal characterization of suprasegmental features. Indeed, it is the linguist's intuition when dealing with a new language that justifies adjustment of the fundamental postulates of phonemic analysis, as Bloch says in the quotation above. But the real task for the linguist must surely be rather to propose a set of criteria that determines when and whether certain features are or are not suprasegmental.

Our object, ultimately, will be to produce a notion of "suprasegmentals" which makes specific predictions about the behavior of these elements. We shall try to show real differences between features per se and supra- or (auto-) segmentals, such as the automatic spreading of autosegmentals over the longest domain possible (in a sense defined below in Chapter One). Features by themselves do not spread; they merely identify a segment for what it is. The domain of association of an "autosegment", on the other hand, does spread, quite automatically.

As I suggested earlier, it is difficult to compare this type of approach with the structuralists', since there were no phonological rules in their system. Certain things comparable can be found. In his discussion of the suprasegmental pitch phonemes of English (1945), Wells says (p. 31):

> Considered phonetically in relation to the segmental phonemes, the length of a pitch phoneme is variable. it can coincide in time with any number of segmental phonemes. Its effect will continue until replaced by another pitch phoneme.

This last statement comes close to our Well-formedness Condition (see Chapter One below); Wells is suggesting that the domain of the suprasegmental specification spreads itself until another suprasegmental specification is found. Wells places a left/right asymmetry on the spreading, since it only occurs to the right in this system, but this asymmetry is not argued for. He continues directly,

> A corollary of this statement is that a pitch phoneme cannot immediately follow itself. For instance, 22 would be merely a notational lapse for 2.

We shall meet this idea gain in Leben's (1973) Obligatory Contour Principle. It is important to recognize, however, that this is not, in fact, a corollary of any notion of spreading such as Wells suggests. Its truth, as an empirical hypothesis, is independent of how the domain of a suprasegmental is determined.

More broadly, the question of how suprasegmental elements inter-
act with each other, and interact with the segmental phonemes, was
not, to my knowledge, rigorously addressed. Most importantly, form
our point of view, what factors determined the placement or coordi-
nation of suprasegmentals with respect to segmental phonemes -- and
especially the changes that these coordinations may undergo -- was
not seriously dealt with in the structuralist tradition. There is
one exception to that, and this concerns the connection of English
intonation patterns with stress.

In this domain, it was frequently explicitly recognized that
there were, in Pikes terms, "contour points" where the intonational
patterns had special points of connection with syllables, or more
generally, with elements of the segmental phonology. Bolinger -- who
consistently evades classification as a structuralist -- most
clearly (1955, 1958) separated out the characteristic a syllable
may have as being the potentially prominent syllable from the point
of view of the shape of the intonation curve (called "stress") from
actually being prominent ("accent")

But having recognized that pitch contours in Engish are associated
with syllables that are stressed -- not, for example, with a syllable
whose vowel is a schwa -- the structuralist analyses went no
further; nor could it, as far as I can see. There is no room for
a distinction, for example, between a rule of feature-assimilation,
on the one hand, and true spreading of a suprasegmental element --
what we will call a "Flop" rule below -- on the other. Again as far
as I know, such a distinction is not necessary for the analysis of
English, at least to the level of detail that has been devoted to
it so far. But the distinction becomes vital in the analysis of
African tone languages, such as Igbo, in Chapter Two below.

As this type of distinction suggests, we are more interested
here in phonological, as opposed to phonetic kinds of analysis.
Again, it will be a feature's behavior with respect to phonological
rules that gives it away as suprasegmental, not any phonetic facts.
Ultimately, it would be extremely interesting to pursue the extensions
of autosegmental representation to phonetics, but this will not be
our goal here. Once again, the nature of the criteria used to prove
a feature "autosegmental" is something which translates very hard
into structuralist terms.

For our part, as I have indicated, we shall take our problem
to include the determination, for any particular language, of what
features are suprasegmental. By the criteria and within the system
developed in this thesis, "stress" -- as a multi-level, or as a binary
system -- does not satisfy the requirements for autosegmentalization.
If stress works as described in, for example, *Sound Pattern of
English*, then stress is not autosegmental and not, in our sense,
suprasegmental.

And yet much of the discussion of suprasegmentals in the American structuralist tradition of the Forties and Fifties included long disquisitions on stress as a suprasegmental. If stress is not "degree of loudness" (as has been one understanding) but is rather a means of modulating pitch, as Bolinger proposed, then this new sense of stress qua accent is a strong candidate for suprasegmental representation (see the section in Chapter Three below on English as a tone language).

The important literature on this subject is reviewed by Bolinger in section 3 of his (1958) "A Theory of Pitch Accent in English". Bolinger remarks,

> According to Bloomfield, "stress -- that is, intensity or loudness -- consists in greater amplitude of sound waves." Jones gives more attention to gestural accompaniment, but he too insists on "the objective impression of loudness."

In a footnote at this point, Bolinger continues,

> Pike (*Intonation of American English*) gives first place to intensity, but allows that duration may be a factor. Trager-Smith correlate stress with loudness, and Edith C. Trager says unequivocally, "there is only one component that matters -- loudness." (p. 21)

Our own position regarding the analysis of English is largely in agreement with Bolinger, and therefore in disagreement with the analysis of the American structuralists regarding the existence of several levels of stress as distinct from pitch.

If we accept, however, the view prevalent in the Thirties and Forties that stress is something closely akin to loudness, then of the classical three suprasegmental traits -- stress, length, and juncture -- none are suprasegmental within the autosegmental system. One might choose to conclude, therefore, that autosegmental and suprasegmental phenomena, as their names suggest, describe entirely different things. As the citations above from Harris, Wells, and Haugen suggest, this would be simply wrong from a historical point of view, for tone, vowel harmony, and nasal spans are thorough-bred suprasegmentals. But as the discussion in Chapter One will show, there are clear restrictions on what features may be autosegmental in a language, and we intend through these restrictions to produce a stronger theory, a theory which is not a theory, for example, of juncture.

Finally, the questions we must ask before we begin are:

(i) Why another theory of suprasegmentals?

For two reasons: first, because no totally satsifactory theory of suprasegmentals has been proposed within any framework yet, and second, because even the rudiments of a successful theory of supra- segmentals isn't to be found in generative phonology. Generative systems have not to date been able to deal with the kind of prosodic systems Harris' system (or its most similar relative across the Atlantic, Firthian prosodic analysis) could handle.

(2) What will count as a successful generative theory of suprasegmentals?

The theory will have to do two different types of things:

(a) make new (and presumably true) claims about the structure of representations at the phonological and/or phonetic levels; and

(b) explain how these different structures undergo phono- logical rules.

The bulk of Chapters One and Two is devoted to answering these questions.

(3) Abstracting away from the difference between structuralist and generative phonological analysis, how does the proposed theory of suprasegmentals differ from the structuralists' theories of suprasegmentals?

There is no satisfactory answer to this question, because ultimately it is not possible to abstract away from the different frameworks in which the theories have been put forward. In any case, autosegmental phonology, as it is developed here, attempts to deal with a proper subpart of the domain that was covered by the term "suprasegmental" in structuralist treatments. Our intention is thereby to give an empirically more contentful theory of these suprasegmental phenomena. In the end, the test of our success will be found in detailed language-specific studies, like that in Chapter Two, and in comparative studies, in line with some of the proposals in Chapter Three. For the linguist already committed to generative phonology, the arguments for our system are those in Chapter One.

CHAPTER ONE

AN OVERVIEW OF AUTOSEGMENTAL PHONOLOGY

1. Introduction

The subject of linguistics being the relation between sound and meaning, its first question must be what the nature and form of the phonetic and semantic levels are. For a generative linguistic system, this task begins with hypotheses about the type of formal representation that counts as a faithful rendering of the phonetic or the semantic aspect of a sentence, word, discourse, and so forth.

Autosegmental phonology is an attempt to supply a more adequate understanding of the phonetic side of the linguistic representation. Viewed in this light, it is a proposal at the same logical level as the idea that phonetic representation is a linear sequence of atomic units -- call them segments; it is at the same level as the suggestion that these atomic units are cross-classified by distinctive features. Autosegmental phonology is a particular claim, then, about the geometry of phonetic representations; it suggests that the phonetic representation is composed of a set of several simultaneous sequences of these segments, with certain elementary constraints on how the various levels of sequences can be interrelated -- or, as we shall say, "associated."

To say that autosegmental phonology is a hypothesis about the geometry of phonetic -- and ultimately phonological -- representations is rather abstract at best. From another, more down-to-earth, vantage point, autosegmental phonology is a theory of how the various components of the articulatory apparatus -- the tongue, the lips, the larynx, the velum -- are coordinated. That is, at the most superficial, observable level, the linguistic signal is split up into a large number of separate information channels. Viewed from the production side, this consists of the specific commands to the larynx, the velum, the tongue, and so on. At an "abstract" level, this information no doubt comes about from splitting up a more unified representation. All this becomes clearer with an example. Consider the word "pin." As the orthography suggests, the linguistic representation of the word consists of three segments linearly ordered, as in (1).

(1)
$$\begin{bmatrix} +\text{conson} \\ -\text{nasal} \\ +\text{labial} \\ -\text{coronal} \\ \cdot \\ \cdot \\ \cdot \\ \text{'p'} \end{bmatrix} \begin{bmatrix} +\text{syllabic} \\ -\text{nasal} \\ -\text{labial} \\ -\text{coronal} \\ \cdot \\ \cdot \\ \cdot \\ \text{'i'} \end{bmatrix} \begin{bmatrix} +\text{conson} \\ +\text{nasal} \\ -\text{labial} \\ +\text{coronal} \\ \cdot \\ \cdot \\ \cdot \\ \text{'n'} \end{bmatrix}$$

The production of this word, however, involves separate, though coordinated, activity by the velum, the lips, and so forth -- expressed roughly in (2).

(2)   A Score for the Orchestration of 'Pin'

Lips: ...Close up....open..........................
Tongue: ...High and front...........touch the palate
Velum: ..Raise.............Lower....................

The standard linguistic assumption regarding the nature of phonological representations -- that they look much as in (1) -- implies that the process of language acquisition and of perception includes the development of the ability to take a representation much as in (2) and slice it vertically into columns, assigning the appropriate feature specifications to each column, ultimately deriving a representation like (1): P-I-N. In short, the normal assumption about phonological representations implies that in processing a signal, we learn to shift around slightly the horizontal alignment of the commands in (2) -- we "justify" it -- and patch it up so that it may be sliced up vertically into the phonologically, and hence psychologically, real segments. Let us call this assumption the "Absolute Slicing Hypothesis."

In the example we considered -- saying "pin" as in (2) -- the process described above is essentially correct. The Absolute Slicing Hypothesis is adequate, then, to the level of detail we have considered in (2). Suppose, however, that we add to the orchestral score in (2) the activity of the larynx that gives rise to pitch. If I am uttering this word "pin" in isolation, and disregarding the actual sluggishness of the vocal folds for the nonce, the syllable will be uttered at a rapidly falling pitch. For reasons that will be developed in detail in Chapter 3, I will represent the Falling pitch as the sequence of a High pitch and a Low pitch. A more faithful version of (2), then, would be (3).

(3)   A Revised Orchestral Score for 'Pin'

Lips: ...Close up....Open..........................
Tongue: ...High and front..........touch the palate
Velum: ..Raise.............Lower....................
Larynx: ...High Pitch.....Low Pitch................

The assumption of the Absolute Slicing Hypothesis fails now, and this failure is in no sense a trivial one. The laryngeal pitch specifications are, we shall argue below, and in detail in Chapter 3, not the result of specification of any of the segments in (1), the phonological representation. The Falling pitch of this utterance is not part of the phonological segments in the same sense that the other commands in (3) (those for the lips, the tongue, or the velum) are.

The speaker of English thus factors out the pitch, and does not attempt to include the pitch features in the huge musical score which is sliced up into the more abstract segments as in (1). In this sense, then, the Absolute Slicing Hypothesis fails: the slicing is not absolute or complete, but rather may exclude some parts of the linguistic signal.
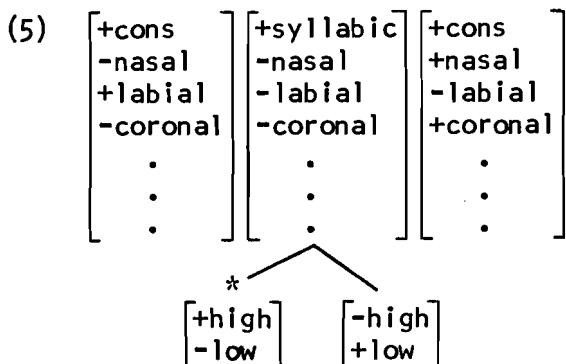
This failure of the Absolute Slicing Hypothesis is non-trivial in two separate, important senses. First, it is only an accident about English that the laryngeal pitch features are excluded from the Great Slicing. Other langauges may well include pitch as a part of the signal which is sliced up into successive segments. Conversely, a language may exclude some features from the Great Slicing that English happens to include. Guarani, as we shall see below, happens to exclude the nasalization feature, while English includes it. Whether a particular channel of articulation is included cannot be specified universally once and for all.

The second sense in which the Universal Slicing Hypothesis fails non-trivially will be the key to the successful formulation of its successor. The articulatory levels that are excluded -- in the present example, pitch -- are themselves susceptible to a "slicing" or segmentation. Thus, while the laryngeal pitch commands in (3) do not correspond to the standard phonological representation in (1), they do in fact correspond to a more abstract segmented level, as in (4).

$$\overset{*}{}$$
(4)   H L   where H = High and L = Low

In the present example, of course, the empirical significance of the representation in (4), including the meaning of the star *, remains open for the present. The important point is this, however: while a language may exclude an articulatory level from the Great Slicing that leads to the principal segmentation (corresponding to P-I-N in (1)), the excluded articulatory levels form themselves segmented domains, domains in which the segments are linearly ordered, and where the segments are cross-classified by feature specifications. In general, the formal properties of

the "phonological" representation, as in (1), will be mirrored on each level. Just as (1) corresponded to (2), then, (5) below -- the revised and now autosegmental representation corresponds to the more adequate orchestral score, (3). And (5), we see, is a synthesis of (1) and (4). The precise significance of the "association lines" linking the tonological and the phonological levels will become clearer shortly.

$$(5) \quad \begin{bmatrix} +cons \\ -nasal \\ +labial \\ -coronal \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \begin{bmatrix} +syllabic \\ -nasal \\ -labial \\ -coronal \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \begin{bmatrix} +cons \\ +nasal \\ -labial \\ +coronal \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

$$\begin{bmatrix} +high \\ -low \end{bmatrix} \quad \begin{bmatrix} -high \\ +low \end{bmatrix}$$

To develope this idea, however, let us turn our attention from these rather general comments on the nature of phonological segmentation to some more pressing problems in generative phonology. A proposed theoretical revision must in general be shown to resolve a dilemma in the standard theory; we shall therefore proceed to several general types of phenomena that are all difficult or impossible to account for in the standard framework, which assumes only one string of segments in a phono-logical representation. We shall show how each phenomenon of this type individually leads toward the same solution, auto-segmental phonology. In the process of doing so, we shall see that these phenomena in general, though not invariably, cluster together, as predicted by the autosegmental solution.

We shall consider:

(i) The existence of contour-valued features, such as rising or falling tones on vowels, or pre-nasalized stops, such as ^{n}d, ^{m}b, etc. Here the argument must be made that in some languages, short[1] vowels bear these dynamic or contour tones, and that furthermore such a contour tone is linguistically the concatenation of two level tones, rather than being an atomic unit itself.

The existence of such short vowels with contour-tones has been argued for recently by a number of linguists, most forcefully by Will Leben (1973). No satisfactory theoretical account has been given, however.

(ii) From contour-featured segments we shall turn to the phenomenon of "stability," or the tendency of a feature value to persist despite the erasure of the major segment (generally, vowel) which appeared to have born that feature. Roughly, that is, we find in tone languages that when a tone-bearing vowel is deleted by a phonological rule, the tone that was being borne does not delete also, but rather shows up elsewhere on a neighboring segment.

(iii) We shall consider next <u>melody levels</u> in the grammar: that is, linguistically significant levels in the grammar which refer to just one or two features in the utterance. We have already seen a foreshadowing of this in the discussion of the difference between the pitch features in English and the other features. If we view the musical scoring for the utterance "pin" as in (3), and many more like it, we will notice a difference in the different lines (or "voices," in the musical sense). Some, like the line of instructions to the velum legislating nasality, will appear rather uninteresting when viewed alone. There are no generalizations to be made by looking at nasalization alone in English. On other lines -- such as the level of instructions for pitch -- there are linguistic generalizations to be gained by restricting our attention to just those features. There are, that is, strictly pitch-feature regularities; -- which is not to deny that further insight is to be gained by observing how the pitch-level is coordinated with the other levels, of course. The point here, however, is that certain subsequences of features <u>do</u> form linguistically significant "melody levels," while others <u>do</u> not.

(iv) Then we will see how the heuristic notion of "floating tone" can be rigorously defended in autosegmental terms. The floating tone has served well in practical terms for linguists dealing with tone languages; it has, however, remained with a tainted reputation because of its apparent anomalous nature from a theoretical point of view.

(v) Lastly, we will consider some processes of automatic spreading of features both to the left and right, over segments unspecified for those features. The accepted notation for phonological rules implicitly predicts that spreading should be simpler if it occurs only to the left or only to the right; we find, however, many cases of bidirectional spreading. When viewed from an autosegmental perspective, the reasons become clear.

This is, perhaps, an appropriate moment at which to point out some differences and similarities between autosegmental phonology and other treatments of what have been called "suprasegmentals." Much of what is covered in this present analysis would be called "suprasegmental" by the criteria implicit in the literature. What is suggested by the term "suprasegmental"? Two things, I would submit, which are entirely different but which have been continually confused in the recent linguistic tradition.[2] Calling tone "suprasegmental," firstly, immediately distinguishes it from the "segmental" -- that is, it correctly views pitch as different from and not part of the phonological segmentation as in (1) -- the segmentation into phonemes, roughly speaking. But this first observation has wrongly led to the assumption that "suprasegmentals" could not be segmental in their own right -- the second, more fundamental sense of "segmental," which is overlooked by the term "suprasegmental." If the "suprasegmental" of pitch, for example, does itself form a sequence of tonal segments, then "suprasegmental" is a misleading label. A more accurate picture, we are suggesting, is parallel sequences of

segments, none of which "depend" or "ride on" the others. Each is independent in its own right: hence the name <u>auto</u>segmental level.

2. Contour Tones

Let us proceed, then, to particular problems within the theory of phonology. Our first task is to make quite clear the reason why two tonal specifications on a single vowel is incontrovertibly in contradiction with the standard theory.

I shall first sketch the argument heuristically, and then redo it in more technical language.

Suppose we analyze the tonal workings of a language and decide that where contour tones occur, they are really the concatenation of level tones.[3] That is, Rising tones are actually composed of the sequence Low tone and High tone; Falling tones are composed of High tone and Low tone. Suppose further that such contour tones occur on short vowels. We shall consider several actual examples shortly. A long vowel, we know, may be analyzed as a sequence of two short vowels, but a short vowel has no such analysis. How can we represent the contour tone of the short vowel?

Our conclusion will simply be that we cannot represent it if we stay within the assumptions of the standard theory. Let us assume, for purposes of exposition, that tone can be represented in terms of binary features[4] (although the assumption of the binary nature is not strictly relevant at this point). Suppose then, we try to represent /a/ with a Falling tone -- that is, /â/ -- a High tone and a Low tone.

$$(6) \quad \begin{array}{l} \\ \\ \\ \\ \\ a\{ \\ \\ b\{ \\ \end{array} \begin{bmatrix} \text{+syllabic} \\ \text{+constricted} \\ \quad \text{pharynx} \\ \text{-high} \\ \text{-round} \\ \vdots \\ \text{+Highpitch} \\ \text{-Lowpitch} \\ \text{-Highpitch} \\ \text{+Lowpitch} \end{bmatrix}$$

Now this clearly won't do. This curious segment (6) is both +Highpitch, and -Highpitch; it is both +Lowpitch and -Lowpitch. The features in the part labeled (a) seem to indicate a High tone; those in (b) seem to indicate a Low tone. But putting one on top of another does not order them, as we require. In particular, we must distinguish a Falling tone from a Rising tone; but to do so, a new and theoretically momentous set of conventions must be introduced in order to make <u>vertical</u> ordering of features within a segment play the role of left-to-right (temporal) ordering of segments elsewhere.

Or one could attempt an equally radical revision of the notion of segment with the introduction of a notation as in (8).

$$(8) \quad \begin{bmatrix} \quad\quad\quad \text{+syllabic} \\ \quad\quad\quad \text{+CP} \\ \quad\quad\quad \text{-High} \\ \quad\quad\quad \text{-round} \\ \\ \begin{bmatrix} \text{+Highpitch} \\ \text{-Lowpitch} \end{bmatrix}_{A}^{A} \quad\quad \begin{bmatrix} \text{-Highpitch} \\ \text{+Lowpitch} \end{bmatrix}_{B}^{B} \end{bmatrix}$$

One has here, however, what is to all appearance a category error. The non-tonal features have one relationship to the entire segment: they are, let us say, "features-of" it. The tonal features are features of a sub-segment (there are two subsegments here, A and B); and the sub-segments bear some entirely other relationship to the entire segment: they "subcompose" it, let us say. Consequently, tonal features are "features-of a segment which subcomposes" the entire segment; in particular, the tonal features are no longer features-of the segment. Here, too, then is a radical suggestion intended to make possible the representation of a contour tone.

Now these two attempts -- the second of which is rather a commonplace in the literature -- are not misguided; quite the contrary, even though both are wrong. They recognize that the existence of contour tones is a _problem_ for the standard representation, and problems are the stuff of which advances are made. Our task is to outline a third proposal, one which is conceptually simpler and which has a number of direct empirical consequences: autosegmental representation. First, however, I would like to review the argument, in more technical terms, that demonstrates the dilemma the standard theory finds itself in.

Segments are atomic elements ordered linearly left to right; this simple and elegant hypothesis lies at the foundation of standard phonology. A feature-specification of a segment is, then, a property or an attribute of the segment. The property of being voiceless "belongs," we may then say, to the first segment in "pin," just as the property of being "+front" belongs to the second segment. More technically, we would say that "pin" has a representation at the phonological level as in (9).

(9) $\check{S}_{75}$ S_{31} S_{33}
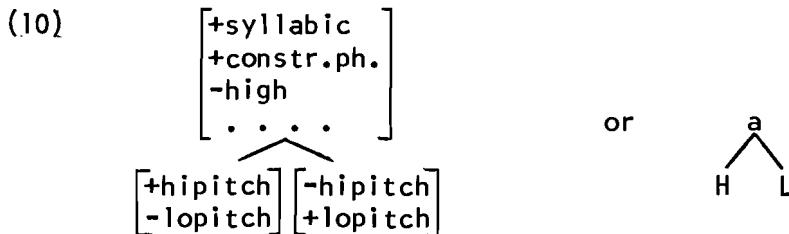
The subscripts indicate nothing but that these are primitive elements in the phonological vocabulary. We would further specify that there are various "feature-projection" maps -- characteristic functions, in mathematicians' language -- that tell us what the feature specifications for each segment are. For example, there is an F_{voice} function which maps S_{75} in the present example to - (minus). $F_{voice}(S_{75}) = -$. All this amounts to saying that the first segment in the representation of _pin_ in (9) is voiceless.

Having made precise our notions of segment and feature-specification (following essentially Chomsky's _Logical Structure of Linguistic Theory_), it is clear that there is no way that a segment can be specified both + and - for some particular feature. In particular, no single segment can be contour-toned unless there is a single feature which is "Falling" or "Rising," contrary to hypothesis. If a single segment were both +Highpitch and -Highpitch, then $F_{highpitch}$ would not be a function, again contrary to assumption, since it would be two-valued.
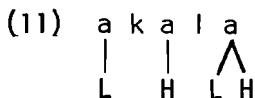
If we wished to resolve this formal conflict by saying that there were two separate characteristic functions associated with each phonological feature, one mapping the specification + (plus) and the other to the specification - (minus), we would avoid the immediate contradiction, but lose the formal representation of the fact that the two values are values of the same feature; such an approach would clearly be the wrong direction to proceed in.

No, the correct position to take here is that, indeed, the
formalism demands that no segment may be doubly specified for a
feature. If tones are features of the vowel segment, then, there
may be no Low-High sequences on a single vowel, and that's all
there is to it. Some other basic axiom must be abandoned.

The formal approach we will take toward a resolution of
this problem is to deny that tonal features are features of the
vowel in the case of the problematic contour tone. Rather, the
tonal features are factored out to another level; feature speci-
fications on the other level constitute segments, but their
relation to the vowels with which they are associated is merely
one of simultaneity in time. We represent this dynamic element
by association lines. The falling-toned /a/, then, is repre-
sented as in (10).

(10)
$$\begin{bmatrix} \text{+syllabic} \\ \text{+constr.ph.} \\ \text{-high} \\ \quad . \ . \ . \end{bmatrix}$$
$\begin{bmatrix} \text{+hipitch} \\ \text{-lopitch} \end{bmatrix} \begin{bmatrix} \text{-hipitch} \\ \text{+lopitch} \end{bmatrix}$

or

$$\overset{\displaystyle a}{\underset{\text{H} \quad \text{L}}{\wedge}}$$

A more complex example, like the word àkálá would be represented
as in (11).

(11)   a k a l a
       | | ∧
       L   H   L H

Before proceeding with the development of this notation, let
us turn to an example from a tone language.

2.1

Our first example is from Qhųhų Igbo, as described in Green
and Igwe (1963). These facts will be discussed in much greater
detail in Chapter Two. For the present we will content ourselves
with finding a contour tone on a single element, and showing that
it is indeed the concatenation of two tones; we shall then see why
the notation as in (10) or (11) does more than represent these facts,
but predicts and explains the intuitively clear sense of left-to-
right order of tones that is revealed inside of contour tones by
the tonological rules involved.

There is a simple tone rule that occurs in the "I Root"
form. Throughout the Igbo language, the particular tone of the
verb stem, suffixes, and prefixes is determined by the "form" of

the clause. We shall devote Chapter 2 to an investigation of the
details of this system; for the moment, we may focus our attention
on the I Root form, in which the verb-stem is Low-toned, in non-
exceptional cases. Consider, for example, a simple sentence in
the I Main form where the subject is a pronoun:

(12) I Root form

(a) ọ́ cị̀ àkhwá   "He must carry some eggs" or
                             "He was carrying some eggs"

he carry eggs

(b) ọ́ zà ụ́lọ̀   "He must sweep the house"

he sweep house

Ḿ cị̀ ánụ́   "I was carrying meat"

I carry meat

We see that in the I Main form, the pronominal subjects are
H in tone, and the verb stem is uniformly Low in tone. When the
subject noun phrase (NP) -- not a pronoun -- ends in a low tone,
the tone pattern of the sentence is just what we would expect on
the basis of the sentences in (12). The verb stem is Low; the
subject bears its inherent (isolation) tones.

(13) Ézè cị̀ àkhwá   "The chief was carrying eggs"

Chief carry eggs   Eze = chief (HL)

Ùwà cị̀ àkhwá   "Uwi was carrying eggs"

Uwa carry eggs   Uwa (a name) (LL)
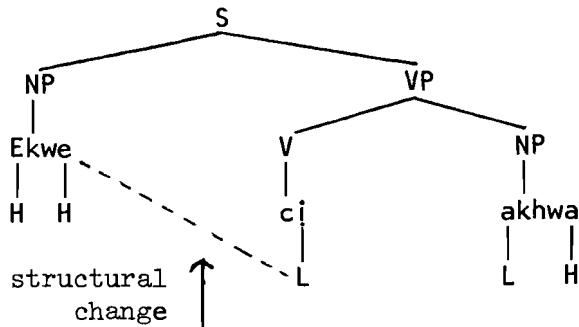
However, when the subject NP would normally (i.e., in iso-
lation) end in a High tone, then it undergoes a slight tonal
change here. As we see in (14), the final H becomes a Falling
tone.

(14) Ékwê cị̀ àkhwá   "Ekwe was carrying eggs"

                           Ékwé (a name) (HH)

Àdhâ cị̀ àkhwá   "Adha was carrying eggs"

                           Àdhá (a name) (LH)
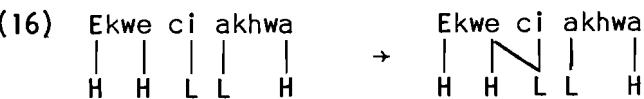
What is happening is clear if we watch the tonal melody. In
(14), the drop from the High on the last syllable of the subject
NP to the Low of the verb stem is shifted leftward, or "anticipated"

on the last syllable of the subject. Using the autosegmental notation, this change is represented as in (15), where the dotted line indicates that the association line was added. (16) separates the Before and After states; it says the same thing as (15) in more familiar but less perspicuous notation.

(15)

```
                              S
                    _____/  _____
                  NP                      VP
                  |                    __/  \__
                Ekwe                 V          NP
                |  | `~~              |          |
                H  H    `~~          ci        akhwa
                                      |         |  |
            structural  ↑      `~L     L        L  H
              change   _|
```

(16) Ekwe ci akhwa   Ekwe ci akhwa
      | | | | | → | |\| | |
      H H L L   H   H   H L L   H

The point to observe is that the final syllable in *Ekwe* is associated with two tones -- H and L -- in the derived structures; this is what is identified as a contour tone.

Before proceeding to the behavior of the irregular verbs in the I Main form, let us see what this autosegmental notation has provided us with. By setting the tonal segments or tonemes off on a separate autosegmental level, we were forced to introduce "association lines" to coordinate the two levels in time. We hoped, thereby, to represent contour tones, which is <u>prima facie</u> a phenomenon purely internal to the vowel in question. The notation captures precisely, however, the fact that the tone associated with a vowel on the right -- here, the verb stem -- may associate with a vowel neighboring on the left, causing a change in the <u>right-hand</u> side of the latter vowel. Thus we find processes like (18a), but not as in (18b), as the notation predicts.

(18) a. V́ V̀ → V̂ V̀

      b. V́ V̀ → V̌ V̀

The difference between two such rules is intuitively plausible, and in fact borne out in empirical work in tone languages throughout the world. (18a) is in some sense a proper "assimilation"; (18b) is not. The representation in (16) makes clear what that sense is; alternative proposals, such as perhaps (6) or (8), do not.

Having called the tonal phenomenon in (14) an assimilation, an having characterized it as a "flop" rule as in (16), we have

committed ourselves to a view with certain predictions. If the verb stem should be on a Mid tone, for whatever reason, then the falling tone of the subject's final syllable must fall to Mid, rather than to Low. This is in fact true, and is borne out in three separate cases, discussed in Chapter 2. Here I shall just mention one representative example.

A Low-toned suffix immediately following the stem causes the stem to raise to Mid, as in (19a). The pitch of the utterance is as drawn; the autosegmental representation is as in (19b).

(19) a. Hã dìrì n'ǫ́r̃ǫ́   b. Hã diri n'or̃u

H M L   H H
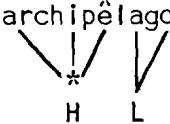
where V̂ means "High falling to a Mid which is level with the following vowel, which must be Mid." The reason for this compli-cated statement lies in the nature of Downstepping tonemes (the Mid here is one); these are discussed in Chapter 2.

2.2 English

Our next example is from English, a language that may feel more familiar. The neutral intonation pattern in English is H̆L or MH̆L, with the High tone associated with the accented syllable; this, in fact, is the significance of the star *.
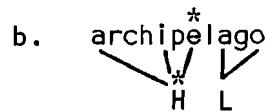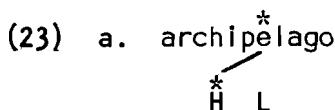
Consider a polysyllabic word like "archipelago." It has a tone pattern roughly as in (20).

(20) archipélago   archipélago

H   L

This is derived by linking the starred elements in either level; the Well-formedness Condition of autosegmental theory then comes into play and does the rest. Thus we start with a representation as in (21), with no linkage between the levels; then rule (22) comes into effect, creating (23a). The Well-formedness Condition creates (23b) from (23a).

(21) archipélago   (22)   *
                              V
                              •
                              •
     H̆ L   •
                              *
                              T

In (22), as elsewhere, "T" stands for any arbitrary toneme, and "V" for any arbitrary vowel; a dotted association line indicates that the rule adds that line.

(23) a. archipélago b. archipélago

  H L H L
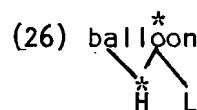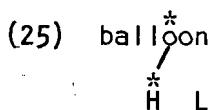
(24) <u>Well-formedness Condition</u> (initial statement)

(1) All vowels are associated with at least one tone;
All tones are associated with at least one vowel.

(2) Association lines do not cross.
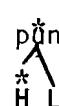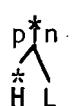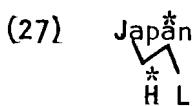
Note that the Well-formedness Condition is in the indicative, not the imperative. A derivation containing a representation that violates the WFCondition is not thereby marked as ill-formed; rather, the Condition is interpreted so as to change the representation minimally by addition or deletion of association lines so as to meet the Condition maximally. In the case at hand, (23a) is changed to (23b) minimally by addition of four association lines.

Holding aside one or two small points until the extended discussion in Chapter 3 (Section 3.2), the description just given accounts accurately for the assignment of neutral intonation to English words. So far no contour tones have been encountered, but that is because the accented syllable was not final in the example considered. If we consider a word like balloon, we get a structure as in (25), corrected to (26) by the Well-formedness Condition.

(25) balloon (26) balloon

  H L H L

What the theory is predicting about English now is something curious, something more than a bit reminiscent of tone languages, but something nonetheless true: non-final accent is realized as a High tone (as in (24)); final accent is realized as a Falling tone.
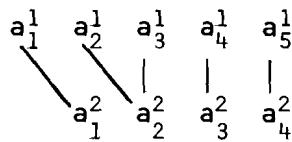
This conclusion is true independent of the length of the final vowel, too; so our search for contour tones on short vowels has led us home to English with a clear example. See (27) (the reader will recall that were the final vowels in (27) long, they would undergo diphthongization and Vowel Shift).

(27) Japán hát pín pún

  H L H L H L H L

2.3 Excursus On Formalism

With the initial statement of the Well-formedness Condition (24) for tones given, the central aspects of the autosegmental theory are clear. In this excursus, I would like to restate some of these provisions formally. A formalization of these ideas is simple, almost too simple, in that it requires a certain amount of familiarity with the concepts involved before the axiomatization makes as much sense as the simple diagram notation.

Be that as it may, let us proceed. Each autosegmental level is a totally ordered sequence of elements, a^i_j: this is the jth element on the ith level. Call the set of segments on the ith level level L^i. We shall restrict our attention to the case with two levels for now. So we have:

$$
\begin{array}{ccccc}
a^1_1 & a^1_2 & a^1_3 & a^1_4 & a^1_5 \\
& & | & | & | \\
a^2_1 & a^2_2 & a^2_3 & a^2_4 &
\end{array}
$$

In addition to these two sequences of segments, there is a totally ordered sequence of pairs -- essentially the association lines, from the geometric point of view.

$$(a^1_1,\ a^2_1)\ (a^1_2,\ a^2_1)\ (a^1_3,\ a^2_2)\ \ldots$$

Call the set of these pairs "A."

It is clear that A in a sense organizes the other levels. Define projection π_1 from 2^A (the set of subsets of A) to 2^{L^1} (the set of subsets of L^1) in the natural way:

$$\pi_1(\{(a^1_j,a^2_k),(a^1_l,a^2_m),(a^1_n,a^2_p),\ldots\}) = \{a^1_j,a^1_l,a^1_n,\ldots\}$$

That is, the 1st projection π_1 picks out the set of first elements of the pairs. Likewise, the second projection π_2 picks out the second element of each pair, and so forth. Define the inverse of these projections: π_1^{-1} may be defined as follows (or any other way that makes clear that π_1^{-1} is the inverse of π_1 in the natural sense):

π_1^{-1} is the largest subset α of A such that

$$\pi_1(\alpha) = a.$$

Then the Well-formedness condition is:

(WFC) π_i and π_i^{-1} preserve connectedness.
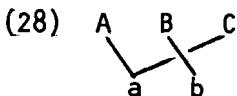
Recall that we have a notion of connected already defined on each level L^i and A since we assumed they were linearly ordered. This ordering naturally induces a connectedness in which the set circled in (26) is connected, but the one in (27) is not, and so forth.

(26) a b (c d e) f   (27) a (b) c (d e) f

I will briefly run through why the statement of the WFC above has the effect of the Well-formedness Condition in (24).

Suppose we have a situation as in (28).

(28) A   B   C

        a   b

This violates (24-2); we shall see that its corresponding formal representation does not have a proper projection operator preserving connectedness.

L^1 for (28) is, evidently, A B C. There are, rather trivially, 7 connected subsets: ∅, A, B, C, AB, BC, and ABC.

Similarly, L^2 is a b, with four connected subsets: ∅, a, b, and ab. We will see that there is no formulation of the set A with the correct properties. There are only two candidates: (29) and (30). Consider (29) first.

(29) (A,a) (B,b) (C,a)

Since {a} is a connected set, $\pi_2^{-1}(\{a\})$ should be connected. $\pi_2^{-1}(\{a\}) = \{(A,a),(C,a)\}$, which we see is not connected in (29). Contradiction.
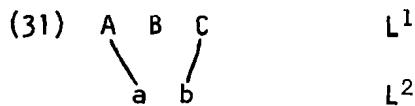
If we try (30), we get a similar result.

(30) (A,a) (C,a) (B,b)

Now $\pi_2^{-1}(a) = \{(A,a),(C,a)\}$ __is__ connected.

However, in (30) $\{(A,a),(C,a)\}$ is connected, so $\pi_1(\{(A,a),(C,a)\}) = \{A,C\}$ should be connected -- i.e., A and C should be neighboring. But they are not; contradiction.

Thus, by extension, any situation like (28) where lines cross must be ill-formed; and so (24-2) is derived.[5] Similarly in (31), which violates the first part (24-1) of the Well-formedness Condition, we find WFC violated also.

(31)   A   B   C L^1

         \   /

          a   b L^2

If π_2^{-1} and π_1 preserve connectedness, then $\pi_1 \circ \pi_2^{-1}$ must preserve connectedness too.   But $\pi_1 \circ \pi_2^{-1}(ab) = AC$, which is not connected (although ab is); contradiction.   The same argument holds if a = b. Thus WFC implies (24-1).[6]

3.   Stability

The second type of phenomenon we shall consider in motivating autosegmental representation is what I have called "stability." In tone languages, we find that when a vowel desyllabifies or is deleted by some phonological rule, the tone it was bearing does not disappear -- rather, it shifts its location and shows up on some other vowel.   The toneme, or more generally the tone melody, has a stability which maintains it independently of the other aspects of the signal, and thus the tone melody preserves itself despite modifications to the syllabic structure.

Reference to this type of phenomenon in the literature has generally been associated with the notion of "conspiracy" or derivational constraint; in this case, a derivational constraint or conspiracy to move around the tonal specifications from vowel to vowel in order to find, on the surface, the same tone melody that was there underlyingly.   This is not to say that there are no tonal rules that delete or modify the tone melody: there surely are such rules.   But the normal case is where the tone melody survives the effects of phonological rules.

If the tone of a vowel is specified by its features, then the pitch of a vowel is just like any other of its characteristics, like its tenseness or its roundness.   If a phonological rule should delete that vowel, then its tonal specifications are deleted along with all of its other properties.   Suppose we have a phonological rule deleting a vowel as in (33).

(33)   V $\rightarrow$ Ø / - V   V-Deletion

(A common Bantu rule -- see, e.g., Spa, 1973:78 rule (32),) However, we need to save the tonal information of the deleted vowel, because it shows up on the surface.   Looking at tone as a feature of the vowel, we could do this in one of two ways, similar in intent.

(1)   We could posit a special "Tone Copy" rule which copies the tone of the to-be-deleted vowel onto its neighbor -- we could do this, I should add, if we permit two tonal feature-specifications inside a single vowel segment, the position I have just argued , against in section 2.   But in order to consider the logic of

"stability" independently of that of "contour specifications,"
let us permit the latter for the moment to be expressed as in (34)
below.

(34) Tone Copy

$$
\begin{bmatrix} V \\ \alpha hi \\ \beta lo \end{bmatrix} \begin{bmatrix} V \\ \gamma hi \\ \delta lo \end{bmatrix} \rightarrow \begin{bmatrix} V \\ \alpha hi \\ \beta lo \end{bmatrix} \begin{bmatrix} \overset{V}{\begin{bmatrix} \alpha hi \\ \beta lo \end{bmatrix}} \begin{bmatrix} \gamma hi \\ \delta lo \end{bmatrix} \end{bmatrix}
$$

A typical derivation applying Tone Copy and V-Deletion would be
as in (35)

(35)   ...à í ...

                        ...à ǐ ...   Tone Copy

                        ... ǐ ...   V-Deletion

    (2) Second, we could posit a general "derivational constraint"
to apply to all tonal rules either particular to the language, or
as a general condition on the application of rules -- this is the
approach Spa takes in his grammar of Enya, a Bantu language. He
suggests, "when a segment carrying a High tone is deleted or be-
somes incapable of carrying a tone, the High tone is transferred
to the nearest syllabic segment... ((This constraint)) applies each
time any rule whatever meets its structural description" (Spa,
1973:139 et passim). In fact, the correct statement of his
derivational constraint should apply to preserve equally both
High and Low tones. This modification both simplifies his phono-
logical system and generalizes the derivational constraint.

    Solution Number 2 is explicitly global if taken as a language-
particular rule, and therefore suspect within received generative
theory: a theory countenancing global rules approaches vacuity.
This solution introduces a general global condition on vowel-
affecting rules, and while this seems like an improvement, in that
it is a generalization, it is nonetheless worse theoretically
because we now permit not only global rules, but a whole new kind
of object which is global and applies anywhere during a derivation,
outside the set of ordered rules. As a general condition on rules,
it meets no such criticism, but it stands as an admission that the
original framework does not adequately model the behavior of the
segments under the application of rules. Solution (1) holds out
promise, and yet what we find in actual work is that for every
case of vowel-deletion or desyllabification, we must set up
another case of tone-copying, and what we have then is a missing
generalization. The generalization is precisely solution (2), the
general derivation constraint. A paradoxical situation: to
effect a satsifactory linguistic solution, we need to state a
generalization; but inclusion of this generalization within the

standard theory amounts to either a serious weakening of the theory of phonology, or an inadequacy in the standard representation.

We might note that even if we did include the derivational constraint, in the belief perhaps that constraining a theory must always play a more minor role than stating a generalization, three important questions would be left unanswered: first, why are tonal features copied, but not the other features? What makes them special? Second, what is meant by a representation with two feature specifications inside a single vowel (the issue pursued in section 2)? Is there indeed a connection between the fact that there can be contour tones and the phenomenon of stability -- apparently so different? Third, and most telling, the "conspiracy" to preserve tonal melodies extends past a derivational constraint that whisks the tone off of a sinking vowel: in fact, in a tone language where the Derivational Constraint seems to generally hold, what we find is that vowel assimilation rules like (36) copy all vowel features up to, but <u>not</u> including, tone features.

$$(36) \quad V \rightarrow \begin{bmatrix} \alpha High \\ \beta Back \\ \gamma Round \\ \delta ATR \\ \cdots \end{bmatrix} \quad / \underline{\quad} \begin{bmatrix} \alpha High \\ \beta Back \\ \gamma Round \\ \delta ATR \\ \cdots \end{bmatrix}$$

A rule like (36) certainly exists in Igbo and Yoruba, and in Enya according to Spa (pp. 47, 57). So if two vowels come together, each with its own tone, then either one vowel is deleted and its tone gets retained, as in (33), or one assimilates in quality in every regard save tone. The only empirical difference lies in the length and syllable quality of the remaining vowel(s). From the point of view of tone -- and its conspiracies -- the same fate has come to pass. Yet the derivational constraint speaks only to the case with deletion, not the case of nearly-complete assimilation: thus missing the generalization.

This is the logic of the situation. Let us look at some actual cases in more detail.

(A) Consider two articles by Julie Lovins (1971a,b) on Lomongo, whose tonological rules, she suggests, "conspire, individually or in concert, to derive surface tone patterns on words and phrases without changing the underlying melody." Central to the analysis is what Lovins calls "tone composition," in which "the tones stay where they are when segmentals are deleted." She continues with an example, "if two vowels are juxtaposed, within a word or across word boundary, it is usual for the first vowel to be elided. Its tone remains and combines with that of the following vowel." For example,

(37) bàlóngó bǎkáé → bàlóngǎkáé   'his book'

 bánà bǎmǒ   → bánǎmǒ   'other children'

 bǒmǒ bòtámbá   → bǒmǒtámbá   'another tree'

 bǎtswá là èmí   → bǎtswêmí   'you who lead me away'

     With a number of similar examples, Lovins concludes, "The only
derived forms that occur are the ones that preserve the underlying
melody... and the only way to get these derived forms is to posit
a species of rule application that many linguists find objectionable."
She is certainly correct, given the standard framework, and she is
exceptional among writers on this subject in recognizing the
implications for phonological theory of the type of rule she posits.
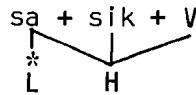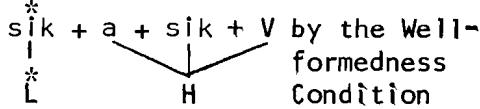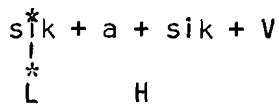
     The existence of the tone melody's "stability" is our concern:
how can it be that a tone refuses to be deleted when its vowel is
deleted: In autosegmental formalism, this is precisely what is
predicted. In any theory of generative phonology, and this one is
no different, a deletion rule deletes a <u>segment</u>. Now, if a rule --
(33), for example, V-Deletion -- should delete a vowel, it does not
delete any of the tone segments that the vowel is associated with,
since those tone segments are quite separate segments. The worst
that can happen is that the tone segments will be left "orphaned"
or free, without a vowel associated with it. That will be the
interesting case to look at in detail.

     The point we have just seen should be emphasized: the
stability phenomenon, formerly paradoxical, is a natural consequence
of the autosegmental system -- not by proposing a constraint on
rules, but rather by proposing in effect a new geometrical shape
(in a somewhat abstract sense) for formal representations.

     Let us consider in more detail the reduplication treated in
Lovins' papers. Verbs are lexically marked for tone, H or L;
the stem is reduplicated and an /a/ infix is added between the
two copies of the stem. An L or H desinence then follows.
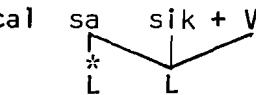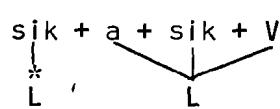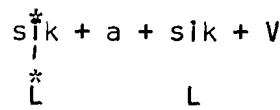
(38) L-toned stem /sik/ 'stop'

H-toned desinence   L-toned desinence

sĭk + a + sik + V   sĭk + a + sik + V
| underlyingly |
L   H   L   L

sĭk + a + sik + V by the Well-   sik + a + sik + V
| ＼＿／   formedness | ＼＿／
L   H   Condition   L '   L

sa + sik + V   Phonological   sa   sik + V
┌＿   ＼＿／   rules   ┌＿   ＼＿＿
L   H   L   L

[săsíkV̌]   [sàsìkV̀]
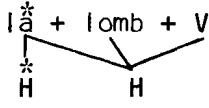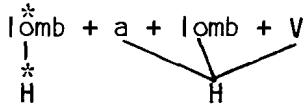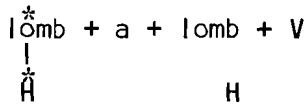
The last stage is reached by pure phonological rules:
k → Ø and ia → a.
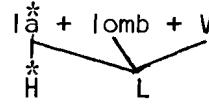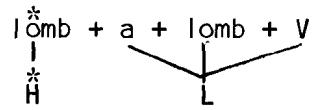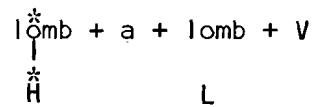
(39) H-toned stem /lomb/ 'be shy'

H-toned desinence   L-toned desinence

lŏmb + a + lomb + V   lŏmb + a + lomb + V
| |
H   H   H   L

lŏmb + a + lomb + V   lŏmb + a + lomb + V
| ＼＿／ | ＼＿／
H   H   H   L

lă + lomb + V   lă + lomb + V
┌＿   ＼＿／   ┌＿   ＼＿／
H   H   H   L

[lálómbV́]   [lâlòmbV̀]

In short, from the notation we get the "conspiratorial" results
automatically by keeping the syllabic and the tonal levels formally
separate. It may be noted that we get the desinence-tone spreading
automatically, too, as well as a formal understanding of the notion
"contour tone." Furthermore, the process of total vowel assimilation
-- construed as, e.g., (40) -- has the desired property of copying
all features up to, but not including, tone features, since tone
features aren't features of vowels.

(40)   X • +syllabic • +syllabic • Y

     SD: 1 ˇ   2   3   4 →

     SC: 1   3   3   4

The resolution of the paradox in tonal stability was derived
from viewing tones as segments on an equal rank with "phonological"
segments. This parallelism can be pursued; in fact, we find in
general perfect formal symmetry between the two levels. The "dual,"
then, of vowel deletion would be tone-deletion, followed by re-
association with the deleted tone. This in fact occurs; see the
discussion of the II Main form in Igbo (Chapter 2), and Sanskrit
(see May and Goldsmith (1975)).

4.1 Melody Levels

The third type of argument for autosegmental theory comes
from the existence of "melody levels." As explained above on p. 20,
these are linguistically significant levels in the grammar which
refer to just one or two features in the utterance. *Faute de
mieux* this has been taken sometimes to indicate tone features as
"features of an entire morpheme," in some unexplained sense.

Let us begin with an example from Mende, a Mande tone
language. The analysis is due to Leben (1973).

On short vowels in Mende, we can find Low, High, Rising,
Falling, or Rising-Falling tones. Morphemes are one to three
syllables long, and if the distribution of tones over these
syllables were random, we would expect to find five tonal classes
of 1-syllable words, 5^2 or 25 classes of 2-syllable words, and 5^3
or 125 classes of 3-syllable words: 155 types in all. In fact,
there are 5 classes for each, not 5^n, and they are of a very par-
ticular sort. Leben explains this by proposing that there are
only 5 available underlying melodies in Mende, and that the melody
is mapped from left to right onto the word. The five possibilities
are:

(41)   H   pélé kɔ́

        L   bèlè kpà

       HL kényà mbû

       LH nìká nàvó mbǎ

     LHL nìkílì nyàhâ mbǎ

Such an analysis, we might note, supports the contention that
contour tones are the concatenation of level tones, and that short
vowels (and these in (41) are) may bear several ordered level tones.

4.2 Tiv Verbal System

The next example is somewhat more complex, and more interesting. The Tiv verbal system has been the subject of a sequence of more and more refined analyses, including D.W. Arnott's exposition (1964), McCawley's reanalysis (1970), and Leben's proposal (1973). The present reanalysis departs from Leben's in one particular way: I shall indicate that the inclusion of Leben's principle (which I shall call the Obligatory Contour principle) leads to unnecessary complications, and that it should be abandoned. This leads to certain particulars that are different than their counterparts in Leben's analysis; it should be kept in mind that although I am using autosegmental notation here, the principal empirical differences between Leben's analysis and mine result from the inclusion or rejection of the "Obligatory Contour principle," which may be stated as:

> Obligatory Contour principle (Leben): At the melodic level of the grammar, any two adjacent tonemes must be distinct. Thus HHL is not a possible melodic pattern; it automatically simplifies to HL.

As I say I shall argue against the inclusion of such a principle within phonological theory at the phonological or tonological level.[8]

In describing the Tiv verbal system, Arnott begins by explaining that the tonal pattern for each tense need only be illustrated by a small number of examples: six, to be exact. There are twelve "tenses":

> General Past
> Future
> Imperative
> Habitual 1, 2, 3 & 4
> Recent Past A & B
> Subjunctive
> Continuous
> Past Habitual

The verb stem can in general be one, two, or three syllables long; and -- it is important to note -- for each length there are at most two (generally two) possible tone patterns. Thus, for example, the General Past is illustrated as in (42).

(42) GENERAL PAST ve 'they'

! = Drop; see text

High	Low
vé!vá	vé dzà
vé!úngwà	vé vèndè
vé!yévèsè	vé ngòhòr

(As of yet, the reason for calling the first column "High" and the second "Low" is unexplained.) Normally two successive High tones are at the same pitch; sometimes, however, there is a "Drop" of one or two semitones between the successive Highs. This drop is indicated by the raised exclamation point.

The point has been emphasized (in the three analyses cited above and elsewhere) that the six tonal melodies in (42) are not unrelated; that the language learner need not memorize six separate patterns for each tense. While this is true and important, we should emphasize equally the following point: whether the speaker knows six formulas for the General Past or just one, these are melodies <u>of</u> the General Past, not of the word. For each tense, there is a corresponding chart as in (42), and from tense to tense the particular tone pattern on a particular verb varies wildly. All verbs of the same syllabic structure have the same tonal pattern in a particular tense, however. The conclusion is that there is a level in the grammar which contains as entries in it various melodic patterns independent of the words or vowels; these entries can be assigned to particular tenses.

This then is the argument for autosegmental representation here: the tonal pattern found on each verb is linked in the lexicon not with the verb but with the verb tense; thus the melody must be a separate entity in the lexicon. The only system capable of merging into one simultaneous utterance two lexical (phonological) entries is autosegmental phonology.[9]

We might then conclude that for each tense, the speaker of Tiv has learned six tonal melodies, two for verbs of one syllable, two for those of two, and two for those of three syllables, as in (43). Which of the two possibilities a particular verb selects must be arbitrarily marked in the lexical entry of the verb.

(43) 2. General Past

1 syll:	$^!$H	L
2 syll:	$^!$HL	LL
3 syll:	$^!$HLL	LLL

Such a tentative conclusion, though, is implausible, for the six forms in (43) have too much in common for it to be an accident. Arnott points out that in general a "drop $^!$" is imperceptible before L; one could as well say that the drop $^!$ occurs before all forms, and lessen the amount of arbitrariness in the tone formulas.

Similarly, Arnott points out, it is only the first syllable whose tone must be arbitrarily memorized, and the distinction here is between H and L. Modifying Arnott's notation slightly,

we could summarize both 2-syllable tone patterns as $\underline{!BL}$. B here
is a variable that ranges over H and L. Thus the abbreviation
stands for !HL, or !LL, which, as we noted, is indistinguishable
from simply LL.

If we adopt this formula for the 2-syllable forms, the
Well-formedness Condition accounts for the tones of the 3-
syllable verbs starting with the same formula. That is, if the
first two syllables of yévèsè were treated as a bi-syllabic verb,
the third syllable would automatically become L (see (44)).

(44)   y e v e s e
            | |
         ! H   L

Through the use of the B variable, we can generalize across
the L and H tonal forms (as Arnott points out): through the use
of the Well-formedness Condition, we can generalize over the 2-
and 3-syllable cases. One would hope that the 1-syllable forms
would be expressible by means of the same formula, for then we
could conclude that the Tiv speaker need memorize only one tonal
formula for the General Past.

The formula !BL works, we observe, for the L-verb, since
LL = L in effect; however, for the H-verbs this formula predicts
!HL (i.e., Falling) rather than !H, the correct form. In a
sense, the difficulty has been created by the Well-formedness
Condition, which causes the L to associate with the single
vowel.

(45)   va
           /\
        ! H L

The theory now requires us to set up a language-particular rule
to delete the L in this context.
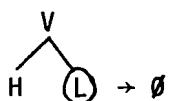
There are two reasons to believe that this is, in fact, a
language-particular rule and not a failing in the theory. First,
recall that in the preceding example from Mende, the same melody
HL does in fact map onto a monosyllable as HL; thus the simpli-
fication to H is certainly not universal. In fact, we have seen
the same thing in English ((27) above) where HL does not auto-
matically simplify.

Secondly, we need a process to delete the L toneme in (45)
anyway, assuming that !BL is the correct formula. As Arnott
suggests, and McCawley argues, the Drop ! is really a Low toneme.
Independently of whether such a claim is true in all cases, it is
superficially true that a pitch drop does occur on the second of
two Hs when they are separated by an L. Thus an unattached L
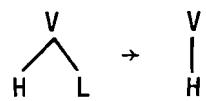
(diagram (45) if the dotted line were not present) should cause
a drop in the pitch of the next tone if that tone is an H. In
fact, no such drop occurs; therefore the L has been deleted. In
sum, the fact that the Well-formedness Condition in this case
forces us to draw the association line in (45) is not unfortunate
in any sense, for we need a rule to delete the L here independently
of any conventions like the Well-formedness Condition. In fact,
it is the inclusion of the Well-formedness Condition that makes
the Deletion rule simple to state, as we shall see.

Our rule, then, will be (46), Fall-Simplification.

(46)   V   V   V
        /\   which abbreviates   /\   → |
       H (L) → ∅   H   L   H

With the inclusion of this rule, all the forms in (42) or (43)
fall under the single formula ˧BL.

Our formula needs one more modification before it can be
adopted finally. With a theory of downstep as described below in
Chapter 2 (Section 1), we expect a downstep or Drop toneme that
has the properties of the sequence "!H" (we have already seen this
briefly in Figure (19) above). If we call this toneme "Drop" or
D, then the formula written ˧BL actually stands for DL (when B =
High) and LL (when B = Low); again using a suggestion of Arnott's,
we summarize these two formulas as B*L, which now replaces the
formula ˧BL.

    To repeat: B is a variable that ranges over H and L;
                 B* is a variable that ranges over D and L;

in particular, B* = D if B = H.

Let us now look at the rest of the Tiv verbal tone patterns,
a bit more rapidly. It is not our intention to prove that in all
tenses, the formulas reduce to one, as they did in the General
Past. To succeed in doing so would, it appears now, require a
fairly powerful set of rules, and to claim that the learner of
Tiv prefers to learn, say, two basic formulas for a tense and
stay with formally simple tonological rules permits us to posit
an evaluation metric with which we can come very close to our
ultimate goal: that is, explaining why this particular solution
was arrived at on the basis of the data.

For purposes of exposition, I shall give for each of the
remaining tenses only the tonal patterns (as in (43)) rather than
the actual spellings of the verbs with tones marked diacritically.

(47) 2. Continuous

	High	Low
1 Syll:	HL	HL
2 Syll:	HLL	HLL
3 Syll:	HLL	HLL

In the 1 and 2 syllable forms, a suffix is added to the stem, so two tones are indicated for the "1 syll" forms, and three for the 2. But the conclusion is rather evident here: there is only one tone formula, HL.

(48) 3. Future

	High	Low
1 Syll:	D	L
2 Syll:	HL	LL
3 Syll:	HLL	LLL

On the basis of the 2 and 3 syllable forms, the tone formula should be BL; given this formula, in fact, the 2 and 3 syllable forms work exactly as expected, and so does the Low 1 syllable form. As in the general past, the problematic form is the 1-syllable H form; here, the sequence HL reduces to D.

But given the revision that the General Past formula for H-toned verbs is DL, the revised form of (46) is D͡L → D. Using the features suggested on p. 54 below (and identifying the Drop toneme D with Mid tone), we write the revised Fall Simplification rule (49).

(49) Revised Fall Simplification Rule

$$\begin{Bmatrix} \widehat{D} \\ H \end{Bmatrix} L \rightarrow D, \quad \text{or} \quad [+\text{High}][+\text{Low}] \rightarrow \begin{bmatrix} +\text{High} \\ +\text{Low} \end{bmatrix}$$

Thus, having simplified the already-needed rule of Fall Simplification (simplified in that one feature specification has been eliminated), we are left with BL as the Future tonal melody.

(50) 4. Habitual 1

	High	Low
1 Syll:	D	D
2 Syll:	DH	LH
3 Syll:	HHL	LHL

The 3-syllable forms reduce to one formula, utilizing the B-notation, BHL.[10] Maintaining left to right mapping as in Mende, the Low-toned 2-syllable verb works as expected, then, for LHL is mapped as $\underset{\displaystyle L \;\; H\overset{\frown}{} L}{V \quad V}$. The HL contour is simplified to D, and, as we know, the sequence L D and L H are indistinguishable.

Nonetheless, this formula (BHL) does not work for the remaining cases, not is there a simply-formulatable rule that can derive the correct surface form from underlying BHL. Rather, we posit the melody B*H for the 1 and 2 syllable forms. This in turn requires positing a complementary contour simplification rule (51).

(51)   LH Simplification

$\overset{\frown}{L\ H} \rightarrow D$

(51)   **5. Habitual 2**

	High	Low
1 Syll:	D!	D!
2 Syll:	DH!	LH!
3 Syll:	DHH	?

Maintaining the same principles we have seen so far, the tone formulas for this form are straightforward. Arnott notes that it is not clear to him whether there is, in fact, a 3-syllable Low form in the Habitual. The other tri-syllabic form inexplicably does not have a following !; aside from that curious feature, the five forms cited above reduce to the formula B*H!. The reader will recall that for the Low-toned monosyllabic form, this formula gives LH! and that this reduces to D! by (51) LH Simplification. If in fact Arnott is correct about the absence of downstep following the tri-syllabic form, then either a minor rule must be posited to delete the Drop there, or two formulas -- B*H! and B*H -- must be posited. At present, these two cannot be compared on empirical grounds, and our evaluation metric is not subtle enough to decide for us.

(52)   **6. Habitual 3**
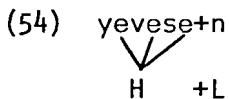
	High	Low
1 Syll:	DH	DH
2 Syll:	DHH	LHH
3 Syll:	DHH	LHH

Again, as in the continuous form above, the 1 and 2 syllable forms have a suffix, and so an extra tone-bearing element. With no further ado, this form simplifies to the melody B*H#H.

(53) <u>7. Habitual 4</u>

	High	Low
1 Syll:	HL	HL
2 Syll:	HHL	HHL
3 Syll:	HHHL	HH(H)L

In this form, the suffix /n/ appears to bring and bear its own tone; thus the + in the tonal formula. We find, then, forms like váán, dzáán, úngwáń, véndáń, yévéséň, ngóbóř. I shall tentatively presume that the suffix, preceded by a phonological boundary (represented here by +) has a corresponding toneme (L) and the boundary on the phonological level is matched by one on the tonological level. Thus we set a pattern like:

(54)   yevese+n

H   +L

But, with this assumption, the tonal formula for this form is just H#L.

(55) <u>8. Past Habitual</u>

	High	Low
1 Syll:	D +L	D +L
2 Syll:	DH+L	LH+L
3 Syll:	DHH+L	LHH+L

In this form, we have the same appearance of a Low-toned suffix as in the form preceding. Again, with the set of simplification rules posited above, these six forms reduce to the formula B*H#L.

(56) <u>9. Recent Past A</u>

	High	Low
1 Syll:	D	D
2 Syll:	DH	LH
3 Syll:	DHL	LHL

The 3-syllable forms, considered alone for the moment, reduce immediately to the familiar formula B*HL. This formula, however, does not apply to the 1 and 2 syllable forms, which derive from the formula B*H. We shall not posit any ad hoc rule to derive one from the other; such a move does not, in fact, seem possible in this case. Some sort of parenthesis notation may well be available for these lexical melodies: the two would then be

expressed B*H(L)$_3$, where the subscript "3" indicates that the L̲ is chosen when the word is tri-syllabic.

The next two tenses have the same tonal melodies, a puzzling set of tones:

(57) Recent Past B / Subjunctive

	High	Low
1 Syll:	H	H
2 Syll:	HH	HD
3 Syll:	HHL	HHL

What strikes us as immediately odd is that there is no tonal distinction between the two columns for the 3-syllable forms, but there is for the 2-syllable forms. On the basis of the 3-syllable forms, we would rather naturally expect a melody HHL; this generalized correctly to the Low-2-syllable form, but not to the High-2-syllable form or the 1-syllable forms. The 2-syllable forms, considered alone, would derive from only one of two possible melodies, given our assumptions: HHB or HBH. Neither of these, however, generalize naturally to the 1- or 2- syllable forms. Adopting the parenthesis notation that seemed plausible in the previous form (Recent Past A), we see that the melody HHB that is appropriate for the 2-syllable forms can extend to the 1-syllable forms thusly:

(58) H H (B)$_2$

When the subscript option is not taken -- that is, in the case of a 1-syllable stem -- the formula is HH, which is the same as H on one vowel, as required.

Pursuing this, we would attempt to see if the formula (58) has any natural relation to the formula for the 3-syllable forms, HHL. The answer seems to be clearly yes; the relationship may be stated as in (59).

(59) Recent Past B / Subjunctive

$$H\ H \begin{Bmatrix} B_{2\Sigma} \\ L_{3\Sigma} \end{Bmatrix}$$

This shall be the formula we posit, then, for the Recent Past B and the Subjunctive,

(60) 12. Imperative

	High	Low
1 Syll:	H	H
2 Syll:	HL	LH
3 Syll:	HHL	LHL

This is the last of the 12 tenses to be considered, and the one for which we can offer the least satisfactory explanation. Given our rules so far, the formulas underlying 1 and 3 syllalbe forms are simply H (or conceivable HH, but there seems to be nothing supporting this); the 3 syllable forms are derived from the formula BHL.

The problem is the 2 syllable forms. As Arnott points out, the generalization here, if there is one, is that the tone of the form is the "Base tone" (b) followed by its opposite -- thus L follows H, and H follows B. If we accept this proposal, then the 2-syllable formula is, in Arnott's terms, B-opp(B).

This is less than satisfactory for two reasons. First, it is the first and only time such a device -- "Opposite" -- need be introduced in this statement of the Tiv tonal system. From this we could not conclude, of course, that a notion like "tonal opposite" does not occur in tone languages in general; in fact, we shall see a clear example of such a thing in Igbo below (Chapter 2). Thus, while it is not impossible as a solution, we would hope to see more reason why the "opposite"-function could play a role in Tiv.

Second (and leading out of the first point), all the other forms showed some coherence among the 1, 2, and 3 syllable forms; the adoption of the "opposite" notation breaks this generalization.

The question, then, is to determine when a language learner, presented with data as in (60), will memorize independent melodies for each syllabic structure (61), and when she or he will adopt a simpler melody supplemented by tonal adjustment rules. To the extent that the tonal adjustment rules necessary for the second approach seem very complex, the independent memorization position is supported; to the extent that they are fairly simple, the second, general melody approach is supported. However, the notion of "simple" is internal to a theory of tonological rules, not external to it, and so we shall not be able at this early state to reach a conclusion about the actual form of the Imperative in Tiv.

(61)   Tiv Imperative a la memorization:

1 Syll:	H	(essentially Arnott's
2 Syll:	B + opp (B)	solution)
3 Syll:	BHL	

If we attempt to fit all these patterns into one melody, we observe that the 3-syllable formula, BHL, predicts the correct form for the 2-syllable Low-toned verb, for the melody LHL spreads over a bisyllabic word as L ĤL, which becomes L D, which is (we have observed) indistinguishable from the reported LH.
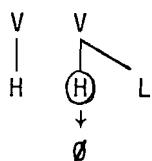
However, such an approach does not work straightforwardly for the H-toned bisyllabic forms, for HHL would give HD, rather than the required HL. Suppose we could resolve this problem; the formula for the 1-syllable forms could (rather trivially) be derived from the 3-syllable formula with parenthesis notation, leaving us with the Imperative Formulas as in (62):

(62)　Tiv Imperative with Readjustment Rule

$$(B)_{2,3} \; H(L)_{2,3}$$

The readjustment rule that is necessary, as we have observed, is for the H-toned bisyllabic form, as in (63).

(63)　Minor rule for H-bisyllabic imperative:

To repeat, the adoption of solution (61) or (62) or some intermediate solution must await further theoretical work.

To sum up this section, we have considered in some detail the verb tone system of Tiv, an African tone language, indicating how the tonal melodies must be described at a particular tonological level in the grammar.

5.　Floating Tones

Next we shall consider the nature of "floating tones," a device that has proven useful in working with tone languages but whose theoretical status has always been suspect. A floating tone is, in essence, a segment specified only for tone which, at some point during the derivation, merges with some vowel, thus passing on its tonal specifications to that vowel. This is, in any event, the traditional view; and this traditional view, framed within the standard theory, fixes the floating tone as one of the segments, and therefore linearly ordered amongst all the other, more completely specified, segments of the phonological representation.

Thus it has been suggested that certain affixes are purely tonal -- Bird (1966), for example, posits a floating L tone for the Bambara definite marker on nouns.

We shall say -- and we shall see -- that floating tones are melodic levels, much as in the previous section, that map onto the syllabic structure in a slightly more complex way than the Tiv forms. Often floating tones associate after some phonological/tonological rules have applied, though this is not always true.

Goldsmith 46

Igbo presents several interesting cases of floating tones, which we shall look at in detail in Chapter 2. For the present, let us consider the general outlines of their behavior.

We will look at the pre-verbal floating H tone, a mark of subordinate clauses. In Chapter 2 we will consider in detail when the H tone occurs; for now, we will content ourselves with observing its behavior when it is there. If a tone "floats" when it has no vowel associated with it, let us say that the process of associating a floating tone is "docking."

We shall see four separate cases of "docking" of this pre-verbal H tone -- four kinds of docking that are predictable on autosegmental grounds, and on no other.

To describe these effects, we must review some of the basics of the structure of Igbo. It is a Subject-Verb-Object (SVO) language; singular subject pronouns (though not plural) have two forms, non-cliticizing or Strong, and cliticizing or Weak. The cliticization of the subject pronoun, however, includes a syntactic movement. In certain tenses, that is, the verb stem is normally preceded by the prefix /a+/ when the subject is a full noun phrase (NP). Thus, if the subject is ányì ('we,' not a clitic) and the stem is /za/, we find (64).
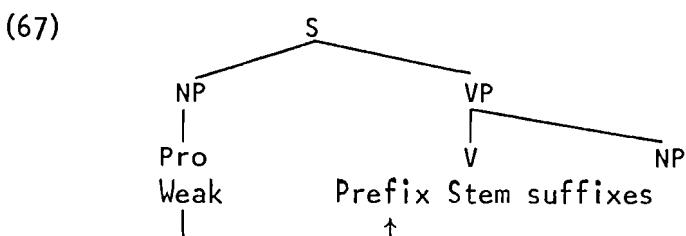
(64)   ányì ázàá àlà   /za/ = sweep

        we   sweep floor   (stem followed by suffix /+a/)

        'we swept the floor'

If the subject cliticizes, the prefix /a+/ disappears, because, we shall see, the cliticization puts the clitic subject into the /a/-prefix's position. Thus we get (65).

(65)   ọ́ zàá àlà   'he swept the floor'   /o/ = he, she

We may formalize this operation as in (66), illustrated in (67).

(66)   [Pronoun] [Verb]   Subject Cliticization
        [Weak] [Prefix]

              1   2   →
              Ø   1

(67)   S
                   ┌────┴────┐
                  NP   VP
                   │   ┌──┴───┐
                  Pro   V   NP
                  Weak   Prefix Stem suffixes
                   └────────────↑
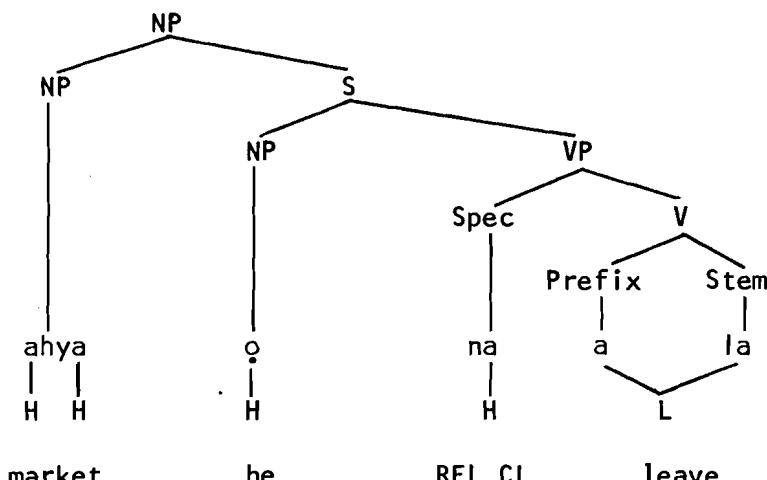
Such a rule must be a local rule, in Emonds' sense, given the general structure-preserving framework (see Emonds (1970) and (forthcoming)). Therefore, if any element could appear in between the Weak subject and the prefix position, it would block this cliticization process, and thus keep the /a/ prefix from disappearing.

The /na/ relative-clause marker does precisely that. In a structure such as in (68), the clitic subject stays where it starts, and it cooccurs with the /a/ prefix.

(68)

```
                              NP
                ┌──────────────┴──────────┐
               NP                          S
                │              ┌───────────┴──────────┐
                │             NP                      VP
                │              │              ┌────────┴──────┐
                │              │             Spec             V
                │              │              │          ┌─────┴─────┐
                │              │              │        Prefix      Stem
                │              │              │          │           │
              ahya            o             na          a          la
               │ │            │              │          └─────┬─────┘
               H H            H              H                L

            market           he           REL CL          leave
```

The picture in (68) is important; around it will revolve our argument. The tone attached there with the Relative Clause marker /na/ is the floating tone H. We shall see that under other circumstances, that same H tone docks on other vowels, depending, we might say, on what is "closest" to it in derived structure.

We have seen first, then, that the /na/, when present, is H-toned. The /na/ may not be present for two reasons: first, in a relative clause, the /na/ is optional; thus we may simply delete it and see what the derived structure is. As we observed in Section 3 above, when discussing tone stability, we expect the H tone to appear elsewhere. Second, and quite differently, while the /na/ always marks a relative clause when it appears, the H-tone appears in other subordinate clauses indicating adverbial dependency (that is, in sentences translatable as "Lest X happen,..."). In this second case, the H tone is not a remnant of a deleted /na/, for the /na/ could not have been there in the first place.

In both such cases, when the subject of the clause is a normal noun (non-cliticizing, that is), the final tone of the subject is raised. For example, the noun ọ̀nụ̀ (a type of yam) shows up as ọ̀nụ́ in the relative clause ọ̀nụ́ rèré éré (rèré éré is a complex predicate meaning 'is rotten'). Here, as elsewhere in Igbo, the contour tone LH generally simplifies to what is called

a Mid tone in Igbo, but is precisely the same as the Drop (D) in Tiv. Its tone is slightly lower that the first preceding High tone. But since any High tone is slightly lower than a preceding High tone if the two are separated by a Low tone, High and Mid (or Drop) are indistinguishable after a Low tone. Thus in this position, LL nouns become L+Rising or L Mid; HL nouns become H+Rising, or more commonly H+Mid, as in (69).

(69) ázù rèré èré   "the fish that was rotten"
             ázù 'fish'

Nouns that end in a High tone do not change; see (70).

(70) ánú rèré èré   ánú 'meat'

       àkwhá rèré èré   àkwhá 'eggs'

In all these cases, the /na/ could have been present, captured the High tone, and then these nouns would have displayed their isolation tones as cited. In the following sentences, the na may not appear, since these are not relative clauses, and the same subject tone raising occurs as noted above.
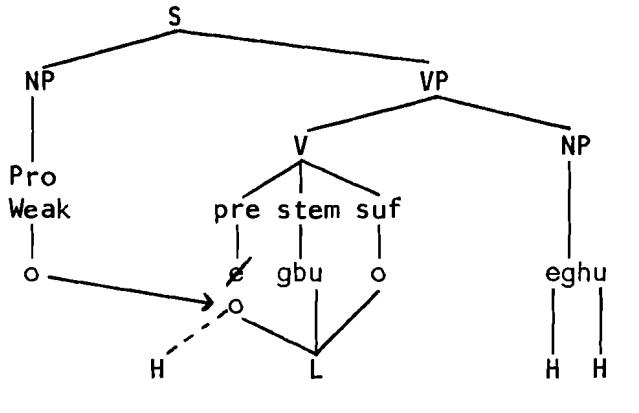
(71) Khwàchíé úzò

      Shut the door lest....

éghú ègbùò àghú	/éghú/	'leopard' (a) HH
leopard/ kill/goat		
òké àtàà àkhú	/òké/	'rat' (b) LH
rat/eat /palm kernels		
úzè àtàà yá	/úzè/	'squirrel' (c) HL
squirrel/eat /them		
èŋwó àtàà yá	/èŋwó/	'monkey' (d) LL
monkey/eat /them		

If in one of these "lest..." clauses, as in (71), the subject were a weak cliticizing pronoun, it would undergo the cliticization expressed in (66) and (67) above. Looking at the tree in (68), we see it would then be to the right of the H tone, in a certain sense; in as much as the cliticized pronoun partakes of the verbs stem's L-tone just as the /a/ prefix in (68), the H-tone docking rightward now will give rise to a <u>Falling</u> tone! All this is clear geometrically: see (72).

(72)
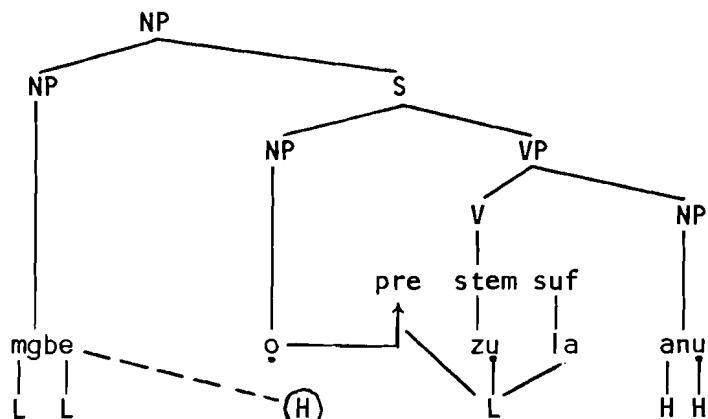
'lest he kill the leopard'

[ô gbǜò éghú]

Lastly, when the relative clause or subordinate clause is constructed appropriately, the H-tone can dock right out of its clause. Thus, if the subject is a clitic, and moves by rule (66) over the floating H (much as in (72)), the floating H may then dock leftwards onto what is to its left. If we have a relative clause where the head deletes into object position in the relative clause, we get a situation as in (73), and one can say either (74a) maintaining the /na/, or (74b), where the floating tone tone docks leftward onto the head m̀gbè.

(73)

m̀gbé ọ̀zụ̀là ánú,...

'whenever he buys meat,...'

(74) a. m̀gbè ọ́ náàzùlà ánụ́ ...

      b. m̀gbé ọ̀zùlà ánụ́ ...

In summary, then, the floating H tone appearing before the verb in certain tenses in Igbo docks onto various vowels depending on the derived syntactic structure: the end of the subject of the clause it is in, or the beginning of the subject clitic pronoun of that clause, or the grammatical particle /na/, or the head of the relative clause. The further behavior of this floating tone, and its interaction with other processes in Igbo, will be covered in detail in Chapter Two.

6. Automatic Spreading

The fifth and last argument for autosegmental phonology here comes from the phenomenon of bidirectional spreading and, we would suggest, its ungoverned nature in these cases -- that is, the spreading is not due to a specific phonological rule, but rather to the geometry of autosegmental representations, and its Well-formedness Condition (or WFC as expressed above in Section 2.3).

It should be clear by now from the examples in each section how the Well-formedness Condition creates the spreading of tonemes over various syllables automatically; we have seen examples from English, Mende, and Igbo. In this section, I will look at a particularly interesting example, one in which the autosegmentalized level is not tone, but rather nasalization. The language is Guarani; the data and much of the initial arrangement of it comes from Lunt (1973) and especially Rivas (1974,75).

Let us begin by considering the forms that morphemes may take. We observe in (75) that there are basically two forms for the affixes /no/, /ro/, and /I/: each have a basically nasal form and a basically oral form.

(75) a. ^{n}do + ro + ha+hú + I   (oral stem)

   b. nõ + r̃õ + h̃ẽ^{n}dú + I   (oral stem with nasal in it)

   c. nõ + r̃õ + nũp̃ã́ + T   (nasal stem)

The prefix nõ thus alternates with ^{n}do, as does r̃õ with ro, and T with I. Which alternate appears is determined by the stem, the morpheme which contains the accent in the examples above. In general, a vowel with an accent in Guarani is also one endowed with a nasality melody or specificaiton. That is, if we say nasality is autosegmentalized in Guarani, we specify that the nasality melodies (essentially just Oral or Nasal) are initially associated with accented vowels, just as we saw in Section 2.2 that the tone melody in English was initially associated with the

accented vowel. In addition, true nasal consonants are endowed
with an inherent Nasal specification on the autosegmental level
(these true nasal consonants are m, n, ŋ, ŋʷ). We shall re-
present a stem as in (75a) by (76a); a stem as in (75b) by (76b).

(76) a. h a i h ǔ̆   b. h e D ǔ̆
                    ⁎ | |
                    0 | ⁎
                                         N 0

The symbol "D" represents all the normal feature specifications
of an /n/ except for nasality -- raised tongue, and so forth.
When associated with a [+nasal] autosegment, then, the two of
them are an /n/. "0" symbolizes [-nasal]; "N" symbolizes [+nasal]
(read "Oral" and "nasal," respectively). As above, the star ⁎
indicates an accent; corresponding autosegments are accented
(starred) also.

The Well-formedness Condition must apply to the forms in
(76). The form in (76a) becomes all oral, as in (75a); so much
is correct. The form in (76b) becomes the form in (77), which
would be transcribed as hẽnu non-autosegmentally.

(77) h e D ǔ̆
          \|| |
           N 0
             ⁎

The empirical difference between the form created by the Well-
formedness Condition and the actual form is that the Condition does
not create forms like /ⁿd/, the prenasalized stops. This is not,
however, a bad result; the prenasalized stop is a complex configura-
tion, and the universal theory should not produce them automatically.
Rather, the language particular rule (78) creates the form as in
(75b) from the form in (77). When we apply (78) Postoralization
to (77), we get the correct form.

(78) C   V   Guarani Postoralization Rule
      ⌐⟍⌐⌐⌐⌐
      N   0
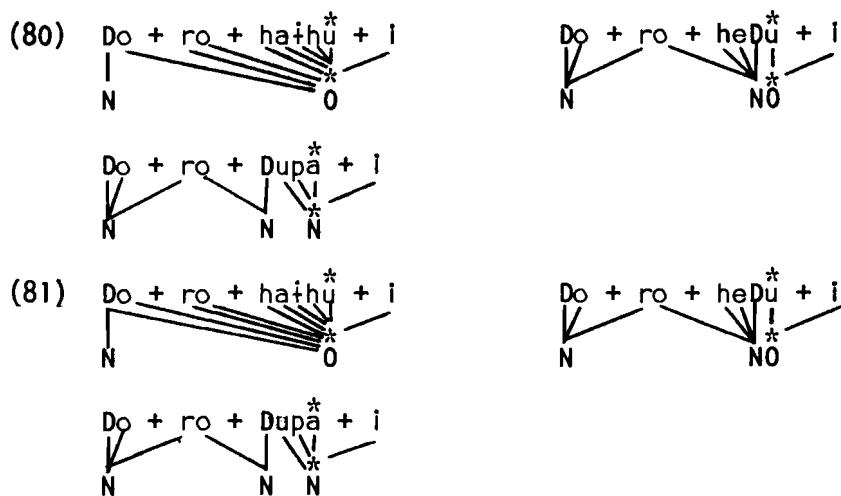
(The reader will recall that the dotted line in an autosegmental
rule indicates the addition of an autosegmental line by that rule.)

With this autosegmental interpretation of the data, let us
derive the forms in (75) precisely. We note that when the Well-
formedness Condition applies here, if there is an ambiguity as to
whether a starred or an unstarred autosegment spreads, it is the
starred element that does.

(79) Do + ro + haihǔ̆ + i   Do + ro + heDǔ̆ + i
      | | | | |
      N   ⁎   N | ⁎
                   0   N 0

      Do + ro + Dupǎ̆ + i
      | | |
      N   N ⁎
                   N

(79) represents the underlying forms. The Well-formedness Condition changes these to (80); rule (78) Postoralization creates (81), the correct output (compare (75)).

(80)

(81)

As expected, we have seen that prefixes and suffixes adopt their nasality from the stem they are attached to. These prefixes are found with the usual + morpheme. Note that this bidirectionality of spreading is expressed only with additional specification if handled by a phonological rule; within autosegmental theory, the bidirectionality of the spreading is the result, in a sense, of the fact that there are prenasalized stops! That is, the existence of the prenasalized stops required the autosegmental analysis, which in turn brought with it bidirectional spreading. In summary, borrowing from Rivas (1974) #8.3: suffixal nasality spreading is as in (82).[11]

(82) $\widetilde{\text{Stem}}$ + suf → $\widetilde{\text{stem}}$ + $\widetilde{\text{suf}}$ (nasal)

$\overline{\text{Stem}}$ + suf → $\overline{\text{stem}}$ + $\overline{\text{suf}}$ (oral)

$\overset{\sim}{\overline{\text{Stem}}}$ + suf → $\overset{\sim}{\overline{\text{stem}}}$ + $\overline{\text{suf}}$ (nasal-oral)
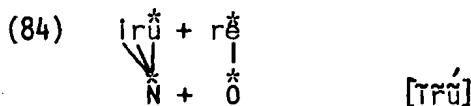
There are also suffixes separated by word-boundary (#) endowed with their own nasality-specification, according to Rivas (1974, 75), as in (83).

(83)   ## Stem # Suffix ##
       || | ||
       ## αNas # βNas ##
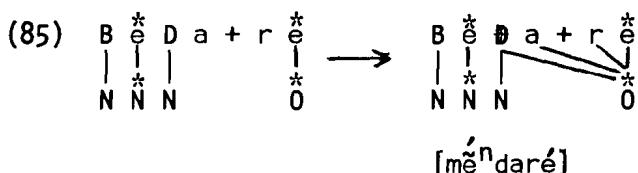
We see in (83) what has been only implicit up to this point: the universal convention that corresponding word-boundaries, but not +boundaries, are connected by association lines.

Another logical possibility is predicted by the notation so far: why couldn't there be an accented suffix -- that is, one with its own nasality melody -- but one which has only a +boundary, not a #boundary. With a +boundary, there will be no association line, and the nasality autosegments of the stem and the suffix will interact. We predict, then, the following type of suffix -- one which meets the description given by Rivas for /+ré/ suffix.

/+re/, endowed with an oral melody on its stressed vowel, gives [Trūré], as in (84). Autosegmental theory does not predict whether the second r is nasal or oral; when surrounded by vowels with conflicting specification as in (84), such a determination is not possible.

(84)

[Trū]

When /+re/ is placed on a nasal stem with a nasal consonant to the right of the accent (such as /mēnã/) we get the correct result as illustrated in (85), a remarkable form.

(85)

[mẽ^{n}daré]

7. Conclusion

We have traveled a long way since our early remarks about the abstract geometrical shape of phonological representations, but in another sense we have surely never left that topic. In these five general cases reviewed, phenomena that are puzzling on the standard view of phonological representations become quite tractable and reasonable when viewed from an autosegmental perspective: and we see, furthermore, connections between these that we never saw before. If we may draw some morals from this, perhaps they are that advances in phonological theory may start from an interested attitude towards low-level articulatory facts; that if they are interesting, they go beyond these superficial facts to unexpected phenomena; and most importantly, that we do not begin our research with an understanding of the nature of the most elementary linguistic observables. The most astounding revelations may be those that change our conception of what we thought were the observables, either in phonology or semantics. We would not restrict our attention to constraints on rules -- phonological, syntactic, or semantic -- at the risk of missing the very nature of the items involved.

FOOTNOTES FOR CHAPTER ONE

[1]The requirement of "short" here is meant to avoid the possibility that a long vowel might actually be two successive short vowels, each with a level tone. This, of course, does happen frequently; it is, however, irrelevant to the present context. The situation is much the same for pre-nasalized stops: in some languages, these stops are clearly <u>one</u> segment, even though they have two successive feature-specifications for nasality.

[2]Leben (1973) wrote, "Suprasegmental tones are by definition independent of any segments -- rather than being expressed as features on segments, they are features on larger linguistic units" (p. 26). Compare this with Troubetzkoy (*Principes de Phonologie*, p. 95-6)" "Les particularit és phoniques qui forment des oppositions distinctives dans les diverses langues peuvent être réparties entrois classes: particularités <u>vocaliques</u>, <u>consonantiques</u>, et <u>prosodiques</u>. Les phonèmes vocaliques consistent en particularités vocaliques distinctives et les phonèmes consonantiques en particularités consonantiques distinctives. Par contre il n'y a aucun phonème qui consiste exclusivement en particularités prosodiques. Ces particularités sont plutôt liées, selon les langues, à un phonème vocalique déterminé, ou à un phonème consonantique déterminé, ou enfin, à toute une suite de phonèmes."

[3]One could, in this regard, accept the suggestion of Woo (1969) that <u>all</u> contour tones are necessarily such concatenations. The logic of the situation would not be much changed; we would simply not need to investigate the particular language in question to come to the conclusion mentioned in the text.

[4]We shall -- more for the sake of being definite now than for anything else -- assume a theory of tonal features that is defended at more length in the following chapters. We adopt two binary pitch features: +Highpitch and +Lowpitch.

$$\begin{bmatrix} +\text{Highpitch} \\ -\text{Lowpitch} \end{bmatrix}$$ will represent <u>High tone</u> (H)

$$\begin{bmatrix} -\text{Highpitch} \\ +\text{Lowpitch} \end{bmatrix}$$ will represent <u>Low tone</u> (L)

$$\begin{bmatrix} +\text{Highpitch} \\ +\text{Lowpitch} \end{bmatrix}$$ will represent <u>Mid tone</u> (M)
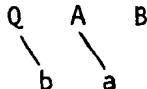
the fourth possible combination, $\begin{bmatrix} -\text{Highpitch} \\ -\text{Lowpitch} \end{bmatrix}$, will tentatively be excluded as a universal principle. The distinction between tone and pitch will be drawn in Chapter Two below.

[5]Strictly speaking this is not true: the one case we cannot rule out is: A   B. There are, however, good reasons

$$A \times B$$
$$a \quad b$$

not to want to rule this out, for if we could rule it out we would have given an inherent sense to each level, not just total ordering. In fact, in every possible linguistic case, there will be at least three segments on some line, and there WFC above works correctly.

[6]Again a proviso must be made. (32) is not in violation of WFC; in general, we find, when we inspect the definitions provided here, that if segments on the end of some L' are not associated, WFC is not violated. In fact, this odd formal characteristic appears to be exactly reflected in the behavior of certain floating tones. That is, to the extent that WFC differs from (24) above -- WFC making curious predictions -- it is WFC which is more correct.

(32)   Q   A   B
            \   \
            b   a

[7]The phenomenon of stability is the subject of my "Tone Melodies and the Autosegment," from which this section draws heavily. Further examples not mentioned here are treated there (it appears in the Proceedings of the 1975 African Languages Conference, in the Ohio State Working Papers in Linguistics No. 20).

[8]In particular, I claim that it is not true at levels in the grammar where phonological and tonological rules apply. Leben's system calls for there to be sequences of identically-toned vowels (CV́CV́CV́) after his "tone mapping"; thus, for him, the Obligatory Contour principle holds in the abstract levels but not at the superficial levels. I shall suggest that rather the opposite holds, in Chapter 4.

[9]One could imagine logically the position that each verb has its tone marked on it in the lexicon, arbitrarily selecting one of the tenses' tonal patterns as the underlying one; the other tonal patterns might then be derived by tonological rules sensitive to the semantic mood and tense of the clause. While not impossible within current phonological theory, such an approach does not in fact work.

[10]Here is the first case where the Obligatory Contour principle is violated, for HHL does not satisfy the principle.

[11]See Section 6 of Chapter 3 below for a further discussion.

CHAPTER TWO

IGBO TONOLOGY

1. Introduction

Our course so far has been to consider a series of arguments, each of which focused on a recurrent phenomenon in phonology that points toward autosegmental representation. Adoption of autosegmental tiers, we argued, resolved these fundamental problems in phonological theory.

In extending autosegments to linguistic theory, we have, it seems, made it an offer it could hardly refuse, once aware of the problems discussed in Chapter One. Still, even if we adopt autosegmental representation, we will want to know what it is that we have just taken on. What will a grammar look like now? Our most fervent hope would be that the inclusion of autosegmental levels will simplify our conception of phonological rules, rather than render it more complicated; but this cannot be decided in advance.

There is only one way to settle the question, and that is to investigate in depth a system that has a non-trivial autosegmental component. This chapter is devoted to such an investigation, wherein we shall look at the tonological system of Igbo, a Kwa language spoken in southeastern Nigeria.
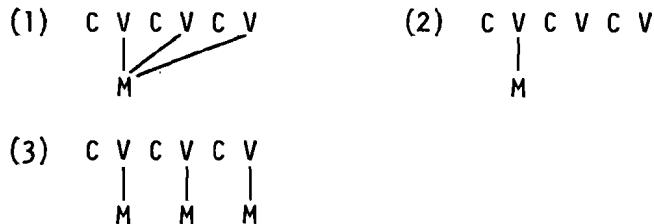
This chapter is intended, at the same time, to begin the development of a more articulated autosegmental theory. Autosegmental theory as presented in Chapter One is a family of theories, rather than just one. Because it represents such a fundamental change in phonology, the ramifications of autosegmentalism lead to a whole series of new questions when we try to implement it in a specific linguistic theory. Just what features may be autosegmentalized, for example? In a given language, may a certain feature be autosegmentalized at one point in a derivation, but not at another point: Once we begin to consider the questions that arise, each one virtually leads to ten more.

This is, I believe, a highly desirable situation for phonology to find itself in. For it is only against the background of a choice among hypotheses that the truth may emerge; and only when alternative hypotheses may be expressed, entertained, and argued against can we make a scientific advance. The point, then, is that although the decision to revise phonology to include autosegmental levels opens, in turn, a series of further questions, the autosegmental proposal is in no sense vacuous; it is, rather, theoretically fruitful. It leads us to see, pose, and answer questions that we were not aware of before.

The remainder of this thesis is devoted to my own view of the answers to some of the questions that autosegmental theory raises. It would be fair to say that of the number of linguists

who have begun to work within the autosegmental framework, the
major area of agreement rests within the proposals made in
Chapter One. From here on, we consider one autosegmental theory
among several..

This chapter develops a number of specific claims. For
example, I shall argue that a representation as in (1) is
empirically distinct from that in (3), and that when a repre-
sentation like (1) occurs (produced, perhaps, from (2) by the
Well-formedness Condition), it cannot be reinterpreted as (2).

(1) C V C V C V   (2) C V C V C V
     | / / |
     M   M

(3) C V C V C V
     | | |
     M   M   M

Beyond that, in this chapter, I shall illustrate how
virtually the entire class of rules that are distinctly
autosegmental[1] do show up in the autosegmental grammar of Igbo,
and we shall see that these rules interact in the interesting
ways we have come to expect of phonological systems. In sum,
we shall be primarily concerned with the hatuce of autosegmental
derivations and the rules found therein.

In Chapter Three, we shall turn our attention from the
nature of the derivation to the bootstrap operation that begins
it: the linking of autosegmental tiers, a process I have
elsewhere called "melodic association rules." Our attention
there will be focused on problems of accentuation and the nature
of the "star," already outlined in Chapter One.

Chapter Four will deal with the question whether the number
of autosegmental tiers is constant throughout the derivation.
The rather speculative remarks made at the beginning of Chapter
One would suggest that the answer to this question is "no"; we
shall elaborate an autosegmental theory consonant with this
negative answer.

2. Guideposts

Our attention here is, admittedly, directed by our interest
in finding in the tonological facts of Igbo characteristics that
will tell us about the nature of autosegments, rather than about
just anything bearing on phonology in general.
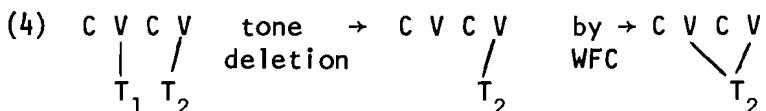
Considered in this light, the phenomena that will interest
us resolve into three classes:

(i) rules that insert or delete segments on one auto-segmental level, but not the others;

(ii) rules that add or delete association lines; and

(iii) rules operating on a single autosegmental level which would be inexpressible (and thus, in theory, impossible), were the remaining features not autosegmentalized. In this last group, only one example occurs; it is a very interesting one, however, pertaining to the interaction of the cliticization of /a/ with the docking of the floating H tone, a process already briefly discussed above in Chapter One, Section 5.

We have already considered the deletion of tone-bearing vowels in an autosegmental framework in Chapter One, Section 3. We shall see below a case of what we would expect equally to find: deletion of a (vowel-bearing) toneme and consequent re-association of the now-stranded vowel, as in (4).

$$(4) \quad \begin{array}{c} C \; V \; C \; V \\ | \;\; / \\ T_1 \; T_2 \end{array} \quad \begin{array}{c} \text{tone} \\ \text{deletion} \end{array} \rightarrow \begin{array}{c} C \; V \; C \; V \\ / \\ T_2 \end{array} \quad \begin{array}{c} \text{by} \\ \text{WFC} \end{array} \rightarrow \begin{array}{c} C \; V \; C \; V \\ \backslash / \\ T_2 \end{array}$$

We shall, furthermore, find a rather large collection of fairly idiosyncratic tonal melodies in the different "tenses" or "verbal forms" of Igbo. Which of these underlying tonal melodies is in fact appropriate is determined by the interaction of semantic and syntactic factors in each sentence. Negative sentences thus choose different tonal patterns from affirmative; and the tone pattern of relative clauses depends on whether the element that is relativized in the lower clause is the subject Noun Phrase or not. Subordinate clauses quite generally have different tonal patterns from non-embedded, or, in Emonds' terminology, root sentences. We shall propose to account for this much as we did for the case of Tiv in Chapter One: the melodies are constructed by the morphological component -- in the lexicon -- and inserted as any lexical material is. We shall require, furthermore, that this process of lexical insertion be sensitive to derived syntactic structure rather than deep structure. This leads naturally to the hypothesis that lexical insertion more generlally occurs at surface structure, or at the end of the cycle. We shall return to this hypothesis in Section 4. As we shall see, although the "surface lexical insertion" hypothesis is not a logical consequent of our other conclusions, it does permit a much tighter, and quite plausible, theory of possible tonal rules.
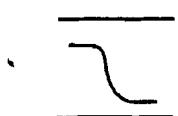
3. The Tonal Inventory

Igbo tones display a three-way contrast superficially, and we shall adopt what is therefore the most "concrete" analysis: there are three tones in Igbo, which we shall call High, Low, and Mid (or, as we called it in Tiv, "Drop"). This three way contrast is illustrated in (5).

(5) a. ʃ m ʃ 'nose' b. ʃ m ʃ 'to crawl'

c. ʃ m ʃ 'to bear fruit'

The Mid, or Drop, tone illustrated in (5b) is a curious one, however. Two successive Mids, unlike two successive Highs or two successive Lows, are not at the same pitch; we see this illustrated in (6).

(6) ó d h ú ó k è   'tail rat' = rat's tail

odhu oke
H M M L

An understanding of this "downstep" phenomena requires that we draw rigorously a distinction between tone and pitch. Although virtually all of the ideas in the immediately following discussion can be found in various places in the literature going back two decades, a strange misunderstanding, it seems, in recent theoretical Africanist analyses has led to many a pointless discussion of the supposed problems that the existence of downstep causes for the theory of binary features, phonological rules, and so on.
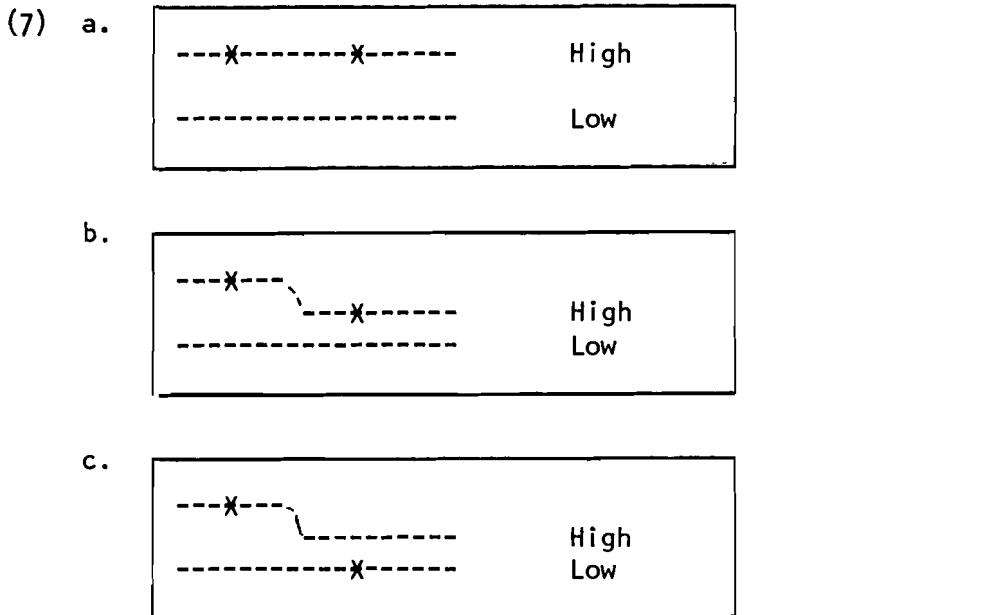
Tone is a characteristic of segments determined by a small set of features (minimally two, but probably larger). These tonal features may (or may not be) autosegmentalized. We shall maintain the hypothesis that all features are binary.

Pitch, on the other hand, refers to a much more superficial analysis of an utterance. It is only a bit more linguistically abstract than the instrumentally-observable fundamental frequency. Unlike pitch, however, the fundamental frequency itself is affected by such factors as consonantal characteristics (voicing, for example), vocal chord sluggishness, and so on.
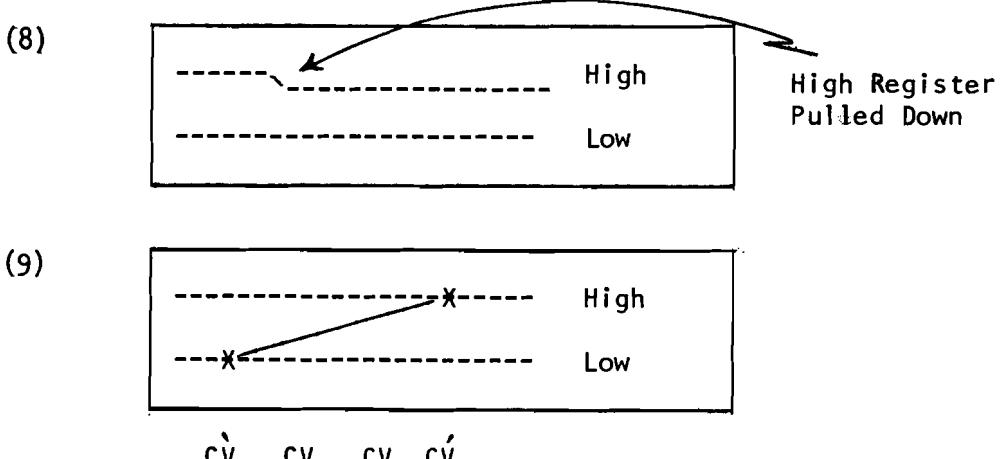
We shall return to the traditional notion of pitch registers; the word "pitch" is used here advisedly. We will consider languages with two pitch registers, for simplicity's sake, but languages may have more than two registers; Yoruba comes to mind as an example with three.

The formal output of the pitch register 'component' is not a segmental representation, but rather a graph like those we have already seen -- in (5), for example. The pitch registral component will predict where the graph will be level, where it will dip and drop, and so forth, on the basis of the segmental representation input to it.

Considering, as we have said, a language with two pitch registers, like Igbo, we construct the pitch diagrams as in (5) out of slightly more abstract pictures which include both the High and Low pitch registers (see (7)).

(7) a.

```
---X-------X------                      High

------------------                      Low
```

b.

```
---X---
       \---X------                      High
-----------------                       Low
```

c.

```
---X---
       _____                      High
-----------X------                      Low
```

The task of the pitch registral. component is, then, first to determine what causes movements of the pitch registers -- pulling of the High Register down, for example, as in (8); and second, to assign each syllable (or in general, tone bearing unit) to a place on the chart. In general this will be done by placing certain of the syllables on pitch registers, as we have with the x's in (7a,b,c), and then drawing lines connecting these fixed pitches. In tone languages, all syllables are fixed on pitch registers; in non-tone languages, it may be that only certain ("accented") syllables are fixed on pitch registers, while unaccented syllables' pitch is determined simply by the curve that connects the determined pitch of accented syllables on either side of them (see (9)). (This last proposal concerning non-tone languages is due to Mark Liberman; see Footnote 1 of Chapter 3.) For now, we may rest with the simpler case wherein each syllable must be placed on one of the pitch registers or the other, as is the case in Igbo.

(8)

```
------
      _____                  High          High Register
                                                       Pulled Down
-------------------                     Low
```

(9)

```
-------------------X-----               High

--X_____                     Low
```

cv̀   cv   cv   cv́
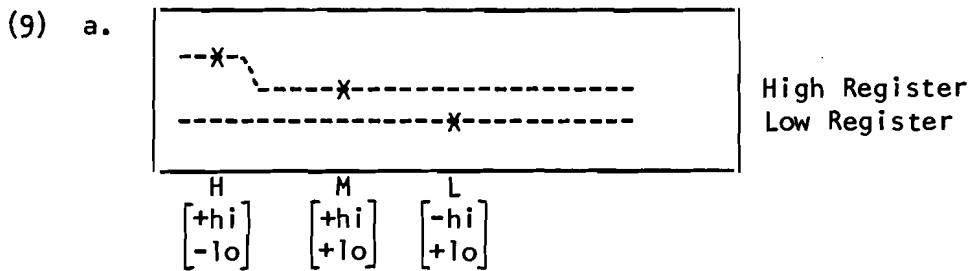
Some of the examples so far have suggested what is, in fact, true: that throughout African langauges, a very common pitch

registral effect is the "pulling down" of the High Register
illustrated in (8). This Drop is rule-governed, and we shall
need a notation for it. Let us denote it as in (10):

(10) High Register → -1 / ... -- ...

That is, we shall use "-1" to mean Dropping of the indicated
Register. This drop is about one to two semitones.

We will specify for Ůhúhu Igbo that High and Mid tones are
realized on the High Register, and Low tones are realized on the
Low Register (see (9a)). Using the feature notation, we may say
that a toneme [αHi] is placed on the αHigh Register, if we call
the High Register the "+High Register" and the Low Register the
"-High Register." Whether or not such a convention is adopted,
the point is that the featuare High determines which register a
tone is realized on.

(9) a.

The High-mid sequence is realized as in (7b). We need mini-
mally, then, a rule approximately as in (12) which performs the
High Register lowering. This rule is not a phonological rule in
the strict sense. It applies in the pitch registeral component,
and reapplies iteratively, left-to-right, as do all rules in this
component.

(12) High Register → -1 /

$$\begin{bmatrix} +hi \\ -lo \end{bmatrix} \quad \begin{bmatrix} \overline{+hi} \\ +lo \end{bmatrix}$$

'H'   'M'

We must ask whether rule (12) could not be expressed more
simply, and by the same token whether it is not part of a larger
phenomenon. The answer to both questions is yes.

There is too much specification in rule (12). A Mid-Mid
sequence, as we have seen in (5b) and (7b), should trigger the
rule as well. The feature "-lo" in the rule made it therefore
too specific. Amending (12) to (13) by dropping this feature
specification, we observe that (13) produces the correct pitch
register modifications for all the examples we have considered
so far.

(13)   High Register → -1 /

$$[+hi] \qquad \begin{bmatrix} \overline{+hi} \\ +lo \end{bmatrix}$$

$$\begin{Bmatrix} H \\ M \end{Bmatrix} \qquad \{M\}$$

It is generally agreed that the existence of Drop tonemes, like the Mid we have been considering, is intimately associated with the existence of the phenomenon of "downdrift," which causes the second of two Highs, when separated by a sequence of one or more Lows, to be slightly lower in pitch than the first High. This is illustrated in (14).

(14)

cv́   cv̀   cv̀   cv̀   cv́
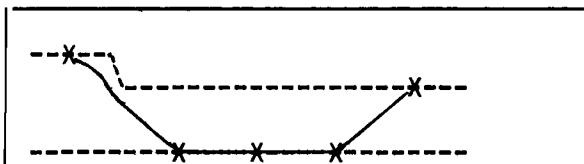
The lowering that occurs between the two Highs is precisely the same amount as occurs between the High and the Mid in (5b).[1a] This intuition is justified when we observe that this downdrift pheno- menon is included in the operation of downstep rule (13) by simplifying that rule -- by dropping the feature specification "+hi" on the second segment. Thus (13) becomes now (15), our final form of the downstep rule for Uhuhu Igbo.

(15)   High Register   -1 /   [-lo] $\begin{bmatrix} --- \\ +lo \end{bmatrix}$

If we take an input like that in (14), rule (15) applies as shown below in (16), lowering the High Register at the sequence H--L. This produces, as we see, the correct final result.

(16)
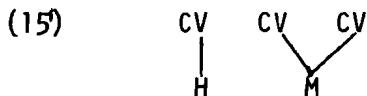
cv́   cv̀   cv̀   cv̀   cv́
rule (15)
applies

Conclusion: From this discussion of the tonal inventory and the operation of pitch registers, two general points important to the tonal analysis of Igbo must be remembered.

First, the sequence Low + Mid is phonetically indistinguishable from the sequence Low + High. This follows immediately from what we have seen so far. The Mid tone and the High tone differ only in
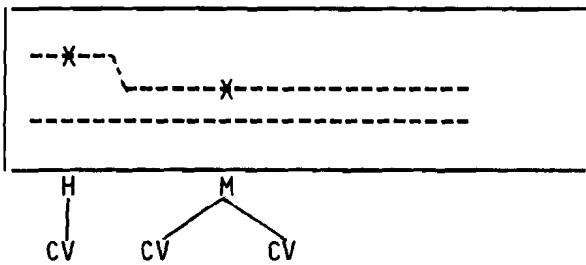
their specification of the feature "low"; and since it is the
other feature, high, that determines which pitch register the
tone is realized on, the specification of the feature lo can
influende only the Register Rule (15). Put differently, tonal
feature specifications can do only one of two things in the pitch
system developed so far: either determine which pitch register
that tone is realized on, or affect pitch registral rules. But
the feature low,,the only feature which distinguishes High tone
and Mid tone (High tone is -low and Mid tone is +low), has no
effect on which pitch register the tone is realized on; that is
determined by the feature high. Furthermore, if the preceding
tone is Low (with the features -hi, +lo) then Rule (15), the
Downstep/downdrift rule cannot apply since its structural
description is not met. Under these conditions, then, High and
Mid are distinct when they follow either High or Mid tone.

The second point that must be clear is what the phonetic
shape of a representation like (15') is.

(15')   CV   CV   CV
               | \   /
               H   \ /
                       M
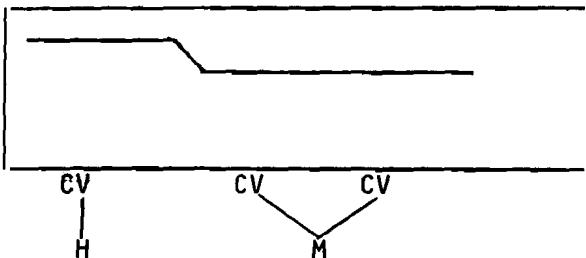
If a Mid tone is associated with two syllables, as in (15'),
then these syllables will be level in pitch. The first of the
two will be downstepped with respect to the High-toned syllable
in (15). In short (15) will produce a pitch registral graph as
in (16').

(16') a.
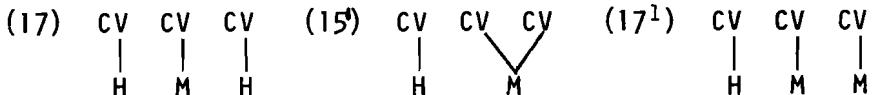
b.
i.e.

Looking at (16b), we see that this is the pitch contour that
would have been created by the sequence in (17) (compare this
with (15'), repeated here).

$$(17) \quad \begin{matrix} CV & CV & CV \\ | & | & | \\ H & M & H \end{matrix} \qquad (15') \quad \begin{matrix} CV & CV & CV \\ | & \diagdown\diagup \\ H & M \end{matrix} \qquad (17^1) \quad \begin{matrix} CV & CV & CV \\ | & | & | \\ H & M & M \end{matrix}$$
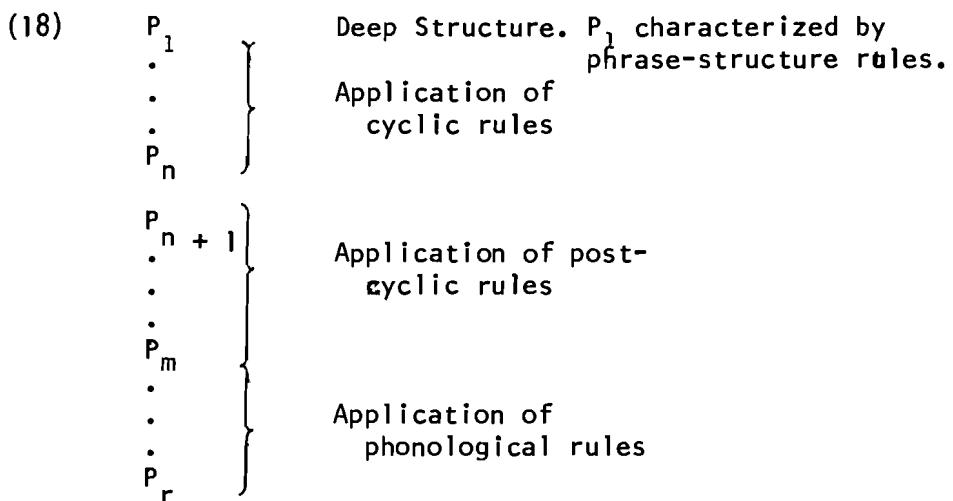
The point, then, is that a Mid tone spread over two or more syllables is phonetically indistinguishable from a Mid follow-ed by a sequence of Highs, though it is distinct from (17^1) (see (6) above).

We shall find, throughout this chapter, that these two surface neutralizations, though they may cause trouble to linguists, do not present problems to the language learner if the learner is inffact endowed with the autosegmental linguistic faculties we are positing. Information about the general systematic structure of the language makes it clear to the language learner which form is occurring, as we shall see.

4. The Derivation

Accepting the essentials of Chomsky (1965, 1970) and Chomsky and Halle (1968), we view the structure of a derivation as in (18).

$$(18) \qquad \left.\begin{matrix} P_1 \\ \cdot \\ \cdot \\ \cdot \\ P_n \end{matrix}\right\} \quad \begin{matrix}\text{Application of}\\ \text{cyclic rules}\end{matrix}$$

Deep Structure. P_1 characterized by phrase-structure rules.

$$\left.\begin{matrix} P_{n+1} \\ \cdot \\ \cdot \\ \cdot \\ P_m \\ \cdot \\ \cdot \\ \cdot \\ P_r \end{matrix}\right\} \quad \begin{matrix}\text{Application of post-}\\ \text{cyclic rules}\\ \\ \\ \text{Application of}\\ \text{phonological rules}\end{matrix}$$

Each pair (P_i, P_{i+1}) corresponds to the application of a linguistic rule, where, in particular, P_i meets the structural description and P_{i+1} satisfies the structural change of that rule when P_i is the structural description.

P_i is a phrase-marker, or tree, with no lexical items as terminal elements. Lexical items are inserted by the process of "lexical insertion," whose place in the derivation we return to in a moment.

Some phonological rules may be best viewed as applying inside the lexicon, and the domain of the application of such rules would be at most one lexical item (it might, of course, be smaller -- just one syllable, for example). There are other phonological rules which operate across lexical items -- there may be an inter-

action, that is, between the phonological material of successive words. Let us call this second sort of phonological rule, which cannot be construed as a lexical process, wide-domain phonological rules. The discussion below concerning the ordering of lexical insertion in the derivation will be concerned with these wide-domain (or sandhi) phonological rules, rather than with lexical phonological rules.

The application of rules can be grouped into syntactic and wide-domain phonological (which includes tonological), with the application of all syntactic rules preceding the application of all phonological rules.[2] Syntactic rules can similarly be divided into cyclic and post-cyclic rules; some would argue that the latter class is empty.

Our concern, as we have noted, is the section of the derivation from P_m to P_r, the phonological part. We may note that the insertion of the phonological information for each lexical entry in the derivation must have been inserted by the time of the first application of a phonological rule. We must ask if we can determine the point of lexical insertion in the derivation any more precisely.

In Chomsky (1965, 1970), Chomsky proposes that lexical insertion consists of a set of transformations applying before all other transformations -- applying in (18) before the cyclic syntactic rules noted there. For our purposes there are two empirical criteria for deciding where in the derivation lexical insertion occurs. First, the phonological information which lexical insertion provides must occur before P_m, as we have seen, the first phrase marker in the phonological derivation. The earlier phonological information is inserted before P_m, however, the more ad hoc conditions must be placed on the functioning of the grammar to prevent interaction of this phonological information with the functioning of syntactic rules. No transformational rule, that is, is conditioned by the phonological character of the items involved. We want to find it no accident that phonological information becomes necessary to the derivation, then, at the point where it in fact becomes available to it. As I say, this is what we would hope, in the best of all possible worlds, to find; it may or may not be true.

The second empirical criterion for determining where lexical insertion occurs depends on the type of contextual information needed by the rules of lexical insertion. Again, in the best of all worlds, lexical insertion will crucially have to be stated either at essentially deep structure level before the application of transformations -- or at "shallow structure," defined either as P_n or P_m in (18), or some other well-defined linguistic level. That is, if the correct point for lexical insertion is at deep structure, we will find evidence that the context for lexical insertion must be stated before the application of transformations;

if the correct locus is at P_n in (18), we would find cyclic rules which must create the context for lexical insertion, but post-cyclic rules would in turn destroy the necessary context. In each case we would find a level in the linguistic derivation where the contexts for lexical insertion were optimally stateable.

Again, if ours is the best of all possible worlds, we will find that the two empirical criteria -- one based on the necessity for phonological information, the other based on proper context -- will point to the same level in the grammar for lexical insertion. This, of course, need not be so; if it were not true, one might, as a second best, hypothesize that lexical insertion occurs necessarily at one level, but is contextually defined by some other level, in a sense that would then have to be made more precise.

Our conclusion based on the evidence from Igbo tone shall be that the second best need not be settled for, that both criteria are compatible with lexical insertion occurring at the point labelled P_m in (18) -- in short, that lexical insertion occurs at surface structure.

Our argument goes like this: considering a very wide range of Igbo tonal data, we are led to posit a set of about ten simple autosegmental rules. These rules relate the superficial forms with a more abstract level. The tones for nouns at this more abstract level are simply the lexical tones. The situation is more complex for verbs and their affixes. We find that the underlying tonal patterns for the verbal system are determined by a complex -- and essentially idiosyncratic -- interaction of lexical, syntactic, and semantic factors.

We are led -- much as was Chomsky in *Remarks on Nominalizations* -- to propose a reason for the difference in character between the well-motivated phonological/tonological rules, and the essentially arbitrary rules necessary to derive the tonal formulas for the verb and its affixes. The verbal tones, we shall say, are inserted at the point of lexical insertion,and thus are the product of the morphological component.

This lexical insertion, however, must be sensitive to the derived syntactic structure -- the structure after relativization, in particular. We would be led by our second criterion to posit lexical insertion at the frontier between the syntactic and the phonological components of the grammar, or at the end of each cycle.

One consequence of this is that there will never be found phonological rules which have left traces in the segmental phonology of the word-order that existed prior to the application of some syntactic transformation.[3] For example, there could not exist (the claim goes) languages like the following, which leave the trace of the pre-passive-rule form on the segmental phonology.

(a) Suppose there was a language much like English as we know it, with a rule palatalizing t → č before y. This phonological rule, furthermore, applies before the passive transformation. Thus the underlying form "Did she meet you" becomes, by this palatalization, "Did she meeč you?"

If the passive transformation were to apply, this would produce a derivation like (A).

(A)   She past meet you   underlying
      She past meeč you   palatalization
      You were meeč (by her)   passive

(b) Consider another example, again an odd dialect of English with a rule devoicing word-initial voiced stops after a voiceless stop:

Devoicing   [+stop] → [-vcd]/ $\begin{bmatrix} -vcd \\ +stop \end{bmatrix}$ ## ___

Such a language would produce derivations like (B), if Devoicing could apply before the Passive rule.

(B)   You met Betsy   Underlying

    (a) [You met Petsy]   Devoicing

    (b) [Betsy left last night.]
        Someone met Betsy at the station   Underlying
        Someone met Petsy at the station   Devoicing

    (c) [Petsy was met at the station (by someone)] Passive

As I say, it is language systems like these that are ruled out as impossible by the ordering of wide-domain phonological rules by the ordering of phonology after syntax.

The question of the placement of lexical insertion with respect to the application of transformations is a subject deserving of more than these few pages here. Indeed, the extent to which cyclic rules should be viewed as transformations rather than "lexical redundancy rules" -- the term is perhaps less than apt -- is called into question if lexical insertion is carried out after cyclic transformations. The remark calls to mind certain recent suggestions made by Joan Bresnan (1976), but they are equally reminiscent of the older proposals of such diverse linguists as Bowers and McCawley.

Some speculations on this issue are worth dispelling. It has been suggested, for example, by Bresnan that there is a clustering of rules that are (i) structure-preserving, (ii) bounded, and (iii) lexically-governed (Bresnan suggested this,

for example in a public address (1976)). If this is true, it could be explained, as she suggests, by saying that these (cyclic) rules are not transformations at all, in Chomsky's sense, but are rather lexical rules relating the uses of the same or related lexical items. Thus there would be <u>no</u> transformational relation, on this account, between the active and the passive forms of a sentence.

If this was true, furthermore, my suggestion that lexical insertion occurs after all cyclic transformations becomes trivially true, since there would no longer be any cyclic transformations.

The notion of "structure-preserving" transformation would, the reasoning continues, be eliminable, since the "reason" that the direct object is moved into the subject position by passive -- a paradigm example of a structure-preserving rule -- is simply that the phrase-structure rules determine the positions into which lexical insertion may place NP's, whether these NP's be logical subjects (as in active sentences) or logical objects, but now placed in subject position with a verb marked in the passive voice.

Such a proposal, while attractive on the face of it, is ultimately not acceptable. Even if we ignore the question of how to deal with idiom chunks -- while this has traditionally been an argument for a passive transformation, Higgins (1974) presents evidence that this is a non-argument -- the claim that rules cluster in the way noted above is highly dubious. In fact, given the eight categories that the three binary distinctions listed above give, either seven or all eight have attested members.

Structure-Preserving	Bounded	Lexically Governed	Rules
No	No	No	?Topicalization (see note 1)
No	No	Yes	"So-preposing,""the more the merrier" movement (but see note 1)
No	Yes	No	Subj-Auxiliary inversion, Verb-fronting in German
Yes	No	No	?Topicalization (see note 1)
No	Yes	Yes	Evidemment-movement in French (see note 2)
Yes	No	Yes	(see note 1)
Yes	Yes	No	Stylistic Inversion (French)
Yes	Yes	Yes	Passive, Raising to Subject

Note 1:   Whether Topicalization is a movement rule is unclear,
so its placement in this chart is doubly unclear. Furthermore,
whether Topicalization (if a movement rule), *So*-Preposing, and
so forth, are structure-preserving or not depends on one's ana-
lysis of the putative Complementizer position (see Bresnan (1972)
for one position on this).

Note 2:   This is the rule that changes *(Que) il est evidemment
venu* to *Evidemment qu'il est venu* in French. See Dubuisson and
Goldsmith (1976) for an analysis of this rule.

The conclusion of this is simply that if lexical insertion
occurs after cyclic rules apply, it is not in the trivial sense
that Bresnan's hypothesis would lead to. We maintain that there
are transformations that are structure-preserving, that are
lexically-governed, and that actually apply before lexical
insertion. However, we cannot pursue this topic here; we return
to Igbo and its tones.

5.   Verbal Forms

The discussion so far in this chapter has been preparatory
and anticipatory; we turn now to a detailed analysis of Igbo tone.

Most nouns in Igbo are bisyllabic, with the form VCV. There
are also a fair number of monosyllabic nouns of the form CV, and
some with more than two syllables. We shall restrict our atten-
tion, as Green and Igwe[4] do, to one and two-syllable nouns.

Monosyllabic nouns are all High in tone; bisyllabic nouns may
be High-High, High-Low, Low-High, or Low-Low. The tonal behavior
of a noun is fully determined by its lexical tone, in the sense
that what tonal rules a noun undergoes is determined by its tone
only and not by its non-tonal characteristics.

Verbs divide into two categories, the Low and the non-Low
(or High). This distinction is a lexical one, and it frequently
occurs that the distinction is neutralized on the surface. When
the distinction is maintained on the surface, the Low class is
usually Low, while the non-Low class may be either High or Mid,
depending on the specific context. All verbs are CV, except com-
pounds, which are CVCV.

There is a highly versatile set of verbal suffixes which mark
various aspectual and other core semantic functions. These verbal
suffixes divide into two categories, which, following Williams, we
shall analyze as toneless and tonally-specified (Class I and II,
respectively). The toneless Class I is much more common than the
Low-toned Class II.

There are two verbal prefixes which occur in a fixed order when they co-occur. The first is the optional relative marker na, which marks the clause as being a relative clause. The second is a (or its counterpart e when the verb stem triggers tense vowel harmony, changing the a to e), which is again a sort of aspectual marker. The presence or absence of the prefix a, however, also marks which of two series, the I Complete or the II Incomplete, the tonal pattern is slected from. That is, the tonal pattern of a verb and its affixes is determined essentially by whether the sentence is embedded or not, and, if so, what sort of embedded clause it is. For each of these types, though, there is a I Complete form, and a II Incomplete form. The II Incomplete form, furthermore, is marked underlyingly by the a prefix.

The tonal melodies of the verb and its affixes consists of four components, each of which has a specific syntactic locus (see Figure (19) below). The first is the floating H tone that was discussed in Chapter One, Section 5. The second is the tone of the a prefix; the next is the verb stem tone, and the last is a suffixal tone.

The Chart (19) below summarizes our analysis of the tonal melodies of the eight simple affirmative forms. As we have observed, the tonal melody depends in part on whether the clause is embedded or not. Non-embedded sentences are called here "Root," following Emonds' terminology (Green and Igwe call these "Main" forms). Among embedded clauses, we must distinguish relative clauses and non-relative clauses. This is simple; relative clauses are those embedded clauses which are preceded by a head, or antecedent, noun; and in just these cases the head noun will also semantically correspond to an empty position in the embedded clause. We must furthermore syntactically distinguish between relative clauses where the function of the antecedent in the lower clause is as a subject (as in the English *the man who was here*) and those where the antecedent plays any other role in the relative clause (e.g., as in *the man I saw* --). We shall call the first type "Relative/ subject," and the second type "Relative/non-subject." Note that there are no words in Igbo corresponding to wh-words or a comple-mentizer like *that* in relative clauses.

The Chart below is divided up into two major halves, one for the I Complete and one for the II Incomplete. The numbers I and II are Green and Igwe's; I have added the terms "complete" and "incomplete" to suggest parts of the semantic difference between the two types of sentences. These English words are perhaps more misleading than helpful, but there are none better that I could find, and it is helpful to have some mnemonic label to dis-tinguish the two series of melodies.

(19) Simple Affirmative Melodies

I Complete

	Floating /	a-prefix /	Stem /	Suffix
Root	-	-	L	-
Subord.	-	-	B	H
Relative/ Subject	H	-	M	H
Relative/ non-subj.	H	-	L	-

II Incomplete

	Floating /	a-prefix /	Stem /	Suffix
Root	-	B⁻	B*	H
Subord.	H	-	L	▾
Relative/ Subject	H	-	M	H
Relative/ non-subj.	H	-	L	-

"B," here as in Chapter One, stands for "Base"; see Chapter One, Section 4.2.

6. I Complete/Root

The normal underlying tone pattern for this form consists, as (19) indicates, of just a Low tone on the verb stem.

A simple sentence of this form, constructued from the lexically High verb cí 'carry' and the object àkhwá 'eggs' is (20).

(20) ó cì   àkhwá

he carry eggs   'he must carry some eggs'

A lexically Low verb also appears with a Low tone in this form -- for example, zà "sweep."

(21) ó̩ zà   ú̩lò̩

he sweep compound   'he must sweep the compound'
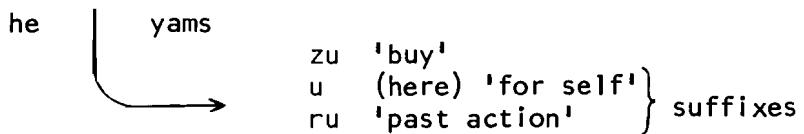
In general, then, verbs are Low in tone in this form, despite their lexical tone specification. There are a small number of exceptions to this which we shall return to below.

When Class I suffixes, the toneless suffixes, are added to this form, we expect a situation as in (22), which, by the Well-formedness Condition, should become (23), which is what we find. Class I suffixes attached to the verb stem in the I Root form are Low in tone, all other things being equal.

(22) o   zu + u + ru   (23) o   zu + u + ru
      | | |
      H   L   H   L

ó zùùrù jí   'he bought himself yams'

he   yams

zu   'buy'
u   (here) 'for self' } suffixes
ru   'past action'

No tonal information particular to the I Complete Root form need be specified regarding the verbal suffixes; the Well-formedness Condition is all that is necessary for the type of forms considered in (22) and (23).

Something further does need to be stated, however, to derive correct surface tone patterns for the noun phrase subjects in the I Complete Root form. Consider the pattern in (24). When the subject ends in a Low tone, it displays its lexical tone; when, however, we would expect it to end in a High tone, the subject in fact ends in a Falling tone, as discussed in Section 2.1 of Chapter One. (24) repeats (13) and (14) of Chapter One.

(24) Ézè cì àkhwá   Ézè   'chief'   H L

      Ùwà cì àkhwá   Ùwà   (a name)   L L

      (NP) 'carry''eggs'

      Ékwê cì àkhwá   Ékwé   (a name)   H H

      Àdhâ cì àkhwá   Àdhá   (a name)   L H

As analyzed in Chapter One, these tonal modifications are the results of the I Root Flop rule, expressed in (25).

(25) V   [C V   I Root Flop Rule
         VP   /
              T

This causes the tone of the verb stem to be shared by the final syllable of the subject noun phrase when that subject is outside the VP -- that is, not cliticized onto the verb.

This simple rule can act as a useful example to illustrate the linguistic claims of this notation. Although one may say, as a first approximation, that the degree of linguistic unpredictableness of a rule is proportional to the number of symbols in the rule, or, more accurately, to its measure of complexity, this is only a first approximation.

The rule claims, first, that the language would be simpler if there were no such rule, and if the subject Noun Phrase in (24) always displayed its underlying tones. The existence of the specification "VP" (verb phrase) in the rule claims that, formally speaking at least, the rule would be simpler if it applied to any successive pair of words, rather than only to subject noun phrases. Functionally, it seems clear that such a generalized rule would, however, be useless, since its application would mark nothing at all.
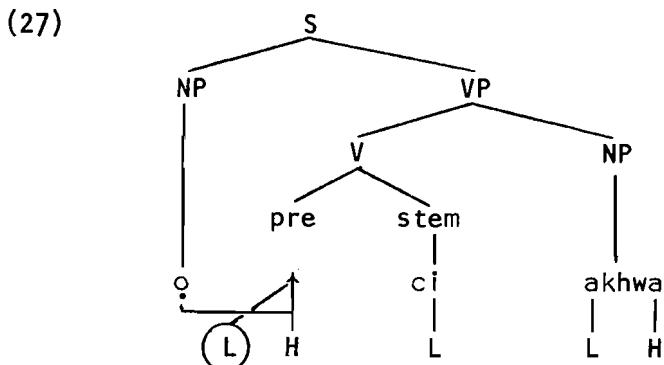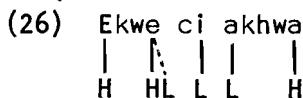
Given, furthermore, that this rule changes the derived tonal pattern, it must do this either by deletion of tonemes, addition (epenthesis) of tonemes, or addition or deletion of association lines. We shall assume that changes of association lines are always linguistically simpler than addition or deletion of segments (a rigorous statement of this assumption will be made below). In this case, this means that the tonal rule creating the contour tones in (24) must add an asosciation line; and since the rule must specify which segments are involved in the addition of the association lines, the four segments mentioned in rule (25), I Root Flop (two Vs, one C and one T) are the minimal number possible. In sum, the conclusion is that although no rule is simpler than some rule, the simplest rule is a Flop rule (a rule adding an association line); furthermore, such a rule requires minimally the specification of V's, C's, and T's as in (25).

We remarked above that the Flop rule must be made formally slightly more complex by specifying some syntactic information about the environment (the expression "VP" in the rule). This addition, it was pointed out, was necessitated by the simple fact that the Flop rule did not apply to any arbitrary sequence of words, but only at certain subject-verb boundaries. We should note now, however, that this additionally explains <u>why</u> the Flop does not apply in (20), (21), or (22), where the subject is a cliticizing pronoun, and thus, as we saw in Chapter One, Section 5, is moved by a cliticization rule into a position <u>inside</u> the VP. Thus there is no VP boundary in the appropriate place, and the Flop rule cannot apply.

In Chapter I, Section 5, a Floating H tone was discussed which, under certain circumstances would "dock" onto the right end of a subject NP. Looking at the data in (24), a linguist or a language-learner might be tempted to impute the Falling tone that is induced in the second two examples to a floating L tone. We should like to propose that, although the linguist may be led to that, the language-learner will not, for the reason that a Flop rule (adding
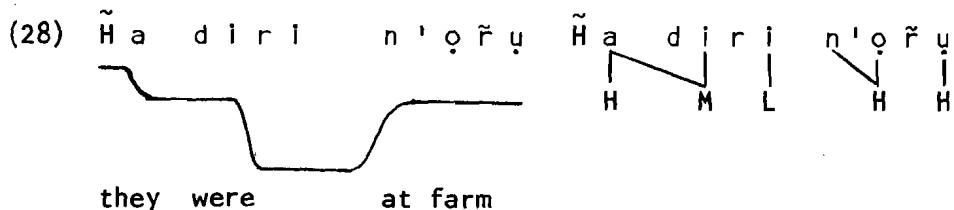
an association line) will always be preferred over (evaluated as
less costly by the evaluation measure) when compared to either an
epenthesis rule or a floating tone as discussed in Chapter One,
Section 5, all other things being equal. In short, the Falling
tone in (24) can be analyzed as a Flop or assimilation, since the
following tone is Low, on the verb stem. Therefore, it will be
interpreted as an assimilation. Such an analysis is not
available for the type of facts discussed in Section 5 of Chapter
One (discussed in more detail in this chapter), and therefore a
floating tone must be posited.

The way to test such a claim is to consider the two alter-
native sets of rules that are consistent with some set of data that
we may assume is learned earlier by language-learners. The claim
that the Flop rule is preferred is tested by determining whether
the Flop rule   makes more correct predictions outside the
original data set than does the alternative solution. We have
already seen that the Flop rule correctly predicts that clitici-
zing pronouns will not undergo the Flop rule. The alternative
account which proposes a Floating Low tone suggests that (24c)
arises from a structure like (26). If the subejct were a pronoun,
as in (20), the Floating Low analysis would predict cliticization
as in (27) (parallel to (72) in Chapter One), and a derived tone
pattern with a Rising (LH) pattern on the pronoun subject, parallel
to the derived Falling in (72) of Chapter One.

(26)   Ekwe ci akhwa
       | | | | |
       H HL L L   H

(27)
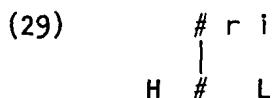
In fact, this is wrong, as we know; the subject clitic tone is High,
not Rising.

A second type of empirical test was mentioned in Section 2
of Chapter One. If the tone of the Verb stem should be anything
other than Low (which is all that we have considered so far), then
the Flop analysis predicts that the subject NPs tone should assimi-
late partially to whatever that tone was (note that the tone speci-
fied for the verb stem in the statement of the Flop rule in (25)
is simply "T" -- i.e., any tone at all). If the verb stem should
for some reason be Mid, then, the Flop analysis predicts that the
subject should take on a contour tone whose second half is level
with the following Mid tone, as illustrated in (28). This repeats
the point made in examples (19a and b) in Chapter One.

(28) H̃a d i r i   n ' o̧ r̃ ų   H̃a d i r i   n ' o̧ r̃ ų

H   M   L   H   H

they   were   at farm
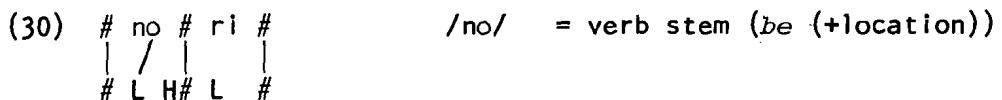
This type of raising of the stem tone from the normal L of the I Root form to a Mid tone occurs in three cases: when the stem is immediately followed by a Class II suffix, as illustrated in (28); when the stem has no suffixes (which is unusual in itself) and is followed by the pronoun ànyì ('us') as an object; and for a small class of exceptional non-Low verbs which appear in the I Root form as Mid rather than Low. We will review these three exceptional cases in order.

The Class II suffixes, unlike the Class I suffixes, have an inherent tone: Low. They furthermore affect the tone of the syllable that precedes them, raising it to Mid or High. We shall ascribe this to the influence of a floating High tone on the left side of the suffix. The general underlying form of the Class II suffix, then, shall be as in (29).

(29)   # r i
            |
    H #   L

(Whether the ri and the L should be associated underlyingly or not we leave open for now.)
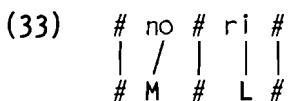
When such a suffix occurs on a normal verb stem in the I Root, then, we get a situation like (30). The first L is the L constituting the I Root melody; the other H and L correspond to the lexical entry for ri as illustrated in (29).

(30) # no # ri #   /no/   = verb stem (be (+location))
      | / | |
      # L H# L #

(I have assumed in (30) that an initial process has linked the verb stem with tonal melody L; the illustration works as well without this assumption. Note the use of association lines to associate corresponding word boundaries, as mentioned in Chapter One, Section 6.)

The Well-formedness Condition adjusts (30) to (31), and a late rule of LH-simplification (32) in Igbo converts (31) to (33), the correct output.

(31) # no # ri #
      | /\ \ | |
      # L H # L #

(32)   V   V   <u>LH-Simplification</u>
          /\   → |
          L H   M

(33)   # no # ri #
        | / | | |
        # M # L #
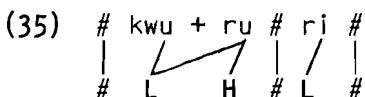
If a Class II suffix should be appended not immediately following the stem, as in (34), we get the expected form.

(34)   kwu = speak; ru = Class I (toneless) suffix

       # kwu + ru # ri #
       | / | | / |
       # L   H # L #   Output: kwùrúrì
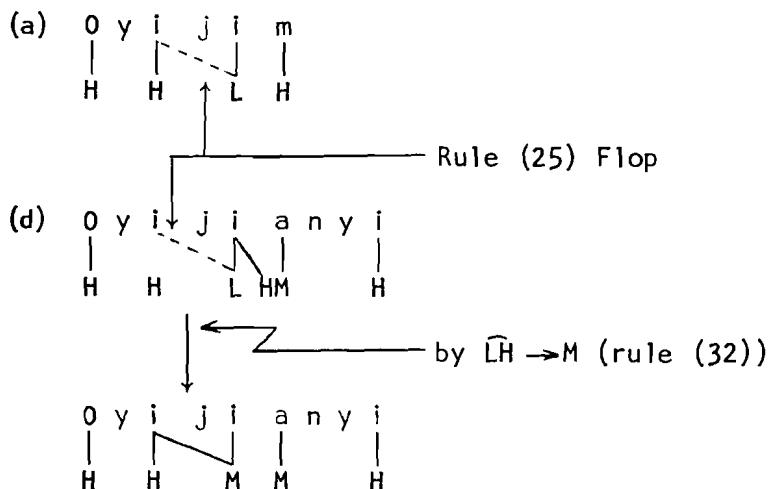
(Note that we can be sure that the L tone of the stem does not spread to the first suffix, giving us (*35), because (32) LH-Simplification would change this to (*36), rather than the correct form (34).

(35)   # kwu + ru # ri #
       | /⎯7 | | / |
       # L   H # L #

(36)   # kwu + ru # ri #
       | \/ | | |
       #   M   # L #   (*kwùrúrì)

The second case raising the verb stem to M is found when the object is anyi ('us') and where there are no suffixes. Consider the paradigm below (37).

(37) a. óyî jì ḿ   "I am feverish"

      b. óyî jì gí   "You (sg) are feverish"

      c. óyî jì yá   "He is feverish"

      d. óyⁱ jⁱ ányí   "We are feverish"

      e. óyî jì únù   "You are feverish"

      f. óyî jì hã́   "They are feverish"

(Each of these sentences literally means "Fever grips me (you, him, etc).)

(a)   O y i j i m
      | | | |
      H   H   L   H

——————————— Rule (25) Flop

(d)   O y i j i a n y i
      | | | | |
      H   H   L HM   H

——————————— by $\widehat{LH} \rightarrow M$ (rule (32))

      O y i j i a n y i
      | | | | |
      H   H   M M   H

As we see, the tone of the verb ji is low everywhere except where the object is anyi. And just there, the tone of the final syllable of the subject falls to Mid rather than Low, as pre-dicted by the Flop rule.

Third and finally, two verbs -- wu, the copula of identity and ka 'surpass' -- take the Mid tone in this form normally. We thus get Ó kà ḿmá rather than Ò kà ḿmà ("it is better" -- it surpasses good," literally).

I have no examples of non-clitic subjects preceding ka or wu in this form but the prediction is clear; we expect Subjects ending on a High tone to fall to Mid rather than Low.

So much, then, for the I Root form, whose melody consists just of an L tone on the verb stem. We turn our attention next to the I Subordinate form.

7. I Subordinate

This form appears only in subordinate clauses, introduced by an element (a complementizer) ka or ma. The tonal pattern here differs in several ways from that of the I Root form. First, the tone of the subject NP is unaffected here in the I Subordinate; second, the tone of the verb stem depends directly on the under-lying verb. An underlyingly Low verb will be Low; an underlyingly non-Low verb will be High. I shall denote this general occurrence by "B," for "Base," as in Chapter One, and as in Figure (19) above. We note, informally for the moment, that suffixes in this form are High-toned. In the following two examples, the preposed clause with the meaning "if" contains the verb in the I Subordinate form.

(38) Low-verb: Ányí fèé únù únù àdý ányí
we pass you   you will follow us

"If we pass you, you will follow us."
(G&E, p. 81)

fè is here the verb stem: a Low-toned verb.

e is a suffix

(39) High-verb: Ányí círí ákhý / ńdí ŋ̀wè yá gà íjídhé
we pick kernels
up
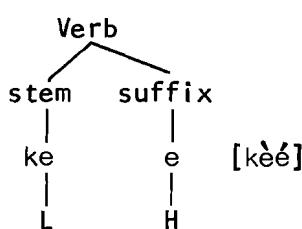
"If we pick up kernels, the owners will
arrest us" (G&I, p. 79)

verb is cí (High toned); ri is a suffix
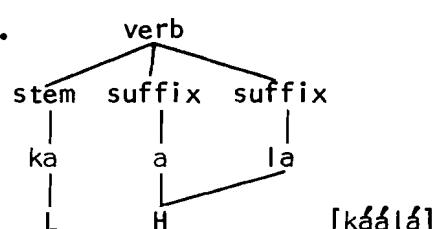
We observed on page 69 above that there are two classes of suffixes (I and II) and that, following Edwin Williams, the more common Class I suffixes are toneless. This corresponds to a formal representation entirely on the non-tonal tier, with no tonal segments in the lexical entry for the Class I suffixes. In the I Subordinate form this assumption leads directly to the postulation of suffixal H tone which is lexically specified only on the tonal autosegmental tier, but not on the non-tonal tier. Thus the H-tone and the Class I suffixes are formally parallel: each is specified only on one tier.

We are led to this suffixal H-tone most obviously by the fact that Class I suffixes are High-toned across the board in the I Subordinate form. The tone of the Class I suffixes is arrived at differently here than in the I Root forms, where the tone of the verb stem was passed onto the Class I suffixes. Thus for Low-toned verbs, we get forms like (40); for High-toned verbs, we get forms like (41).

(40) i.
```
         Verb
        /   \
    stem    suffix
      |       |
     ke       e     [kèé]
      |       |
      L       H
```

ii.
```
          verb
        /  |  \
    stem suffix suffix
      |    |     |
     ka    a     la
      |    |_____/
      L       H        [káálá]
```

(41)
```
          Verb
        /  |  \
    stem suffix suffix
      |    |     |
     gba   tu    la
      |    |_____/
      H       H          [gbátúlá]
      |       |
    stem    suffix
```
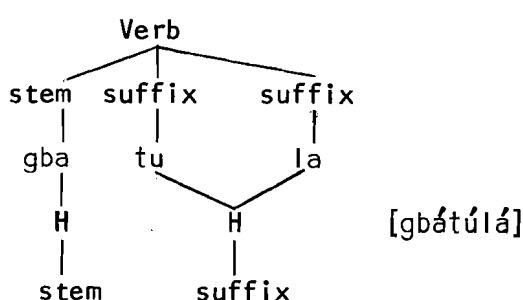
Given that the Class I suffixes are inherently toneless, the language-learner will analyze the High-tone on the suffixes as an autosegment, which here is a High-tone particular to the I Subordinate form.

As Williams furthermore correctly observed, the postulation of an H-tone suffix -- a "floating tone" -- provides an explanation for what seems to be quite a different tonal process. This is what I shall call "Object tone mutation," and refers to a change of tone undergone by the noun phrase immediately following the verb in several of the verb forms (this process of object tone mutation applies in the genitive construction also, as we shall see).

In a word, it is the presence of a floating H tone immediately preceding an NP that triggers Object Tone Mutation. Object Tone Mutation causes the following changes in mono- and di-syllabic nouns:

(42)   Tone Before Object Mutation   After Object Mutation

H	M
HH	H M
LH	M H
H L	H L (no change)
L L	L L (no change)
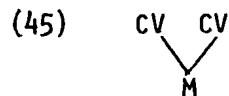
If we attempt to formalize this tonal process, we note first that three forms undergo <u>some</u> change; two undergo none, the two that end in L tone. We might make a first stab at a rule like (43), leaving off the triggering environment, the floating tone.

(43)   H → M / -- # ; or, more precisely, (a)
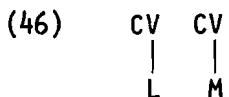
T → [+lo] / -- #   (b)

Such a rule is literally as simple as a rule can be, containing just one feature mentioned.[6] We observe that it is empirically satisfactory for four of the five forms: it expressed the fact that forms ending in L = [-hi,+lo] undergo no change; it also correctly states the change undergone by the forms H and HH. It wrongly predicts, however, that LH will become LM. Recall, too, that LM is superficially indistinguishable from LH: thus (43) as written predicts a "systematic" change for the LH forms, but with no distinguishable surface realization of that change. As (42) shows, however, the actual result is "MH."

Looking at the surface description "MH," we remark that this pattern may reflect either of two structures -- (44) or (45) -- in the autosegmental framework.

(44)   CV CV   (45)   CV CV
          | | \ /
          M   H   M

As noted above, rule (43) applied to the LH forms gives (46),
which, obviously, is neither (44) not (45).

(46)   CV CV
         | |
         L M
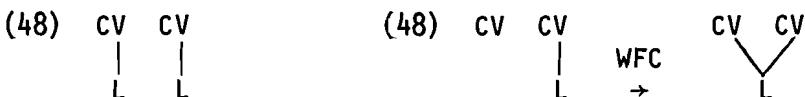
If the effect of rule (43) is augmented by a rule deleting
the L in (46) -- thus, in effect, permitting (43)'s change of
the H to M to be heard on the surface, no longer masked by the L
-- then (45) results, for the M immediately reassociates with the
first syllable once that syllable loses the tone it was associated
with.

The rule required to state such a deletion is again extremely
simple; it is of cost 1, as we see in (47).

(47) L → Ø / ## --

Note that the other place where (47) could operate besides on
the LH form -- that is to say, on the LL form -- its effects are
not noticeable. If the initial L is deleted by rule (47), the
remaining L reassociates, spreading over both syllables, as in (48).

(48) CV CV   (48) CV CV   CV CV
      | | | WFC   \ /
      L L   L →   L

Summarizing, the following analysis for the Object Tone Mutation
is selected by the theory:

(49)   Object Tone Mutation
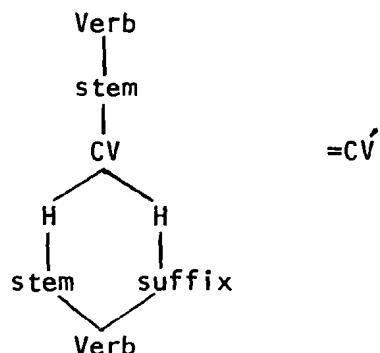
    T → M   / [-low] $_{affix}$]##(T) --#

    L → Ø   / [-low] $_{affix}$]## --

What is important to us now, however, is the existence of the
floating H tone suffix, rather than the precise formulation of the
Object Tone Mutation rule. We must furthermore keep two goals
separate: arguing that the correct analysis for Igbo contains a
floating H tone here, on the one hand; and, on the other, arguing
for the present autosegmental theory on the grounds that it directly
forces the grammar of Igbo to contain this floating H tone. The
former task must await more data; the latter is contained in each
discussion where this H tone appears. We have, for example, already
seen that the Class I suffixes High tone in the I Subordinate form
have required positing the floating H tone there. That the Class
I suffixes were toneless, the reader will recall, was justified on
the grounds that the Class I suffixes appeared Low-toned in the I
Root form, but High-toned in the I Subordinate; and more generally,
either take on the tone of the verb stem or the tone of the
floating H-suffix.
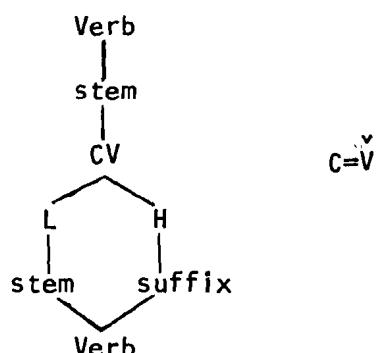
Let us consider a prediction made by the system developed so far. The verb stem is underlyingly assigned its lexical tone, either Low or High; a floating H tone is placed as a suffix. We have not considered any cases so far where there are no suffixes on the non-tonal tier, which I shall henceforth call the *syllabic* tier. A strong consequence of the autosegmental hypothesis, of course, is that the floating H tone may continue to exist whether or not there are syllabic segments associated with it.

We expect to find, then, structures as in (50) for High-toned verbs, and as in (51) for Low-toned verbs.

(50) High-toned Verbs

Verb
|
stem
|
CV   =CV́
/   \
H   H
| |
stem   suffix
\   /
Verb

(51) Low-toned Verbs

Verb
|
stem
|
CV   C=V̌
/   \
L   H
| |
stem   suffix
\   /
Verb

However, we have already postulated a rule (32) of Low-High simplification:

(32)   CV   CV   Low-High Simplification
          /\   → |
          L H   M

This rule would apply in (51), changing the rising tone to a Mid-tone. Thus we expect the forms in (52) in the I Subordinate without suffixes.

(52)   Hightoned Verb: CV́

Lowtoned Verb: CV̀

In fact, however, the tone of the verb in the I Subordinate form without suffixes is more complex than hypothesized by (52). High-toned verbs are, as predicted, High; however, the more interesting Low-toned verbs are Low before a High tone, and Mid otherwise, as illustrated in (53).

(53)   (a)   ányị̧   kè   ánu̧   (HH)   ...meat

       (b)   ányị̧   kè́   àkwhá   (LH)   ...eggs

       (c)   ányị̧   kè   áz̃ù   (HL)   ...fish

       (d)   ányị̧   kè́   àlà   (LL)   ...land

       _____

             we   share out   (direct
                   (verb)   object)

The tones of the direct objects in (53) have been changed from the tonal form indicated in parentheses to the indicated tones by the process of Object Tone Mutation discussed immediately above. From the data in (53) it is clear that on any account there must be a rule affecting the tone of the verb stem, a rule that is conditioned by the tone of the direct object. Our system so far makes it clear that, all other things being equal, the Mid tone seen in (b) and (d) is the expected tone; all is not equal, however, in (a) and (c), where the presence of an initial High tone in the direct object has caused deletion of the floating H tone before rule (32), Low-High Simplification could cause the Low-High sequence to become a Mid. In short, we have a rule of "Floating H Deletion" (54).

(54)   V   Floating H Deletion
           /\
        T • H •# (#) H

        1   2   3→
        1   ∅   3

Floating H Deletion says that a contour tone followed by an H is simplified as shown above. Note, crucially, that such an H-Deletion does not occur in examples such as (40), where the H-tone is associated to its "own" syllable(s).

In summary, we have seen that the tonal formula for this tone is B-tone for the verb stem, and a floating H-tone suffix. A rule of Floating H Deletion has been postulated to deal with more complex tonal alternations, and the Object Tone Mutation phenomena have been introduced.

8.   I Relative/Subject
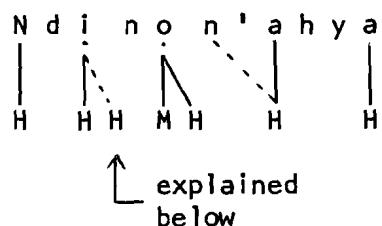
This form is used for relative clauses where the head of the clause is the subject of the embedded sentence, as in "the woman who won the prize," but not as in "the man I saw." The latter would be treated as a Relative/non-subject form.

The tone of the verb stem in this form is invariable M,[7] while the suffixes are H, caused by a floating H-tone suffix, just as in the I Subordinate form considered immediately above. As we would now expect, given the formulation of Object Tone Mutation in (49), direct objects undergo Object Tone Mutation-- the consequence, as we have seen, following Williams, of the floating H-tone suffix. This can be seen in (56); (55) is intransitive. Some examples:

(55) Low-toned Verb: Ńdị̀   nọ̀   ń'áhyá

           people be at market
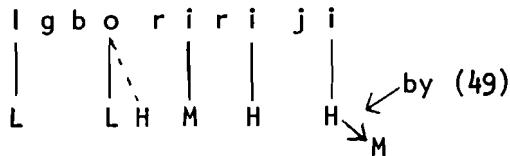
      "the people who are at the market"

```
N d i   n o   n ' a h y a
|   |.    |.     \.  \| |
H   H H   M H     H     H
        ↑ explained
          below
```

(56) High-toned Verb: Ìgbǒ   rirí   jí

           Igbo   eat   yam

      "the Igbo people who eat yams..."

```
  I g b o   r i r i   j i
  |     |\    |    |       |
  |     | \   |    |       |    ← by (49)
  L     L H   M    H      H ↘ ←
                            ↘ M
```

I shall not pursue the matter of the Floating H-tone suffix; its effects on the Class I suffixes and its triggering of Object Tone Mutation are sufficient evidence for its occurrence here. An example of Object Tone Mutation occurs in (56) above; mutation of <u>ahya</u> is regularly blocked in (55) by the preposition <u>na</u> (or <u>n</u>).

A new tonal phenomena arises in this form. The tone of the final syllable of the subject of the clause is raised, as we see above in (56), where underlyingly the subject is <u>Ìgbò</u>, which becomes <u>Ìgbǒ</u> in this form. In brief, the changes undergone are as follows:

(57) <u>Underlyingly</u>   <u>Subject Mutated Form</u>

jí (H)   jí (H)   (no change)

ánú (HH)   ánú (HH)   (no change)

àkwhá (LH)   àkwhá (LH) (no change)

ònù (LL)   ònǔ (L LH) or ònú (LH)

ázù (HL)   ázǔ (HM)

We see clearly what is going on if we look at the change ònù > ònǔ; an H-tone is being added on. As for the L ⌢LH/L H alternation, we have already posited (32) LH-Simplification, which will change L ⌢LH to LM, which is itself indistinguishable from LH. This same rule explains the H L → H ⌢LH → H M change (the last line of (57)).[8]

In fact, we are rather pleased to see surface alternations justifying rule (39). That the forms ending in H don't change tonally follows, of course, if we accept this analysis of tone-raising as resulting from a leftward docking of a floating H-tone, since V is indistinguishable from V.

```
   /\                           |
  H  H                          H
```

If we step back from the analysis a moment, we may note that tonal alternations like those displayed in (57) are precisely the

sort predicted by a theory with floating tones: unidirectional
(raising or lowering) tonal changes on the left or right side
of a word, sometimes realized by contour tones themselves. In
addition, we would expect a systematically motivated syntactic or
morphonological orginal for the floating tone. On this last point.
see Section 13 below.

In summary, then, the I Relative/Subject form is:

Floating	Stem	Suffixal
H	M	H

The precise formulation of the rule that docks the floating H tone
here onto the subject will await the discussion below of the under-
lying position of this H-tone. It is, of course, the floating
tone discussed in Section 5 of Chapter One above.

9. I Relative/Non-Subject

Roughly speaking, this form is used for relative clauses
where the head of the relative clause is not also the subject
of the relative clause--"roughly," because we find cases where
there is no particular head at all; rather it is used as an ad-
verbial clause. We shall consider these below. There are no
cases, however, where the I Relative/Non-Subject is used as a
relative clause where the head of the relative is the subject
of the relative clause; in all such cases, the Relative A form
is used. Conversely, I know of no cases where the Relative A
is used as an adverbial.

The tone of the verb stem here in the I Relative/Non-
Subject form is always L. The suffixes follow on L tone. too,
indicating that there is no suffixal tone in this tense. As
we expect, then, there is no object tone mutation in this tense,
either.

Just as we found Subject Tone Mutation in the I Relative/
Subject form, we find such a tone mutation here in the I Relative
B form. Some examples:

(58) Underlying   Mutated Form

H jí   →   H jí

L H ùbhé   →   L H ùbhé

H H éghú   →   H H éghú

$$H\ L \left[\begin{array}{lcl} H\ H\ \acute{u}khw\grave{a} & \rightarrow & H\ M\ \acute{u}khwa \\ \\ H\ L\ L\acute{a}dh\grave{a} & \rightarrow & H\ LH\ L\acute{a}dh\check{a}\ (personal\ name) \end{array} \right.$$

$$L\ L\ \grave{o}v\grave{u} \quad \rightarrow \quad L\ H\ (=LM)\ \grave{o}v\acute{u}$$

In the examples in (41), the form with surface alternations was LL. Here, in the texts, we find occasional alternations in the HL form, as indicated above. Considering the sparsity of data, I have no way to hypothesize concerning the optionality or obligatoriness of LH Simplification in this context. Again, we attribute this to a floating H tone docking onto the subjet as in the I Relative/Subject; its precise history will become clear in Section 13.

There is a new tonal rule that occurs in the I Relative/Non-Subject form that we have not considered yet. Recall that the I Relative/Non-Subject form is used when a relative clause follows the head NP which is not the subject of the relative clause. Since Igbo is an SVO language, this juxtaposes the head of the clause with the subject of the clause like this:

(59) . . . Head NP/Subject NP Verb Etc.

Relative Clause
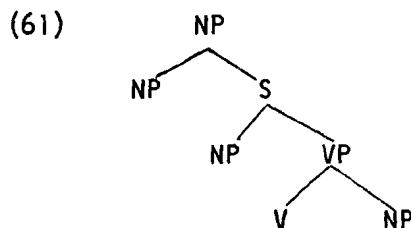
In the following discussion, I shall refer to the first and second NP's in (59) as the Head NP and the subject NP respectively. From the point of view of either useful prepositions or case markings, Igbo is quite impoverished; in place of such things, Igbo uses word order and tonal modifications. In order to help make clear the grammatical relation between the Head NP and the Subject NP--to wit, none--there is a rule, henceforth "Disjoint Tonal Sandhi"--which operates on the head NP's tone. The point of the rule, again, is to point out that the two nouns are not closely related, as distinguished, in particular, from the two nouns in the genitive relation; on this construction, see section 15 below.

The change that occurs is the following:

(60) Disjoint Tonal Sandhi: If the subject NP begins with L, then the Head NP mutates:

From	To
H	L
HH	LL
LH	LL
HL	HL
LL	LL

The syntactic structure to which this rule is applying is essentially (61).

(61)

We shall formulate the Disjoint Tonal Sandhi Rule, but this is tentative, and based on some applications we have not considered yet.

(62)   Disjoint Tonal Sanshi

$$H \rightarrow L \; / \; -\!\!- \; X \; \# \; [\; L \quad \text{Condition: X contains}$$
$$S$$
$$\text{no } \# \text{ and no L}$$

This produces the correct results according to (60), and applies to a structure such as (61), but in no other structures encountered so far. This process of Disjoint Tonal Sandhi, however, plays a rather marginal role in the entire grammar of Igbo; in this way it is quite different from the other tonal rules, which form a very tight and inter-connected system.

In summary, then, the tone form of the I Relative/Non-Subject is:[9]

Floating	Stem	Suffixal
H	L	Ø

10.   Summary

We have considered four forms to this point: Green and Igwe's Affirmative I Root, I Subordinate, I Relatives/Non-Subject and Subject. We shall next consider the corresponding II Incomplete forms, which will give us a clearer idea of the structure

of the phrases we are investigating.

(63)

Form	Floating	Stem	Affix	Special Rules
I Root	--	L	--	I Root Flop
I Subord	--	B	--	
I Rel/Subject	H	M	H	
I Rel/Non-Subj	H	L	--	*

*This form is subject to the Disjoint Tonal Sandhi rule, but this is not due to a special marking of any sort, but because the I Relative/Non-Subject satisfies the structural description of the rule.
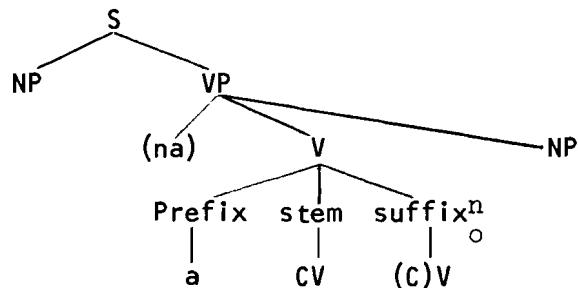
Note that the only specially marked rule is one which applies only in "root" sentences, in Emonds' sense.[10]

11.   The II Incomplete Forms

The form of all the affirmative II forms is similar to the I forms except that there is a verbal prefix /a/ in the II form. (/a/ becomes /e/ if the verb stem is in the "tense" harmony ckass.) In addition, there may appear before the /a/ a relative marker /na/ in the II relative forms.

Thus, schematically:

(64)   II Incomplete -- Schema

When one of the short forms of the singular pronouns--the cliticizing pronouns--is subject, the a/e prefix is suppressing-- unless the /na/ relative marker intervenes, in which case the a/e remains. On this important process, more below.

12.   The II Incomplete Root

Initiating use: This form, the II Root, is divided by Green and Igwe into (i) the initiating, and (ii) the non-initiating, forms; by this is meant (i) discourse-initial use, or (ii) elsewhere.

In the initiating use, the clitic pronouns are used. The following chart indicates the morphological shape of the pronouns throughout the forms. The cliticizing forms are given first, the strong forms second.

(65)   Pronouns

Person	Singular	Plural
1	m/mʉ	-/anyi
2	i/gi	-/unu
3	o/ya	-/ña
indef	a/-	-/-

Let us consider the tonal characteristics of this form when there is a non-cliticizing subject, thus not deleting the a/e prefix.

(66)   High-toned verb cí 'carry'

Prefix	Stem	Suffix
a	cị	a

Égbé àcịá àkhwá   Egbe carried eggs

Àdhá àcịá àkhwá   Adha carried eggs

Ézè àcịá ánḍ   Eẓe carried meat

Ùghò àcịá àkpà   Ugho carried bags

(67)   Low-toned verb zà 'sweep'
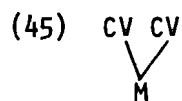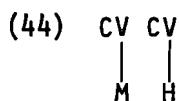
Prefix	Stem	Suffix
a	za	a

Égbé àzàá ámà   Egbe swept the street

Àdhá àzàá ámà   Adha swept the street

Úcè ázàá ámà   Uce swept the street

Ùghò ázàá ámà   Ugho swept the street

Apparently the final tone of the subject affects the tone of the prefix--here, a--for it changes, depending on whether the subject ends in an H or L tone. We can summarize these tones in (55). I shall write "H/M" where we cannot determine which it is because of an L tone immediately preceding.
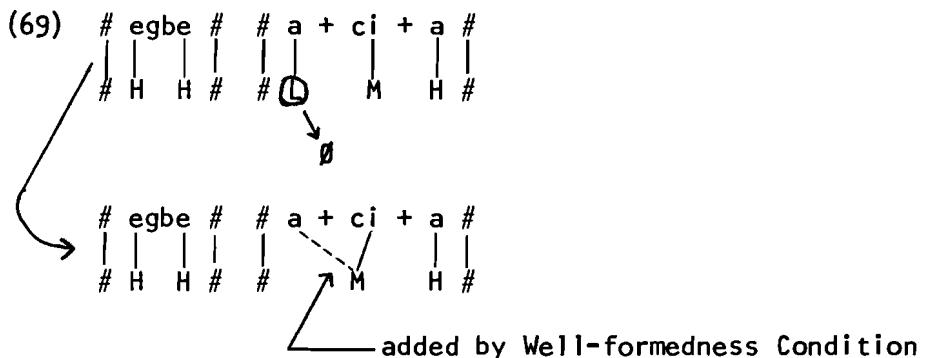
(68) Subject ends in	H-verb Prefix Stem		L-verb Prefix Stem	
(a) H	M	H	L	L
(b) L	L	H/M	H/M	L

Note that the H following the M in (68a) (e,g., as in "Égbé àcíá...") is actually on the same pitch as the prefix which is indicated by M; in general, as we remarked above (44) and (45) are phonetically indistinguishable.

(44) CV CV
      | |
      M H

(45) CV CV
       \/
       M

Indeed, we see that the prefix and the stem are on the same pitch if and only if the subject NP ends in H (the cases in (68a)). Furthermore, this tone is found on the **stem** in the other case (68b), when the subject ends in L.

The autosegmental analysis of this is straightforward. A final H on the subject NP deletes the tone of the prefix, which automatically reassociated the stem tone to the prefix. For example:

(69) # egbe # # a + ci + a #
       | | | | | | | | |
       # H H #   #Ⓛ   M H #
                   ↘
                    Ø

      # egbe # # a + ci + a #
       | | | | | \ / | |
       # H H #   # ↗M   H #
                 /
        └────── added by Well-formedness Condition
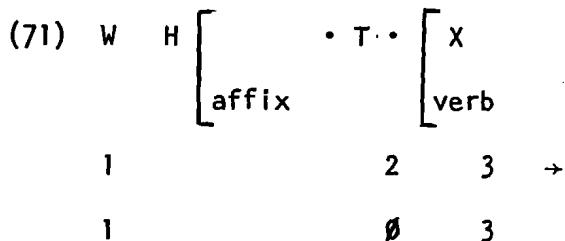
The underlying tonal form, then, is what is given in (68b). The tone of the stem is L with an L-verb; it is M with a non-L verb. We shall continue with the convention given in Chapter One, and abbreviate this dependency by "B*."[11]

The verb stem has tone B*, then, in the II Main form. The prefix for non-Low verbs, as we can see in (68b) is L; the prefix tone for Low-verbs is either M or H--we cannot tell which. I shall arbitrarily pick H, but nothing rides on this choice. Thus the prefix has a dissimilatory tone on it with respect to the stem; I shall denote this as "B⁻."

(70) II Main Tone

Prefix	Stem
B⁻	B*

Let us formulate, then, the prefix-deletion rule illustrated in (69).

(71) W H $\begin{bmatrix} \\ \text{affix} \end{bmatrix}$ · T · · $\begin{bmatrix} X \\ \text{verb} \end{bmatrix}$

1   2   3   →

1   ∅   3

Tonal Behavior of Pronominal Subjects in II Root Form

Cliticizing pronouns behave in the II Root as illustrated (72). They are uniformly H in tone, be the verb H or L. Note that the a/e prefix is suppressed.

(72a) cí 'carry' H-toned verb

í cʲá áṇ̀   "you have carried meat"

ọ́ cʲá áṇ̀   "he has carried meat"

á cʲá áṇ̀   "meat has been carried..." or
                     "one has carried meat"

(72b) zà 'sweep' L-toned verb

í zàá ámà   "you have swept the street"

ọ́ zàá ámà   "he has swept the street"

á zàá ámà   "one has swept the street"

The H-tone on the suffixes, and its corollary object tone mutation, of course, remains constant from (66-67) above.

The disappearance of the a-prefix in the presence of the cliticizing pronoun is due, as we have already said, to the fact that the cliticization process <u>substitutes</u> the pronoun for the a-prefix. The rule was introduced in Chapter One (66), reproduced here.

(73) <u>Subject Cliticization</u>

$$\begin{bmatrix} \text{Pronoun} \\ \text{weak} \end{bmatrix} \begin{bmatrix} \text{Verb} \\ \text{prefix} \end{bmatrix}$$

$$\begin{array}{cc} 1 & 2 \rightarrow \\ \emptyset & 1 \end{array}$$

The reader is urged to reread Section Five of Chapter One here, because the floating tone process sketched there pulls together a larger number of the processes that are being introduced one by one in the present chapter.[12]

By way of review, consider the tone of the underlined verb in the following paradigm (74). The superficially complex tonal alternations reveal themselves as quite natural, given what we have seen so far. We shall show how these follow from the rules we have posited.

(74) (i) Yá jíkhéré, ányì <u>áfừ</u> ámà (from <u>ámá</u> 'street')

he ready   we go out street

'if he is ready, we'll go out to the street'

(ii) àhú khòwé, únù <u>áfừ</u> èzí (from <u>èzí</u> 'outside')

body swells you go outside

'when it gets hot, you will go outside'

(iii) Á bhá yà, yá <u>àbhà</u> ódè

one grips him he grips cocoyam

'when one grips him, he grips the cocoyam'
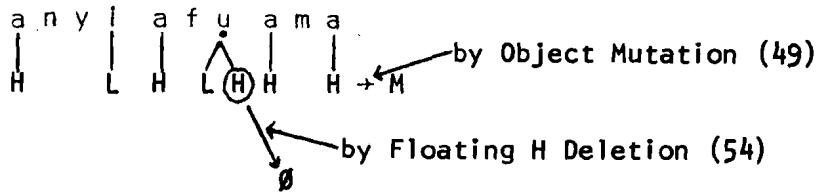
(iv) Á bhá yà, yá <u>àbhá</u> àlà

one grips him he grips ground
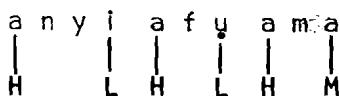
'when one grips him, he grips the ground'

In all four cases, the underlined verb--prefix plus stem--is in the II Main form, and in all four cases is a Low-toned verb. But due to different environments, each has a different tonal pattern. These are derived as follows:

(74i)

```
      a n y i   a f u   a m a
      |     |   | /\    | |
      H     L H L(H)H   H →M      ──by Object Mutation (49)
                    |
                    ↓           ──by Floating H Deletion (54)
                    ∅
```

yields:

```
      a n y i   a f u   a m a
      |     |   | |     | |
      H     L H L H     M
```
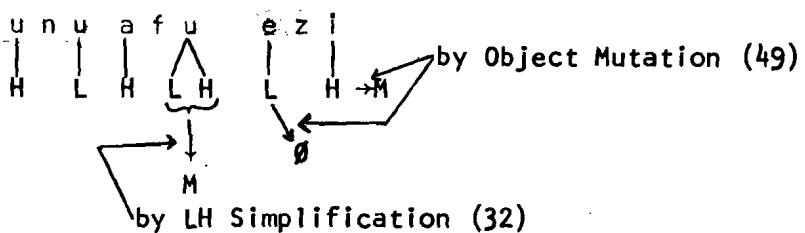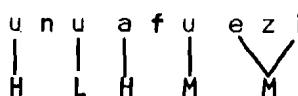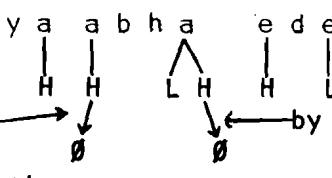
(74ii)
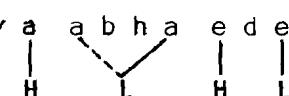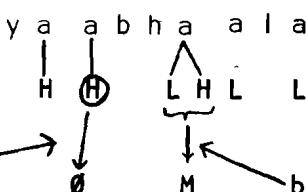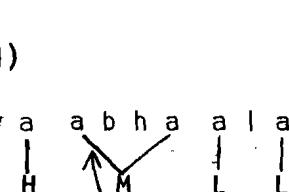
```
      u n u   a f u     e z i
      | |     | /\      |   |
      H L   H L H       L   H →M    ──by Object Mutation (49)
              \_/        |
               |         ↓
               ↓         ∅
               M
            ──by LH Simplification (32)
```

yields:

```
      u n u   a f u   e z i
      | |     | |     | \ /
      H L     H H     M   M
```

(74iii)

```
      y a   a b h a   e d e
      | |   |   /\    | |
      H H   L H H     H L      ──by Floating H Deletion (54)
          ↓       ↓
          ∅       ∅
        ──by II Root
          Rule (105)
```

yields:

```
      y a   a b h a   e d e
      | |   |  \ /    | |
      H     L   H     H L
```

(74iv)

```
      y a   a b h a   a l a
      | |   |  /\     | |
      H (H) L H L     L
            ↓   \_/
            ∅    |
                 ↓
                 M
            ──by LH Simplification (32)
        ──by II Root
          Rule (71)
```

yields:

```
      y a   a b h a   a l a
      | |   |  \ /    | |
      H     M   M     L L
        ──added by Well-formedness Condition
```

These derivations illustrate our rules:

(49) Object Tone Mutation
(54) Floating H Deletion
(32) LH Simplification
(71) II Main Rule

We also have:

(25) I Root Flop
(62) Disjoint Tonal Sandhi
(73) Subject Cliticization

13.   The II Subordinate and the II Relative/Non-Subject

The reader will recall that we observed above that there was some overlap between the uses of the I Relative/Non-Subject and the I Subordinate: sometimes the I Relative/Non-Subject was used adverbially rather than as a relative clause. Nonetheless, a glance at chart (63) indicates a substantial difference between the two forms.

The parallel II forms--that is, the II Subordinate and the II Relative/Non-Subject--bear an even closer affinity to each other. In fact, we shall conclude that they are underlyingly identical from a tonological point of view.

The verb stem and its Class I (toneless) suffixes are L in the II Relative B, just as was seen in the I Subordinate form; the same is true of the II Subordinate. From this we conclude directly, as we did for the I subordinate, that the stem tone in these two forms is L, and that there is no suffixal floating tone; the Low tone that appears on the Class I suffixes is the stem's L tone spreading rightward, as in the I Main form. The Direct object NP does not undergo Object Tone Mutation, as we would expect, since there is no suffixal H tone to trigger it.

Furthermore, in both the II Subordinate and the II Relative/Non-Subject forms the non-clitic subjects undergo the tone-raising on their right-end side as we have observed,. for example, in. (57), which we attributed to a floating H-tone to the left of the verb stem that docked leftward. Thus on the three major tonological dimensions we have defined for Igbo verb forms so far, the II Relative B is indistinguishable from the II Subordinate: both have the floating H tone to the left of the verb stem, and both have an L-tone associated with the stem (and no suffixal toneme). Schematically:

(75)

Pre-Floating	Stem	Suffix
H	L	$\emptyset$

Thus along the three tonological dimensions we have defined for
Igbo verb forms so far, the two forms are indistinguishable.
Furthermore, when the II Relative/Non-Subject is used as an
adverbial, it can appear with the suffix mò -which appears else-
where in Igbo only in the II Subordinate form.

Why distinguish the II Subordinate and the II Relative/
Non-Subjct at all, then? Why give it two names, for that matter,
if it is really only one form?

There are three reasons, really: first and most simply,
because we are following Green and Igwe's classificatory system,
which deals with these two forms quite separately. Second,
because these two forms, as I have indicated, play quite
separate roles; a relative form has one function, an adverbial
Subordinate clause another (although, admittedly, there is
some overlap). The distinction between these two functions is
kept separate in the I Forms, as we saw in chart (63). So it
is just an accident, really, that the two forms have the same
underlying melody. In short, the forms themselves are different;
they simply have the same underlying melodies.

But the real reason for pursuing the differences between
the II Subordinate and the II Relative/Non-Subject is this:
our theory has so far led us quite directly to a specific way of
analyzing the tonal system of these forms; it forces the con-
clusion that the melodies of these forms are as in (75). When
we look now for differences between the two forms--and there
are some--we are hamstrung by the lack of distinctions permitted
by the theory; and so the differences that appear must be "super-
ficial" differences arising out of something entirely different.
Indeed, this is the case, and it has already been foreshadowed
in the discussion of the floating tone in Chapter One.

Consider the paradigms below, which illustrate the tonal
forms in the II Subordinate and the II Relative/Non-Subject.
First are given the forms with non-clitic subjects, and these
forms are essentially identical. Note the presence of the low
toned /a/ or /e/ prefix. Below are the forms with pronominal
subjects, and there we find a sharp contrast: the II Subordinate
has a Falling tone on the cliticizing subjects, while the II
Relative/Non-Subject has Low-tone.

(76) II Subordinate

Subject:

(i) /ághú̞/	ághú̞ è̞gbùò éghú	'lest the leopard kill the goat...'	
	leopard kill goat		
(ii) /ò̞ké̞/	ò̞ké̞ àtà̞à ákhú̞	'lest the rat eat the palm kernels'	
	rat eat palm kernels		
(iii) /úz̃è̞/	úz̃è̞ àtàà yá	'lest the squirrel eat them'	
	squirrel eat them		
(iv) /è̞ŋwò/	è̞ŋwó àtàà yá	'lest the monkey eat them'	
	monkey eat them		

(77) I Relative/Non-Subject

Subject:

(i) /éghú/	Ùbhú éghú àtàà̞là...	'The pears eaten by the goats'	
	pears goat eat		
(ii) /Àdhá/	Ùwé Àdhá è̞tlè̞lè̞...	'The clothes Adha wore...'	
	clothes Adha wear		
(iii) /ézè̞/	Ùwé ézè̞ àzù̞ò̞là...	'The clothes the chief has bought...'	
	clothes chief buy		
(iv) /Àlà/	Ò̞kè̞ Àlá è̞rlè̞lè̞...	'The shares Ala has taken...'	
	shares Ala take		

(78) II Subordinate

Pronoun Subjects:

... m̀ gฮ̀uɔ̀ éghú	lest I kill the goat
... ĩ " "	lest you kill the goat
... ŏ " "	lest she kills the goat
... ę̀ " "	lest the goat be killed
ányī̀ ègbùɔ̀ éghú	lest we kill the goat
únú ègbùɔ̀ éghú	lest you (pl) kill the goat
ńá ègbùɔ̀ éghú	lest they kill the goat

(79) II Relative/Non-Subject

(a) Ṃgbé m̀ zụ̀là ánụ́   Whenever I buy meat

Time I buy meat

___

(b) Ṃgbé ĩ zùlà ánụ́   Whenever you (sg) buy meat

___

(c) Ṃgbé ǫ zùlà ánụ́   Whenever he buys meat

___

(d) Ṃgbé à zùlà ánụ́   Whenever one buys meat

___
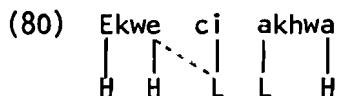
(e) Ṃgbé ányī̀ àzùlà ánụ̣́   Whenever we buy meat

___

(f) Ṃgbé únú àzùlà ánụ̣́   Whenever you (pl) buy meat

___

(g) Ṃgbé ńá àzùlà ánụ̣́   Whenever they buy meat

___

   We are thus faced now with the need to explain the
tone on the cliticizing pronouns in (78)--the II Subordinate
form. In doing this, we will come to understand the origin of
the floating H tone to the left of the verb stem that we have
postulated to account for the raising of the final tone of the
subject Noun phrase in several forms (see (68) and (72), (73)
of Chapter One, for example).

The Falling tone on the cliticizing pronouns in (78) is distinctly different from the only other Falling tone we have encountered so far--that is, the Falling tone created by the Flop rule in the I Main form, as illustrated in (80).

(80)   Ekwe   ci   akhwa

H   H   L   L   H
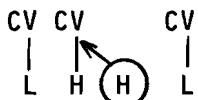
The dotted line represents the association line added by the I Main Flop rule.

In the case of the Falling tone created by the I Main Flop rule, the cliticizing pronouns were precisely the forms that did not take on a Falling tone. In the II Subordinate form, however, it is rather the reverse: here, only cliticizing pronouns take on a Falling tone. And the origins of these two falling tones are quite different: similarities in appearance may be deceiving.
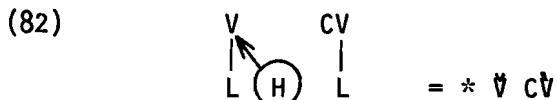
We shall show that the Falling-toned clitic pronouns in the II Subordinate are just the Low-toned clitic pronouns of the II Relative/Non-Subject (see (79a-d)) plus a Floating H-tone docked onto their left-side. This H-tone is precisely the same H tone that causes a raising of tone on the subject NP. This Floating H-tone begins, in a certain sense, in between the subject and the verb, and may dock either leftward or rightward.

Thus we will be arguing that the raising of the subject NP's tone operates as in (81).

(81)   Subject   Verb

CV CV   CV

L   H (H)   L
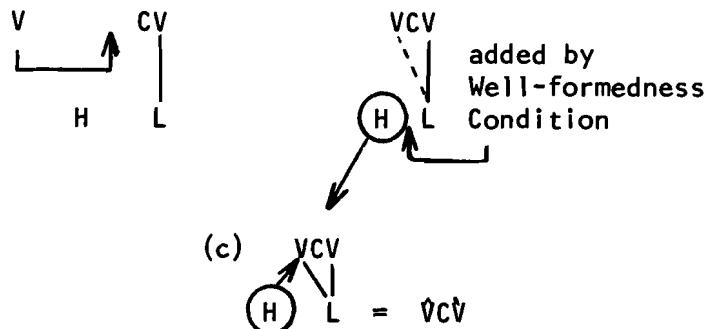
But if a pronoun were just like any other noun phrase, we would then expect a Rising tone on it; as in (82).

(82)   V   CV

L (H)   L   = * V̌ CV̌

In (82), we have not taken into account the cliticization of the pronoun; we shall argue that cliticization brings the clitic pronoun into the verb phrase, as in (83a); the docking of the floating H tone, now, if it be rightward, causes the tone of the clitic pronoun to be a Falling tone, as in (83c).

(83) (a) Subject Verb   (b)

We now precede to argue these two points.

In the discussion of the raising of the tone of the subject NP in the I Relative/Subject (following illustration (57)), we argued that, given our conception of autosegments, and thus floating tones, such a tone mutation must come from a floating tone, docking from the right. In the II Relative/ Non-Subject form, we find confirmation of this as direct as one could hope to find.

In the II Relative forms, it is possible--indeed, pre-ferable--to have an /na/ affix in front of the a-prefix. Thus, for the II Relative/Non-Subject both (84a) and (84b) are possible.

(84a) áñyá éghú náàlà écì   "The market the goat
                                should have left
      market goat leave yes-   yesterday"
                          terday

(84b) áñyá éghú àlà écì   (same meaning)

Na, when present, is always H-toned; and, note well, when na is present, the Subject Raising process does not occur. This was not apparent in (84); see however (85), where the subject /ósò/ becomes /ósò/.
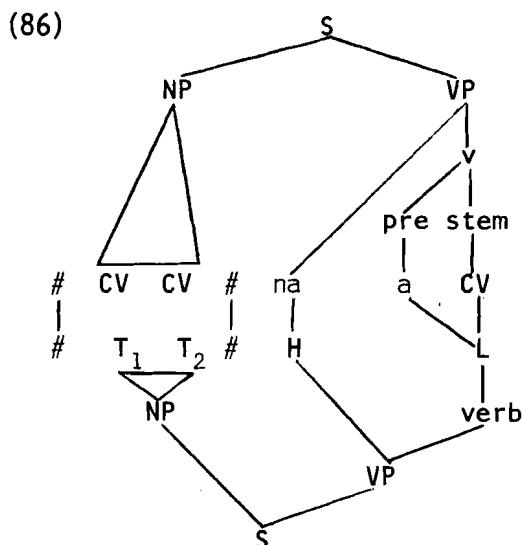
(85a) áñyá ósò náàlà écì "The market the
                              pepper should have
      market pepper leave yesterday left yesterday"

(85b) áñyá ósò àlà écì   (same meaning)

The Subject tone raising we have observed all along appears now to be just a particular example of the more general "stability" phenomenon discussed in Section 3 of Chapter One; when the na deletes, its High tone shifts leftward onto the subject. When the na does not delete, as in (85a), of course, the subject displays its isolation tones.

More generally, this H tone may appear in forms where there is no possibility of the na occurring--as in the I Relative forms and the II Subordinate. The correct way to phrase the generalization we saw above linking the na and the H tone is to say that lexical insertion for stricly tonological material and strictly syllabic (non-tonological) material occurs on parallel syntactic phrase-markers, and if corresponding elements occur, they will be associated. Thus, for a sentence like (85a), we get a tree like (86). Note that the same phrase structure rules govern both the tonological and the syllabic structures.

(86)

However, when the na does not appear--for what ever reason--we have a "floating tone" on the tonological tier: the H tone. It will then ultimately dock onto some vowel.

We will now see why cliticization of the subject pronouns should be viewed as a movement rule, as in (83) (or better--(67)--of Chapter One). This analysis has already, we note, made certain correct predictions regarding the I Main Flop rule. This rule we have formalized as (73), repeated here.

(73) Subject Cliticization

$$\begin{bmatrix} \text{Prounoun} \\ \text{Weak} \end{bmatrix} \quad \begin{bmatrix} \text{Verb} \\ \text{Prefix} \end{bmatrix}$$

$$\begin{matrix} 1 & 2 \rightarrow \\ \emptyset & 1 \end{matrix}$$

Adopting Emonds' framework, this rule must be a local (or minor movement)rule, since it is neither a root rule nor a structure-preserving rule. This implies that there can be no material intervening between the two factors in the rule. If, therefore,

the na should be present, as in (86), rule (73) cannot apply,
and further‚ore, the rule cannot (according to Emonds' principle)
be generalized to (87), 'for example:

(87) $\begin{bmatrix} \text{Pronoun} \\ \text{Weak} \end{bmatrix}$ (na) $\begin{bmatrix} \text{Verb} \\ \text{Prefix} \end{bmatrix}$ $\underline{\begin{matrix} \text{Impossible Variant} \\ \text{of Subject} \\ \text{Cliticization} \end{matrix}}$

$\begin{matrix} 1 & 2 & 3\rightarrow \\ \emptyset & 2 & 1 \end{matrix}$
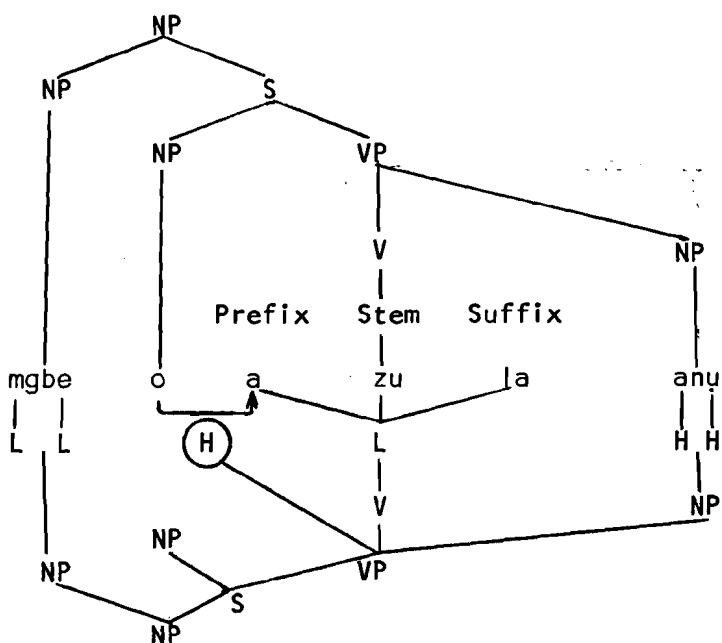
In fact, this is correct; when the na is present, subject pronouns
do not cliticize, and so the Verb prefix /a/ remains. Compare
(88a) and (88b).

(88a) mgbe o a zula anu ———— Rule (73)

[mgbe o̱zu̱la anu]

(88b) mgbe o na a zula anu

[mgbe o na a zula anu] ·Prefix does not delete

Parenthetically, this provides us with a whole new argument
for autosegmental theory, for if the presence of na blocks the
cliticization rule (70), the presence of the H-tone between the
subject and the verb does not. Thus (88a) proceeds as in (89),
regardless of the presence of the floating H-tone; therefore,
the H tone is not in the string to which rule (70) applies, and
must be on another (autosegmental) tier--but this is exactly what
was predicted.

(89)
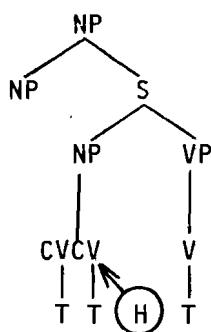
Pursuing this a bit further, we see that rule (73) is a
rule applying to the syllabic (upper) autosegmantal tier, and
the moved pronoun subject takes up the tone of the /a/ prefix
that it replaced, as indicated above in (89). The unmoved
clitic form, as in (88b), is H in tone, but whether this is
from the H-tone or some other source, we cannot presently say.

The rule that docks that floating H tone in (89) and
similar structures is an autosegmental rule; the rule that
cliticizes the subject pronoun (70), is a syntactic rule,
since it falls within the domain of Emonds' hypothesis.[13]
Given our general assumptions about rule-application ordering
and the derivation, the cliticization rule must always apply
first. We assume, as discussed in Section 4 above, that
syntactic rules--minimally, cyclic syntactic rules--precede
insertion of phonological material, and again, preferably
before insertion of all lexical material. Therefore, an
autosegmental rule must follow a syntactic rule in application.
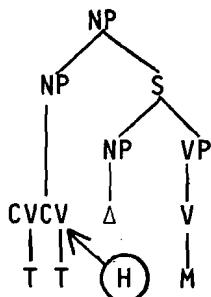
We will close this section with a formulation of the rule
that "docks" the floating H tone. The point of this section--
as of the previous section--was a positive and optimistic
answer to the question posed in the second paragraph of this
chapter: we see that inclusion of autosegm ntal analysis has
led to a simple conception of the underlying tonal form in
Igbo, with a small number of elementary rules. Furthermore,
the complementarity between such processes as Subject Tone
Raising, high-toned na and falling-toned subject clitics is
straightforwardly explained.

In the I Relatives, the II Subordinate and the II Relative/
Non-Subject, we have seen the H-tone dock leftward in the
following circumstances.
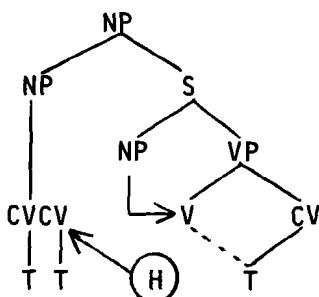
(90)   II Sublordinate, I, II Relative/Non-Subject

(91) I Relative/Subject

```
              NP
            /    \
          NP      S
          |      / \
          |    NP   VP
          |    |    |
        CVCV   Δ    V
         | |\       |
         T T (H)    M
```

(90') II Subordinate, II Relative/Non-Subject

```
              NP
            /    \
          NP      S
          |      / \
          |    NP   VP
          |    |    |
        CVCV  L→V   CV
         | |\       |
         T T (H)    T
```
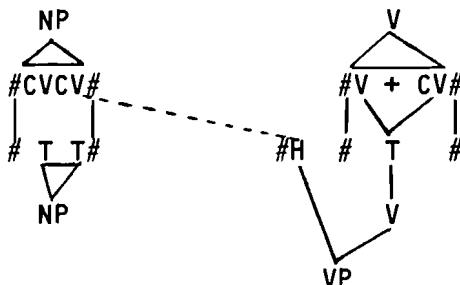
(if and only if the subject is a cliticizing pronoun)

Our main problem in formulating the rule that docks the floating H tone (circled in (90), following Mary Clark's notation) comes not from the autosegmental approach, but something quite different. Present generative theory is quite silent about the way in which syntactic boundaries may be placed in phonological rules, and more importantly, given a certain small data set which indicates that a certain phonological process is syntactically conditioned in part, which aspects of the syntactic structure will be most highly preferred by the language learner to condition the rule. In short, we the linguists have too many choices in viewing a structure as in (90)-(91) when we want to write the syntactic conditions for the docking rule. The question is what conditions strike the language learner as simplest--what will he or she guess the conditions to be on the basis of (90)-(91). And this question is quite independent of any autosegmental proposal, for the same issue could have arisen in a purely segmental phenomenon.

This much said, let us look at the relevant parts of the structures in (90)-(91) in more detail: (see (92)). Here we make explicit references to the convention that corresponding word-boundaries are associated to each other, above and beyond all tone/vowel associations. There is a word-boundary to the left of the floating H-tone that has no corresponding word-boundary on the syllabic tier, and thus remains unassociated.
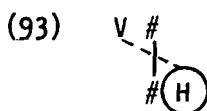
(92)
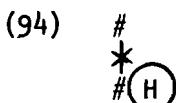
We may note that there is no inconsistency in having an operation inset an association line that crosses a previously present association line. The Well-formedness Condition prohibits such a structure, and the interpretation we have given to the Condition will "change the representation minimally by addition or deletion of association lines so as to meet the Condition maximally" (p. 48 above). Thus the old association line will delete in the process of putting the new (dotted) line in (92). The same type of operation occurred in the I Root rule (above, Section 6).

The absence of a corresponding syllable (na) for the H tone in (92) is crucial for the operation of the docking of the H-tone. Our present notation appears to have an arbitrary asymmetry, in that the _presence_ of association lines is easy to write in as part of the structural description or change of a rule; the corresponding absence is, so far, impossible. This asymmetry, however, is unprincipled, as far as I can see, and we may rectify this by proposing that a notation (i) for deleting association lines as a structural change, corresponding to the dotted line which indicates insertion of association line; we shall use an 'x' in an association line for this, and (ii) a notation indicating that a tone or syllable is associated to nothing at all--circling the segment.[14]
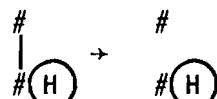
With these conventions, two solutions are unfortunately available. We could write the rule as (93); we could also write it as (94). Note that the circle around the H means it is a floating tone--more precisely, associated with no syllable at that point in the derivation.

(93)


DOCKING RULE

(94)
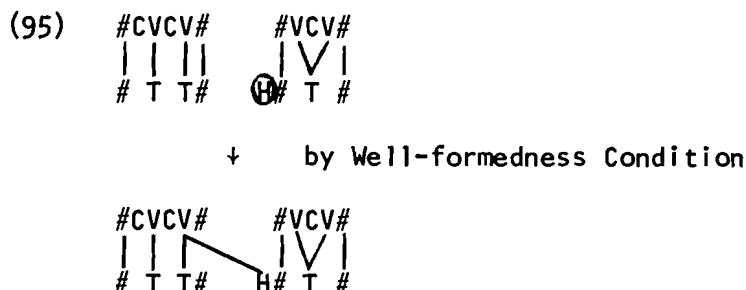
where this means

The effect of (94) is perhaps not immediately obvious; but by deleting the association line in (92), we get a structure as in

(95), which, by the Well-formedness Condition, leads to the association of the floating tone on the final syllable of the subject NP.

(95)   #CVCV#   #VCV#
         | | || | V |
         # T T#   # T #

              ↓   by Well-formedness Condition

         #CVCV#   #VCV#
         | | ⌐ | V |
         # T T#   H# T #

Rather arbitrarily, I shall choose (94) rather than (93) as the appropriate formulation of the docking rule. It must be emphasized that detailed investigation of other tonal systems is the only way to determine whether this choice was correct on the basis of the data.

This formulation of the rule, then, works appropriately for the cases laid out in (90)-(91). We have looked at (90) in detail in (92); the same rule operates in (91).[15]

As the docking rule is written in (94), it docks the floating tone onto whatever is to its left with no further conditions as to what is found there to the left. This is largely, but not entirely, true. We have seen examples as in (90), where the H docks onto the subject NP in the subordinate clause; we have seen examples, as in (91), where ft docks onto the head NP outside the relative clause. In (96), we see examples where the H tone docks onto a complementizer (ka, ma) to the left of the clause. There are no examples, however, of the H tone moving outside the clause and associating with an NP external to the clause that was not grammatically related to the clause.

(96) $ma o sI ya gara ya ozi$ = má ọsì yà gàrá yà ózí
 |⌐ \/ | | | | | |
         L (H) L L L H M H H

       comp she send him go him errand

   'lest she should send him on a mission'
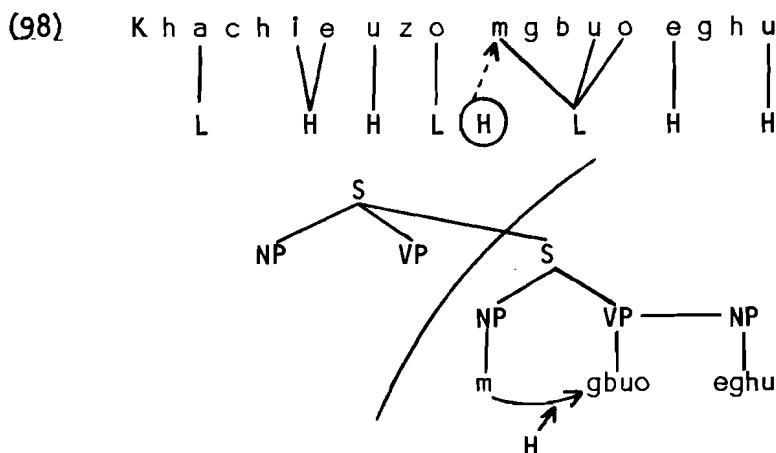
Thus we find no examples like (97); instead, we find the floating tone docking rightward, as in (78); this is illustrated in (89). Although this process is clearly related to the leftward docking rule (94), it is not identical to it. Another rule must be formulated that applies after (94), or when, due to the structural configuration (94) cannot apply. This second rule we shall call Rightward Docking (99). It is the mirror image of (94).

(97) *Khachie uzo mgbuo eghu

L H   H L (H) L   H H

shut   door I kill   goat

'Shut the door lest I kill the goat'

(98) K h a c h i e u z o m g b u o e g h u

L   H   H   L (H)   L   H   H

S
NP   VP   S
NP   VP — NP
m   gbuo   eghu
H

(99)   Rightward Docking

#
(H) #

And yet we have no formal principle why the floating H
tone docks rightward in a structure like that in (98), and
not leftward. There is apparently no intonational pause here,
unfortunately, which would account for this. I shall leave this
question open.

14.   The II Relative/Subject

This form is simple to understand, given the analysis so far.
The tone of the verb stem is Mid; there is no underlying tone for
the verbal prefix a, so it associates with the stem M throughout
the Well-formedness Condition.

This form is used only for relative clauses where the head
noun plays the role of subject of the relative clause. This
head noun undergoes Subject Tone Mutation, if the na prefix
does not occur, just as we have seen in the previous section.
Consider forms like (100) in the II Relative/Subject:

(100)   Eghu na ala ahya n'Ekhe
         H   H   H   M H   H   H   L   H

         goat   leave market on Ekhe
                                (a day)

'Goats which were
expected to leave
the market on Ekhe
day'

         Oso na ala ahya eci
         H L   H   M H H M H H

         Pepper   leave market yesterday

'Pepper which was
expected to leave the
market yesterday...'

         Oso ala ahya eci
         H LH M H   H   M   H H

              ↓
              M

(same meaning)

As I have noted in (100) above, there is a suffixal H tone
which raises the tone of any suffixes and triggers Object Tone
Mutation. In short, the tonal formula for this form is:

(101) | Floating | Prefix | Stem | Suffix |
         |----------|--------|------|--------|
         | H | Ø | M | H |

Summary

There is still more to be said about the verbal forms of
Igbo. We have looked at only half of them so far; we have not
yet approached the negative forms. These present slightly longer
tonal melodies, making considerably more use of the "B*" notation.
They do not, however, present much more information about the
functioning of an autosegmental system, and while the system has
a great deal of inherent interest, it is not germane to our
present questions regarding autosegmental rules.

We should not go without obseving that there is some order
in the chart (19). For example, there is a connection--somehow--
between the functioning of rule (72), the rule which deletes the
prefix tone in the II Root form, and the fact that in the subor-
cinate and relative forms, there is a floating H-tone and no
prefix tone. It is as if the floating H-tone is playing the same
role as the final H-tone of the subject NP that deletes the prefix
tone in the II Root form. Yet in the subordinate and relative
forms, the floating H is <u>always</u> there, and so the prefix tone is
<u>always</u> deleted. How much a regularity should be stated in a
grammar--indeed, whether it <u>should</u> be stated--I leave as open
questions.[16], [17]

(19) **repeated**

| | I Complete | | | |
	Floating /	Prefix	/Stem /	Suffix
Root	-	-	L	-
Subord.	-	-	B	H
Relative/				
Subject	H	-	M	H
Relative/				
non-subj.	H	-	L	-

| | II Incomplete | | | |
	Floating/	Prefix	/ Stem	/Suffix
Root	-	B	B*	H
Subord.	H	-	L	-
Relative/				
Subject	H	-	M	H
Relative/				
non-subj.	H	-	L	-

15. Nominal Compounding

We may pursue the empirical validity of the notational
convention suggested above for tones not associated with any
syllables at a certain point (we circled them at that point).
The significance of this move was that we would then expect
to find "being circled" to be a simple condition for a lang-
uage to place on a factor in a rule, and that certain rules
should become quite difficult to express if we do not have
access to this notional convention.

The process of compounding nouns in Igbo is quite
productive, and corresponds to a genitive or possessive con-
struction in other languages. Thus "John's book" is expressed
in Igbo as "book John," with the two words undergoing certain
regular tonal changes. Edwin Williams (1971) originally
pointed out that the tonal processes in this construction can
quite naturally be explained in terms of the Object Tone Mutation
and Subject Tone Raising processes we have already seen. These
together point toward an H tone intervening between the two
nouns of the compound construction, an H tone which plays much
the same role as the floating H tones we have seen in the verbal
system.

Chart (102) presents the relevant data, and (103) sum-
marizes the tonal forms. However, in the terms we have devel-
oped so far, these alternations can be summarized easily: the

first noun undergoes Subject Tone Raising just in case the second
noun does not begin with an H-tone.

(102)   First Noun

Second Noun	H	HH	LH	HL	LL
H jí	dí jí	ísí jí	àhú jí	ǫ́dhǔ jí	m̀kpǎ jí
HH éghú	ŋwá eghú	ísí éghú	àhú éghú	ǘlǫ̀ éghú	àbhà̰ éghú
LH òké	ŋwá òké	ísí òké	àhú òké	ǫ́dhu òké	àbhá òké
HL m̀bè	ŋwá m̀bè	ísí m̀bè	àhú m̀bè	ǘlǫ̀ m̀bè	àlà̰ m̀bè
LL èŋwò	ŋwá èŋwò	ísí èŋwò	àhú èŋẃo	ǫ́dhu̟ èŋwò	àbhá èŋwò

(103)
                              First Noun
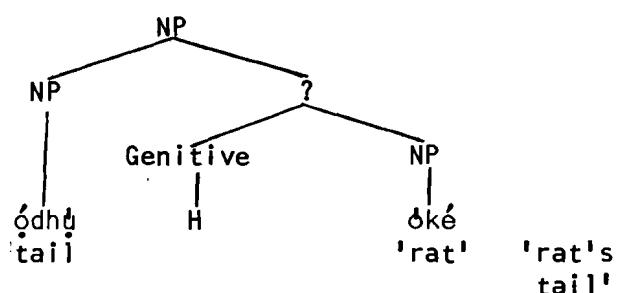
| Second Noun	H	HH	LH	HL	LL
H	H <u>M</u>	HH <u>M</u>	LH <u>M</u>	<u>H</u> L͡H <u>M</u>	L͡ L͡H <u>M</u>
HH	H <u>HM</u>	HH <u>HM</u>	LH <u>HM</u>	HL <u>HM</u>	LL <u>HM</u>
LH	H <u>MH</u>	HH <u>MH</u>	LH <u>MH</u>	<u>HM</u> <u>MH</u>	<u>LH</u> <u>MH</u>
HL	H HL	HH HL	LH HL	HL HL	LL HL
LL	H LL	HH LL	LH LL	<u>HM</u> LL	<u>LH</u> LL
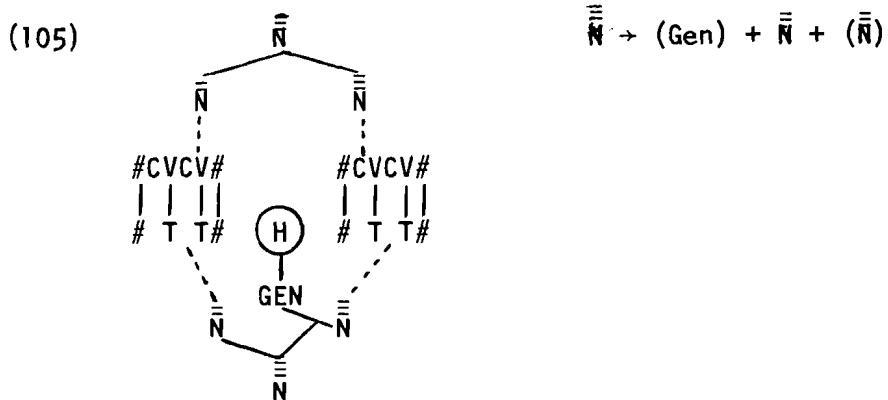
In (103) I have underlined a form where some
tonal modification occurs.

Indeed, let us take (104) as the underlying stucture of the
possessives.

(104)

NP
NP   ?
          Genitive   NP
ǫ́dhu̟   H   òké
tail   'rat'   'rat's
                                tail'
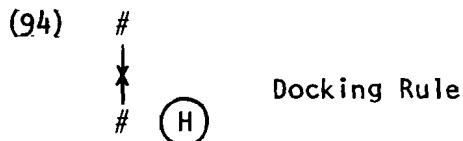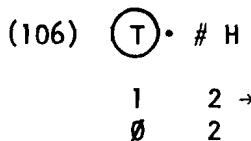
or more accurately (105)

(105)

$$\bar{\bar{\bar{N}}} \rightarrow (\text{Gen}) + \bar{\bar{N}} + (\bar{\bar{\bar{N}}})$$

The rule of Object Tone Mutation (see (40) above) is triggered by the H affix in (105). The floating H tone will cliticize leftward onto the first noun in (105) by the Docking Rule (94), repeated here

(94)

Docking Rule

This H tone must be deleted if the second noun begins with an H, however. We could write this rule as (106):

(106) (T)· # H

    1   2 →
    ∅   2

Indeed, this might be correct. However, Williams argues that the deletion of the H tone here is to be identified with the process summarized in the rule of Floating-H Deletion (54) on page 82 above. The formulation given there makes (106) superfluous; for once the Docking rule (94) applies, (54) Floating-H Deletion applies, undoing the effects of the Docking rule by deleting the H-tone. In short, no new rules are needed at all to deal with the tonal process summarized in (102) and (103).

FOOTNOTES FOR CHAPTER TWO

1.   That is to say, rules that are formally and technically
     "simple" in the autosegmental framework, but which would
     be relatively complex (and thus unexpected) if expressible
     within a standard theory.

1a. Some take this as evidence that the High Mid sequence is
     "really" High Low High. Of course this does not follow,
     although in some particular cases it may be correct. We
     shall see in a moment that the parallel dropping of the
     High pitch register is due to the fact that the same rule
     applies in both cases (essentially a generalization of (13)).

2.   We may say that an application A_1 of a transformation occurs
     before an application A_2 of a transformation if application
     A_1 corresponds to the pair of PM's (P_i, P_{i+1}) and A_2 corres-
     ponds to pair (P_j, P_{j+1}), where i is less than j.

3.   Two types of possible counterexamples have been suggested
     to this ordering of phonological rules (this ordering is
     the Sound Pattern conception, of course, not my own).
     The first is Bresnan's suggestion that the Nuclear Stress
     Rule applies cyclically; the second is the Bresnan-Selkirk
     analysis of the blocking of liaison phenomena. The
     Selkirk analysis--which explains this blocking in terms
     of word-boundaries left behind by transformations in their
     operations--is, in one form or another, not strictly
     speaking a counterexample to the Sound Pattern conception
     of ordering syntactic and phonological rules. The
     Bresnan Hypothesis is clearly contrary to this, however,
     and would require abandonment of the picture in (18) if
     her analysis of the Nuclear Stress Rule were correct.
     However, I do not find her examples convincing, and the
     replies in the literature--especially Bolinger's--are
     telling responses. If the Nuclear Stress Rule were a
     formal rule responsible for the placement of intonation
     (the star, in the sense developed in Chapter Three below).
     then Bresnan would have adduced strong evidence for its
     applying cyclically; the question, however, is whether
     there is such a rule at all that places sentential accent.
     If there is no such rule that looks just at structure--
     and I do not believe that there is--the question of when
     it operates does not arise.

4.   All the data from Igbo, unless otherwise specified, comes
     from Green and Igwe's grammar of Uhuhu Igbo.

5.   In fact the case here is stronger than I have suggested,
     for the connection between the cliticization of the pronoun,
     as in (27), and the non-application of the Flop rule is
     supported by some further facts. It is only the singular,

weak forms of the pronouns that undergo cliticization; the
diagnostic for cliticization is deletion of the a/e prefix
in the II Incomplete forms, as discussed above in Section 5
of Chapter One, and below in Sections 11-14 of Chapter Two.
Thus neither the singular strong pronouns nor the plural
pronouns cliticize. The singular strong pronouns do not
occur in the I Main form, but the plural ones do, and they
do undergo Flop, as illustrated in (28). Thus the non-
application of the Flop rule is directly connected, not
with pronounhood, but with cliticization.

6.   Throughout this thesis, I am using the following evalua-
     tion metric implicity:

     1.   Boundary elements, major category names (syntactic,
          N, A, V, N', A', V',...or phonological, C, V, T) and
          association lines are counted as equally costly; these
          elements, however, contribute a great deal less com-
          plexity than other, feature specifications, which are
          counted on an equal par: one feature, one cost unit.
          Boundary elements thus have a privileged position in
          the writing of phonological rules.

     2.   Epenthesis of a segment costs as 1 unit above and
          beyond the features required in the statement of the
          rule.

     3.   Elements left unchanged by a rule are not counted
          twice, of course, in counting the complexity of a rule,
          but only once--hence the notational convention of
          writing rules that insert association lines in one step
          by using dotted association lines to indicate that line
          is part of the structural change, not the structural
          description (see rule (25) for example).

7.   The statement that Low-toned verbs have a Mid-tone is made
     by Green and Igwe. Their few examples, however, of Low-
     toned verbs here are all without suffix. The present anal-
     ysis, therefore, is indeterminate as to whether the stem tone
     of the Low-toned verb is really M or L, given this limited
     data base. I shall accept Green and Igwe's statement that
     the stem tone is in fact Mid; comparison with the II Relative/
     Subject form strongly supports Green and Igwe's point, as we
     shall see below.

8.   Where the LH-Simplification rule is obligatory and where
     optional cannot be decided on the basis of the limited
     data available to me at present.

9.   As mentioned above, there are also uses of the I Relative B
     without a head, but rather preceded by the complementizer ka

which appears, as usual, with an L tone.

ex: Áñùrù m̀ 'únù kà únú vùsì ńghùghwù ń'ʃʃʃ

    I saw   you as you were   parcels on head
                             carrying   (your)

10. The lack of system in (63) should neither surprise nor disconcert us. It is upon the arbitrariness of such patterns that the argument in Section 4 of this chapter was based, arguing for lexical insertion of tonal melodies at surface structure, or after cyclic rules.

11. "B" stands for base, or underlying tone, as in Section 4 of Chapter One. This is meant in the following sense: there is a two-way underlying distinction for verb tone. Those we call L-verbs are generally realized on an L-tone: $\begin{bmatrix} -hi \\ +lo \end{bmatrix}$. The other class, the non-low or High verbs, are generally realized either as $M = \begin{bmatrix} +hi \\ +lo \end{bmatrix}$ or $H = \begin{bmatrix} +hi \\ -lo \end{bmatrix}$.
In either case, we see the underlying distinction determines the feature "high"; the specification for "low" is either simply plus, whence an M/L opposition; or the specification for "hi" is the opposite sign from that of "lo"--that is $\begin{bmatrix} -\alpha hi \\ \alpha lo \end{bmatrix}$ --whence an H/L opposition.

12. There is a non-discourse-initial form of the II Root. It differs from the discourse-initial form only with respect to the behavior of singular subject pronouns. In the non-discourse-initial form, the non-clitic or Strong forms are used, and they behave like any other lexical item.

(i) ʃ   bhá   yà,

   you grip   him

   'When you grip him,'

                 yá   àbhá   m̀
                 'he   grips me'
                 gʃ   àbhá   m̀
                 'you grip me'
                 mʊ́   àbhá   yà
                 'I   grip him' etc.

So far we have no reason to distinguish systematically (as Green and Igwe do) between the initiating and the non-initiating forms; all we have seen is that in the II Root form, cliticizing subjects are used if and only if the sentence is "discourse initial." All the rest would follow, and there would be no need for this footnote. But

there is one point remaining. There is no non-
cliticizin form of a, the impersonal subject pro-
noun; it is permitted in the II Main non-initiating.
And rather than being H, as in (73a) and (73b), it is L
for both L and H verbs, as in (ii).

(ii) <u>zà</u> 'sweep' (L-verb)

yá byá à zàá úló   'when he comes home, the
                           house will be swept'
he comes one sweep house
                         NB: <u>à zà á</u> is II Main
                                           form.

<u>ńú</u> 'see' (H-verb)

Ébé à gàlà, à ńú yà 'wherever one goes, one
                             sees him'
place one goes one sees him
                         NB: <u>à ńú</u> is II Main form.

This L-tone on the a is connected, no doubt, with the
fact that cliticization of the subject pronoun moves
the pronoun into the place of the prefix, thus taking
on its tone (see (73) above, the cliticization rule);
similar tone shifts occur below in the II Relative/
Non-Subject. Yet this leaves open the question as to
why the clitic pronouns are in fact high-toned in (72).
I have no answer at this point.

13.   Preferably one would have the "because" running in the
      other direction.

14.   Such decisions are ultimately empirically testable,
      although with difficulty. The system of rules in Igbo
      are the sort of evidence we need to decide upon which
      set of notation conventions most adequately reflect a
      language-learner's preferred grammar.

15.   If the relativization rule left behind word-boundaries
      in the subject NP position, some modification would need
      to be made. While we are not using Selkirk's system,
      if one did, it would require that a succession of con-
      tinguous association lines attached to word-boundaries
      be deleted. Nothing particular rides on this, as far
      as I can see.

16.   This particular example seems to be related to the
      more general problem of how to account for the fact
      that underlying forms in a generative phonology often

have as morpheme-structure constraints (i.e.,
redundancy measure) conditions that mimic the appli-
cation of the phonological rules. As I say, I have
nothing to add on this question.

17. For those interested in Igbo, the negtive forms go
essentially like this:

		Floating	Prefix	Stem	Suffix
Root	(i)	-	H	B*	-
	(ii)	-	L	HB*	-
Subord		-	H	B*	
Relative/ Subject		H	-	M	M-(or else stem tone is MM and
Relative/ Non-Subj	(i) (ii)	-	H	B*	suffix is H)

The two Relative/Non-Subject forms suggest very strongly
that there is indeed a synchronic connection between the
presence of a floating H-tone and the possibility of a
tone in the prefix position. There is another signifi-
cant fact here, but just how it should be interpreted
is not clear to me at all. We have equated the position
of the relative marker /na/ with the tonal position of
the floating H tone. In relative clauses, the /na/ is
optional, though preferred. Green and Igwe remark,
however, that they know of no cases where the /na/ is
used in the Relative/Non-Subject (i), which, under our
analysis, does not have a floating H-tone (it is the
only relative clause form which does not, in fact).

CHAPTER THREE

MELODIC ASSOCIATION RULES

1.   Introduction
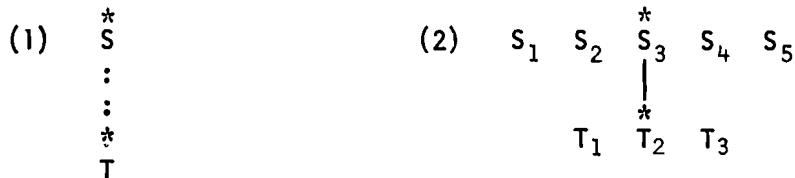
In the last chapter, we considered the nature of the auto-
segmental derivation, and so we looked at a set of rules that
modified the associations between two autosegmental tiers--the
tonological and the syllabic, or non-tonological. The particular
underlying forms chosen were argued for, but the discussion of
the rules that link the two tiers underlyingly--the "melodic asso-
ciation rules"--took a back seat to the discussion of the subse-
quent rules in the derivation.

The purpose of this chapter is to fill in part of this gap,
and to consider possible ways for autosegmental tiers to be
linked. In this sense, the present chapter considers the nature
of the grammar at a point just slightly "deeper" than that part
considered in Chapter Two. Chapter Four below, on the other hand,
will deal with prodesses more "superficial" than those of Chapter
Two. Thus while in Igbo we considered the number of autosegmental
tiers--let us call this the "autosegmental index"--to be fixed at
two, in Chapter Four we shall consider the possibility of the
autosegmental index increasing during the derivation. As we shall
see, furthermore, a stage in a derivation with more autosegmental
tiers is always more "superficial" than a stage with fewer tiers.

In the present chapter we shall consider the way in which two
strings of segments are selected from separate entries in the
lexicon and then associated with each other, thus permitting the
autosegmental derivation to proceed.

We shall assume that such a linking process, or "melodic
association rule," is ordered in the derivation at the earliest
point possible. Such a hypothesis is consistent with all lang-
uages I am aware of, and of course, serious constraints are
necessary for the theory regarding that phonological processes
may occur before any melodic association rule, and thus before
the Well-formedness Condition comes into effect. This last point--
the effects of the Well-formedness Condition--is our major em-
pirical diagnostic for determining the ordering of a phonological
rule before or after the major association rule.

In general, two types of melodic association procedures are
available: accentual and non-accentual. Accentual systems
distinguish themselves by assigning the tonal melody to the
syllabic tier by means of an abstract marker, which we shall
denote by an asterisk or "star." In the simplest case, there is
exactly one star in the domain corresponding to each tonal melody;
and on the tonal level, one tonal segment bears a star also. If
the language then contains the simplest accentual melodic asso-
ciation rule--(1)--then the two tiers will be matched as in (2).

(1)
$$
\begin{array}{c}
\overset{*}{S} \\
\vdots \\
\vdots \\
\overset{*}{T}
\end{array}
$$

(2)
$$
S_1 \quad S_2 \quad \overset{*}{S_3} \quad S_4 \quad S_5 \\
\qquad\qquad | \\
\qquad T_1 \quad \overset{*}{T_2} \quad T_3
$$

We may, however, find either too many or too few stars on the syllabic level. Thus it may be that no star appears, and some provision must be made for aligning the two tiers. On the other hand, there may be more than one star available, and some decision must be made to indicate which star has precedence in the application of rule (1) or whichever melodic association rule occurs in the language.

The effect of accentual association rules is to make a certain vowel or syllable prominent or "distinguished," and this apparent purpose has certain consequences for the Well-formedness Condition. While the Well-formedness Condition requires certain restructurings of the autosegmental representation at times, it may leave open more than one way to do this. The nature of accentual systems leads to a lessening of this ambiguity in ways that preserve the prominence of the starred segments.

The alternative to accentual systems is the type of system we have already seen in Igbo and, more briefly, in Tiv and Mende, in which only categorial information was involved in aligning the tonal and the syllabic tiers. In the simplest case, the tone melody is mapped left-to-right, one-to-one, onto the syllables. This pattern of association has been considered with respect to Etung by Edmundson and Bendor-Samuel (though rejected by them), Margi by Williams, and Mende and other languages by Leben. We shall consider several cases of this and other principles in this chapter.

## 2. English as a Tone Language
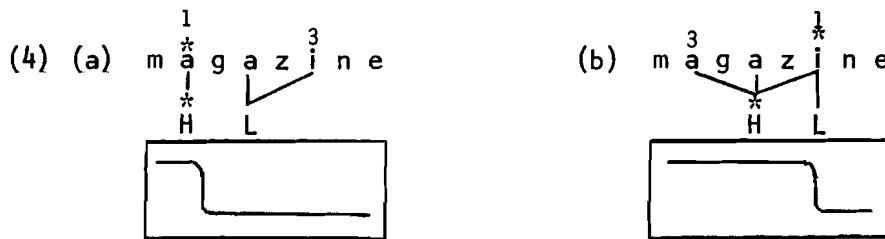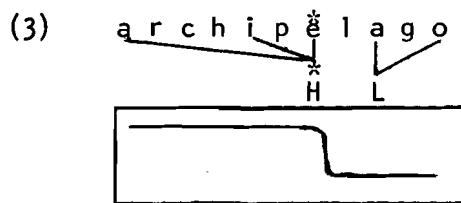
We shall begin with a look at the tone patterns of the simplest part of the English pitch system, the utterance of words in isolation, as at the beginning of a dictionary entry—reserving for the moment judgment as to whether this reflects tonal or pitch registral rules, or something else still different. These citation forms, furthermore, are the forms that standard phonological treatments of English concern themselves with.

A number of characteristically autosegmental traits show up in this system. First, we find contour tones on short vowels which must be analyzed as sequences of level tones. Second, we find independent tone melodies in the lexicon. Third--and here there is a caveat[1]--the tone melody spreads itself over the syllabic

structure in a characteristically autosegmental way.

All of this will lead us to the conclusion that tone is autosegmental in English. We have already seen some of the tonal analysis for English in Chapter One, Section 2. There we found that the neutral melody for English was ĤL (or MĤL). The starred toneme is associated with the syllabic which, in *The Sound Pattern of English*, Trager and Smith system would receive a "l-stress." In Liberman (1975), it is called the "designated terminal element."

In Chapter One we looked at the example of *archipelago*, sketched in (3); the same point--that "l-stress" is the syllable associated with the starred H toneme--is clearly brought out by such minimal pairs as m<u>á</u>gazine/m<u>à</u>gaz<u>í</u>ne, c<u>í</u>garèt/c<u>ì</u>garét, and so forth. What could be represented as the difference between a "l" and a "3" is simply that the pitch contour falls on the l-stressed syllable, as sketched in (4).

(3)   a r c h i p é l a g o

(4) (a) m á g a z i n e   (b) m a g a z i n e

For my speech, these frequencies are physically approximately the following: in initially accented *magazine* (4a), the first syllable is at 100 hz.; the second two at about 84 hz. When the accent is final (4b), the pitch of the first two syllables rises from an initial 92 hz. to a peak of about 102 hz. falling shaply on the final syllable to about 87 hz.

The theory being proposed here is making heavy use of what is called the "l-stress"; the other levels of stress inside the word, 3-stress and no stress, are playing no role. In (4a) and (4b), the H toneme and the L toneme spread themselves through the graces of the Well-formedness Condition, but irrespective of the nature of stress on the syllable level. What is noted as "3-stress" is a segmental characteristic which has an important consequence for vowel quality, but little effect on such suprasegmental matters

as pitch or intonation. A stressed vowel will tend, perhaps, to be a slight bit higher in pitch due to the air pressure, but not significantly.

Both the "1-stressed" vowel and the "3-stress" vowel differ from the unstressed vowel by two major processes that Chomsky and Halle demonstrated in *The Sound Pattern of English*. Put simply, vowels divide into those that are tense and those that are lax; and among lax vowels, those and only those that are unstressed reduce to schwa. Among tense vowels, only those that are stressed undergo vowel shift.

The assignment of the "1-stress," or star, is given by the provisions of Schane's "detail" rule, cited in Halle (1973):

(5) $V \rightarrow * / \text{--} Q$ $(VC_0(+y))$ ## Condition: $Q \neq [+\text{stress}]$

Chomsky and Halle, following Trager and Smith, produce a description of English with several stress levels, where these stress assignments are essentially segmental features. Liberman (1975) proposes a revision that maintains a binary segmental stress feature, but provides an <u>organization</u> of neighboring stresses, an organization of relative weaker and stronger, so that given a word of arbitary length, there can be stressed vowels subordinated to any arbitrary degree with respect to the main-stress.

Our primary interest here rests in showing that pitch phenomena in English display certain autosegmental characteristics. Thus Bolinger's insight--that main-stressed syllables are manifest as pitch extrusions, either up or down--is clarified and extended in an autosegmental analysis, which furthermore provides that the extrusion be essentially level in tone, except in the case where the mainstress is phrase-final, in which case an "extra" toneme may be associated by the Well-formedness Condition, creating a gliding tone. We have already seen examples of this in Chapter One, and we will return to it below, when we consider other melodies.

The existence of non-primary stresses is, therefore, not our central concern, nor their organization in time; but we may ask, if main stress manifests itself as a pitch extrusion, do secondary and tertiary stresses do likewise? If not, should this surprise us?

Chomsky and Halle, who defend a multiple stress-level analysis in *The Sound Pattern of English*, suggest the following, concerning the reality of stress-levels:

We do not doubt that the stress contours and other phonetic
facts that are recorded by careful phoneticians and that we
will study here constitute some sort of perceptual reality
for those who know the language in question. In fact we are sug-
gesting a principled explanation for this conclusion. A
person who knows the language should "hear" the predicted
phonetic shapes. In particular, the careful and sophisticated
impressionistic phonetician who knows the language should be
able to bring this perceptual reality to the level of aware-
ness, and there is ample evidence that phoneticians are
capable of doing this. We take for granted, then, that
phonetic representations describe a perceptual reality.
Our problem is to provide an explanation for these facts.
Notice, however, that there is nothing to suggest that these
phonetic representations also describe a physical or acoustic
reality in any detail. For example, there is little reason
to suppose that the perceived stress contour must represent
some physical property of the utterance in a point-by-point
fashion; a speaker who utilizes the principle of the trans-
formational cycle and the Compound and Nuclear Stress Rules
should "hear" the stress contour of the utterance that he
perceives and understands, whether or not it is physically
present in any detail. In fact, there is no evidence from
experimental phonetics to suggest that these contours are
actually present as physical properties of utterances
perceived (26).

It is to be expected that determined [i.e., not freely chosen
--JG] phonetic features should be quite difficult for the
user of the language to learn to identify, whether they
involve stress or degree of aspiration (where undoubtedly
there are many levels, predictable, at least roughly, by
general rules). The apparent ease with which phoneticians
trained in the same convention can, to a large extent, agree
on the assignment of four or five stresses in utterances
may very well be traceable to their ability, as speakers of
the language, to grasp the syntactic structure of utterances
and to assign to them an "ideal" stress contour by the rules
of the transformational cycle. . .

To summarize this discussion of phonetic representation,
we do not doubt that representations of stress contours and
similar predictable phenomena correspond, up to a point, to
some perceptual reality that can be brought to consciousness
with training and care...the representation of the perceptual
facts is likely to be governed in part by arbitrary convention
or irrelevant cognitive limitations after a certain degree of
complexity is reached. Thus it is impossible to expect (and,
for purposes of investigating linguistic structure, unneces-
sary to attain) a complete correspondence between the records
of the impressionistic phonetician and what is predicted by
a systematic theory that seeks to account for the perceptual
facts that underlie these records (27).

It is difficult to agree with this passage for the following reasons: (1) Chomsky and Halle assume a consensus among a group of linguists ("careful and sophisticated impressionistic phoneticians," in particular) regarding the representation of--if we can find a theoretically neutral term--syllabic prominence. In fact, there is no such consensus or uniformity. There are traditions with their notations; but there certainly is no gereral unanimity.

Vanvek (1960), for example, says, "One is inclined to think that it is possible to perceive secondary stress since phoneticians have written about it. But often, I think, it is really inherent prominence that one hears" (i.e. non-reduced vowel quality--J.G.) (p. 73). Later, "It seems doubtful whether it is at all necessary to mark secondary stress after a full stress in the same word, e.g., *foot-passenger, make-believe.* The inherent prominence of vowels following the fully stressed syllable seems to settle the question automatically in the above cases. Finally, there are cases like *subcommittee, telerecord* (noun), *unliquidated* (attributively), I wonder whether English people really hear a secondary stress here, or whether D. Jones is influended by the words *committee, record* and *liquidate* when he puts the secondary stress mark" (75).

On another line, when dealing with forms like *light-house keeper,* and *light house-keeper,* Bolinger and Gerstman (1957) conclude "Since the disjunctures transparently supply a physical separation whose width corresponds inversely to the semantic bond it follows that the disjunctures function directly to carry the information, and not indirectly as components of a hypothetical stress. Rather than attempt to salvage the stresses it would seem better to reconsider the juncture complex, for it is apparent that the uniform "plus" juncture with which Trager and Smith have marked the examples has overlooked the function of disjuncture."

(ii) The issue is--and has always been--not whether certain differences exist, but rather that the linguistically or psychologically correct way is to represent these differences, at the physical, the phonetic, the phonological, and whatever other levels may be relevant. If the n-stress system has a "perceptual" reality--and a reality dependent on the obsever's familiarity with the language at hand--then the "careful phonetician," whose training is intended to overcome the particularities of a language and attain a language--independent description, would be in a poor, not privileged, position to report on these "perceptually real" stress levels--*qua* phonetician.

(iii) Whether the predicted stress contour matches the physical description of the signal "in a point-by-point fashion" is not the issue; it is certainly possible that the most adequate treatment of English stress contains of a level of n-stress, as

Chomsky and Halle suggest, and then translates this in a complex
way into a combination of physical cues. 1-stress would be phys-
ically realized in a certain way; 2-stress by some other way,
3-stress by yet a third, and so on. While this would be surprising,
it would not affect the validity of Chomsky and Halle's position.
Their position would rest supported if, at a certain level in the
derivation, the n-stress levels were the appropriate representa-
tion, even if this level is not closely related to either the
phonological level or the acoustic level (if we admit an acoustic
"level" in our conception of grammar).

If, however, the relevant properties of the acoustic signal
--syllabic length, intensity, and pitch--can be predicted from
other considerations, like autosegmental spreading of tone, and
placement of word-boundaries, then the n-stress level system
would not be empirically supported. In fact, the question
remains an open one; there is as yet no successful system
that explains English syllable length, intensity and pitch.
However, we shall look at some cases below, and the facts
tend to the conclusion that whatever difficulties remain in
present theories, the analyses would not be improved by
having access to a level with n-stress features in the grammar.
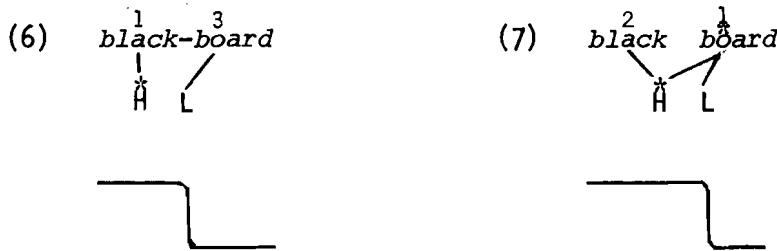
It would no doubt be worth stepping back from this
particular issue for a moment to see the broader consequences
of this issue, for whether there are two levels of stress in
English or four is not of great significance to anyone in and
of itself, I hope. One related issue is the existence of the
transformational cycle; I refer the reader for a discussion of
this to Chomsky and Halle (1968), Chapters Two and Three, and
Liberman (1975), Section 4.1. Another issue is whether there
is a linguistically real distinction between pitch-accent and
stress-accent languages like Japanese (see Haraguchi (1975))
can be dealt with in the same way as so-called stress-accent
languages like English, the distinction is untenable.

Returning one last time to Chomsky and Halle's remarks:
"the representation of the perceptual facts is likely to be
governed in part by arbitrary convention...after a certain
degree of complexity is reached." On this we are agreed; the
question is, now, whether inclusion of "3-stress" levels, and
so on, is arbitrary, or part of the linguistically real system.

We shall argue that (i) the perceived difference between
the 3-stress in (6) and the 2-stress in (7) below is due to the
difference in length; that (ii) this difference is not a general
fact about syllables designated as 2-stress and 3-stress, and not
even a general fact distinguishing compounds from adjective-noun
pairs (as in (6)/(7)), for it occurs in only a subpart of these

forms; and that (iii) therefore the difference in length is not attributable to a difference in an abstract n-stress feature, but to certain straightforward lengthening processes sensitive to word-boundaries (#); we shall see that such contrasts as those between *black-board* (that one wrties on) and *Black board* (a committee presided over by Justice Hugo Black) show that certain actual differences would not be describable if length were predictable from a n-stress feature.

Consider first the compound *black-board*, an object that may as well be green or white as black; it contrasts with *black board*, which is black but likely has nothing to do with chalk (we shall consistently use the hyphen to distinguish these). *Black-board*'s pitch pattern is like that of *magazine* (see (6)); *black board*'s is as in (7).

$$(6) \quad black\text{-}board \qquad\qquad (7) \quad black \; board$$

The 2-stress on *black* in (70) is supposed to indicate a prominence that *board*—a 3-stress—does not have in (6). Although the representations in (6) and (7) do reveal a difference between the two syllables—*black* is H-toned, *board* is L-toned—a glance at (4b) shows that this could not be the relevant difference, since 3-stressed syllables can be H-toned. 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 (6)-(7) minimally different by giving them the same tonal pattern as in the context, "But that's my --," which is roughly as in (8).

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

For some reason—which we shall return to—the *black* in (7) *black board* is considerably longer in duration than the *black* of (6) *black-board*.

Now we shall see that this difference of length in (6) and (7) is not due to a difference in stress patterns, if stress patterns is taken as in *Sound Pattern*. Other 2-stresses, for example, 5o not lengthen as the *black* in (7). Thus (9), when the cyclic application of stress in the SPE system applies to give a 1-2-4 stress pattern, displays no lengthening of *black*

in *black-board*.

(9)   *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 (10), where the adjective is not monosyllabic, displays no lengthening of the first syllable corresponding to the *black* in (7).

(10) $\overset{2}{h}$*acking* $\overset{1}{co}$*ugh*

In general, then, being marked as 2-stress is not a sufficient condition to undergo the lengthening displayed in (7).

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 (6)/(7) 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 to monosyllables surrounded by double word-boundaries (11), if we assume that lexical compounds like *black-board* have the structure *##black# board##*.

(11)   Monosyllabic Lengthening Rule
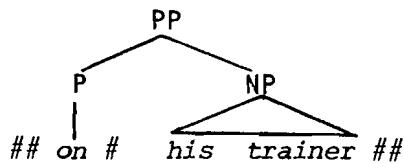
$$V \rightarrow \text{lengthened} / \#\#C_0 - C_0\#\#$$

The assumption that lexical compounds have just the single word-boundary, rather than the double word-boundary proposed in SPE, 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.
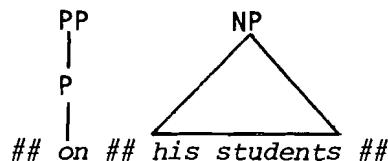
The Monosyllabic Lengthening Rule (11) mentions nothing about stress, and so we would expect it to apply equally to

1-stressed ("starred") vowels and to the 2-stresses we have
seen. This is, as we observed above, true (*The board was bláck*)[1];
the distinction is equally manifested in the pair in (12):

(12)   *The lion turned on his trainer*   (no lengthening)

PP
├── P
│   │
│   ## on #
└── NP
     his trainer ##
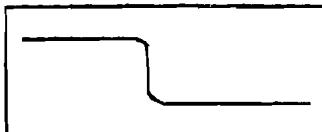
The teacher turned on his students   (lengthening)

PP
│
P
│
on ##   NP   his students

   Furthermore, there are compounds that would be described in
the n-stress system as having 1-3 patterns which display a differ-
ent lengthening. If there were a committee chaired by Justice
Black, such a *Bláck board*, though starred on the first syllable,
thus displaying a tone pattern as in (13), is clearly pronounced
differently from *bláck-bóard*.

(13)   B l á c k - b o a r d
           H   L

   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 Lengthening Rule, suggesting two word-boundar-
ies in (13), 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 re-
cordings 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 0 1 0 Ticonderoga	4 3 0 1 0 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 per-
ceived 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 *Ticongeroga* 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.
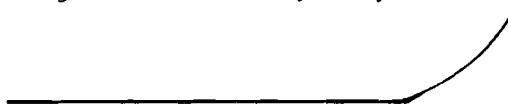
Let us turn to some more complex intonational examples from English, where the tonal melody is not the simple H L. Consider the disjunctive question, the sort where the asker assumes that the person asked wants one of the choices: "Do you want coffee, tea, or milk?" asked in such a way that the person asked feels it incumbent upon him or her to take one of the three. This could be sketched as in (14).

(14)   *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 (15).

(15)   *Do you want coffee, tea, or milk?*

(15) is the intonation of a normal, polite yes/no question, which we shall not consider here.

The choices in a disjunctive question may number anywhere from two up, since to make a choice, at least two options must be present. If the question has N options--in (14), 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 or offer is the final low of that phrase, in the first N-1 options. In short, the pitch pattern is exactly what would derive from the autosegmental formula L H for each of the first N-1 disjuncts. When the disjunct has the star on the final syllable, this is realized, as we would expect, as a rising tone ("tea," e.g., in (14)).

Precisely the same type of observation leads to the conclusion that the tone melody for the final disjunct is H L. To derive (14), the autosegmental representation is (16), where corresponding boundary elements ($) have been linked.

(16) $ *Do you want cóffee* $ *téa* $ *or mílk* $
         | | | |
         $ L̇ H $ L̇ H $ H̊ L $

In each domain defined by the boundaries, the major association rule (1) applies, giving (17), which becomes (18) through the Well-formedness Condition.

(17)

$$\$ \quad Do \ you \ want \ c\overset{*}{o}ffee \quad \$ \quad t\overset{*}{e}a \quad \$ \quad or \ m\overset{*}{i}lk \quad \$$$

(18)

$$\$ \quad Do \ you \ want \ c\overset{*}{o}ffee \quad \$ \quad tea \quad \$ \quad or \ m\overset{}{i}lk \quad \$$$

This (18) describes formally what we sketched in (14). The formula for the disjunctive question, then, is as in (19).

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

There is another possible intonation contour for the disjunctive question in (14), one which is slightly less polite, it seems to me, which looks like (20).

(20) $Do \ you \ want \ c\overset{*}{o}ffee \quad t\overset{*}{e}a \quad or \ m\overset{*}{i}lk?$

The formula for such an intonation is (21); we leave off the derivation, for it proceeds just like the others we have seen.

(21) $\$ \ L \ (\overset{*}{H} \ \$)^n_1 \ \overset{*}{H} \ L \ \$$
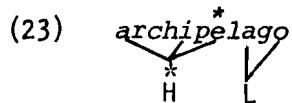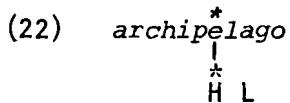
Compare this with (19). It is more than slightly remarkable that these two contours, used in about the same sense, have tone contours which, in this formalism, differ precisely in the parenthesization. Note that the hopping of the star from the L, in (19) to the H in (21) is in effect predictable; every phrase must have a star so that it may be associated with the phonology in English, and if the L had taken the star with it when it left the parenthesis in (19), this would not have occurred.

The Well-Formedness Condition in Accentual Systems

In the discussion thus far of the application of the Well-formedness Condition to accentual systems, it has been made to do more work than we might rightfully expect of it. In the first example that we considered--*archipelago*--the Well-formedness Condition revised (22) to (23).

(22) archipélago
        |
        *
        H L

(23) archipélago
        *
        H   L

While it is true that (22) does not satisfy the Condition and that (23) does, the revision to (23) was not, strictly speaking, forced. (24) and (25) also satisly the Condition; why was (23) selected rather than either of them?

(24) archipélago
        *
        H   L

(25) archipélago
        *
        H L
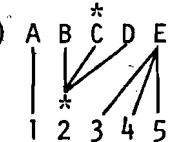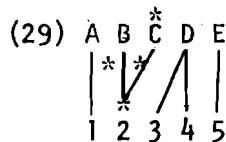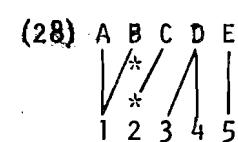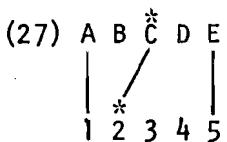
Why (24) was not created from (22) is clear; it contains (23) as a subpart, and (23) does satisfy the Condition. Therefore (24) would not be a __minimal__ way to modify the structure (22) in order to satisfy the Condition. But any difference to be found between (22) and (25) must make reference to the star, for (25) would have been the correct structure, rather than (23), had the star been on the fourth syllable rather than the third.

The principle that we need is one that insures that the presence of the star on a syllable on the syllabic level finds its way into the association structure. One simple way to state this is:

(26) Given ambiguity in ways to fulfill the
      Well-formedness Condition, do not reassociate
      a starred segment.

That is, if there is a choice about associating a segment with a starred or an unstarred element, associate it with the unstarred one. The point is to preserve the prominence of the star, or accent. A structure like (27) will reassociate by this interpretation of the Well-formedness Condition to (28), rather than (29) or (30). If we mark each association line associated with a starred element, we see more graphically that the output (28) is preferred because the starred association is most unambiguously identifiable.

(27) A B C D E   (28) A B C D E   (29) A B C D E   (30) A B C D E

     1 2 3 4 5   1 2 3 4 5   1 2 3 4 5   1 2 3 4 5

We will modify this point somewhat below; see also the discussion of Sanskrit in May and Goldsmith (1975).

It is worth nothing that this is the first use of the star in autosegmental rules so far other than for the initial major association rule.

3. Non-Accentual Melodic Association Rules

Determining the conventions on the interaction of the
star and the Well-formedness Condition is intimately connected
with determining the types of possible melodic association rule,
since in general the more powerful the Well-formedness Condition
is, the less will the language-particular rule need to do. We
have thus far considered only a simple accentual system--that of
English--and proposed a strengthening of the interpretation of
the Well-formedness Condition in line with this.

In general we would not expect all melodic association
rules to be as simple as (1), the "star-to-star" association
rule, if only because not all languages have a star (accent).
In Section 4 of Chapter One we reported on Leben's analysis of
Mende, where he showed that the tonal melodies are mapped
one-tone-to-one-vowel in a left-to-right procedure. Repeating
the tonal classes here, we find five groups:

(31) H   pélé, kɔ́

      L   bèlè, kpà

      HL   kényà, mbû

      LH   nìká, nàvɔ́, mbǎ

      LHL nìkílì, nyàhâ, mbã̀

Whether we say that the one-to-one mapping is a productive
process or not--essentially depending on whether we take these
words to arise from two separate lexical entries, one syllabic
and one tonal--nonetheless it cannot be an accident that the
pattern associates with the syllables in a particular, non-
arbitrary way. And we must propose a formalism that expresses
this intuition. In Luganda, below, we shall suggest that such
a formalism is necessary for a productive part of the grammar,
for the melodies and the syllabic morphemes are underlyingly
separate. The same point has already been observed in the
case of Tiv in Chapter One, Section 3.2.
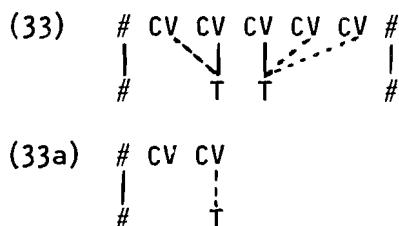
What factors must be stated in the association rule that
are not part of the universal theory so far? The fact that the
mapping is left-to-right, for one thing, follows from nothing
we have seen so far. Yet all the examples we have seen, or
will see, of non-accentual systems bear this property. We may
observe, then, without any principled explanation, that the
unmarked pattern is left-to-right mapping.

We can approach this, however, with certain principled
hypothesis. Consider, for example, the Local Environment Con-
dition, proposed in line with Sanskrit in May and Goldsmith(1975).

(32) Local Environment Condition (tentative): No autosegmental rule may mention in its structural description a segment in a syllable, or a toneme, unless that segment is (in a syllable which is) affected by the rule, or associated with a syllable or toneme which is affected by the rule.

The local environment condition proposes that elements that condition an autosegmental rule must be affected by the rule, in a certain sense. As it stands, the LE Condition is probably too strong (see, e.g., rule (54) of Chapter 2, a counter-example), but it is, I think, in the right direction.

One situation the LE Condition rules out is where judicious use of symbols permits a phonological rule to count. It rules out, for example, the possiblity of a language like Mende but where the melodic association began on the second syllable. This would look like (33), and such a rule would contain within it something like (33a) (leaving aside details we have not considered yet).

(33) # CV CV CV CV CV #

(33a) # CV CV

One cannot speculate too much at this point, given our limited knowledge of devices employed by existing languages. The strategy illustrated in (33), for example, is surprisingly similar to one that might be proposed.

If these suggestions are correct, however, they imply that non-accentual mapping must start from the extreme left end or the extreme right. Once again, while this is intuitively plausible, it is in no sense the null hypothesis.

A case that we wish to rule out--a language like (33)--would reveal itself, as we see in (33), by having tone melodies like Mende's but which began on the second syllable, the first syllable matching the second because of the Well-formedness Condition. The question of the possibility of melodies which begin with two identical tones then arises. The task of falsifying the hypothesis suggested here would then be simpler if we could be sure that tone melodies never begin with two identical tones. This would be a consequence of a proposal made by Leben, cited above in Chapter One as the Obligatory Contour Principle: Adjacent underlying tonemes must be distinct. The system of Etung, however, appears to show conclusively that this hypothesis,

while extremely attractive, is too strong.

The Etung system is sketched by Edmundson and Bendor-Samuel. They observe,

> In an initial analysis of Etung, four contrastive
> pitches are found: high, low, rising, and falling...
> In examining the occurrence of the rising and
> falling tones it is found that these usually occur on
> the final syllable of the phonological word. In
> addition, these tones very rarely occur on three
> syllable words, but rather on words of one or two
> syllables. Thus, the rising and falling tones are
> generally found on the final syllables of one or two
> syllable words.

> ...An analysis was therefore set up stating tone as a
> feature of the phonological word, and this made it possible
> possible to eliminate the two glides [Rising and falling
> --JAG] as separate tone units and so to describe Etung
> in terms of high and low tones plus downstep.

> With this analysis, a limited number of tone patterns
> are set up for the phonological word.

They then consider two ways to state the generalizations more precisely. The first would be to describe "each tone pattern as consisting of a sequence of three tones," and then to express how this pattern is spread out over one, two, and three syllable words.

"A more satisfactory type of statement," they continue, "would set up these patterns as contours marking the phonological word rather than as a succession of three tones which are compressed in certain circumstances [i.e., two level tones compressed to a glide in mono- or bi-syllable word, JAG]. In this way, pattern 1 [L, LL, or LLL--JAG] would consist of a level, low contour. Pattern 2 [LH, L H, L H H--JAG] would consist of a contour starting low and rising to high... Statements could easily be made for the patterns and in addition such statements could show how each contour was realized in words of three, two, or one syllables."
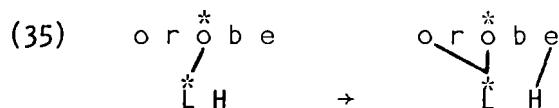
Let us look at their tone pattern classes (I exclude three because they do not seem to be underlying tone classes, but this is irrelevant to the argument at hand).[2]

(34)		1 Syll	2 Syll	3 Syll
1)	L	kpὲ 'even'	ǹjòm 'juju'	èyürì 'dress'
		bὲm 'softly'	ègù 'evening'	ǹkǜì 'corn'
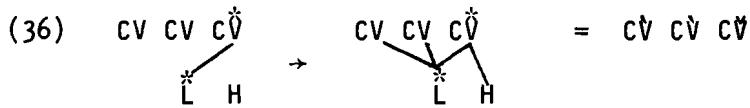
2) LH   kǎ 'to, into'   ǹsí 'fish'   bìsóŋé 'spoon'

         nǒ 'how'   èkát'.leg'   ŋ̀káé 'wife'

3) HL   tîk (future   égòm 'jaundice'   ákpùgà 'money'
              particle)
         nâ 'it is'   ódà 'platform'   bírəmə̀ 'me'

4) H   kpá 'first'   gbáń 'servant'   ŋ́kímí 'prosecu-
                                                       ter'
         tʃén 'all'   ńsé 'father'   ékúé 'forest'

5) LHL   -   òbô 'arm'   m̀bútà 'rain'

                     ètʃî 'head'   m̀búè 'goat'

6) LLH   -   ǹsǐ 'mud'   òròbé 'beam'

                     èbǐn 'farm'   ǹjàé 'hunger'

7) HHL   -   éfô 'cloth'   ŋ́gárè 'pepper'

                     érôp 'spear'   ésébè 'sand'

8) HLH   -   ábǒ 'they'   édìmbá 'pot'

     The conclusion is clear: we have in Etung a small class
of tone formulas that may be spread over words of one, two, or
three syllables, proceding from left to right. These melodies
are L, LH, HL, H, LHL, LLH, HHL, and HLH. The occurrence of
contour tones that motivate Leben's left-to-right mapping in
Mende occur here, but the melodies HHL and LLH make it clear
that the Obligatory Contour Principle is too strong in fact.

     In an attempt to preserve the Obligatory Contour principle,
one might reply that there is no melody L L H: if you find forms
like òròbé, then that must come from LH. The form we find just
proves that, to our surprise, Etung is an accentual language:
the melody L H has an accent--on the L--and òròbé is accented on
the second syllable. Thus:

(35)   o r o̊ b e   o r o̊ b e
           /   \ /
          L H   →   L̊ H

Why is this wrong? While this move is less interesting *a priori* since a highly counter-intuitive ploy had to be called on to preseve the Obligatory Contour Principle, this does not prove it wrong; it is conceivable, of course, that (35) is right. Nor does it prove (35) wrong that this line of explanation leaves unexplained what was explained before: that there are no words with the tone pattern CV̀ CV́ CV̂, which would have, in the non-accentual account, come from the putative pattern LLLH. On the non-accentual scheme, this is an accidental gap, because "final-accented" words with the tonal-pattern L̃ H should be CV̀ CV́ CV̂ (see (36)). So the accentual scheme can't explain this gap; this too does not prove it wrong. There are such things as accidental gaps. The real question is: how does a language learner decide if the language is accentual or not?

(36)   CV CV CV̈   CV CV CV̈   = CV̀ CV́ CV̂
          ╱   →   ╲╲╱
        L̃   H   L̃   H̃
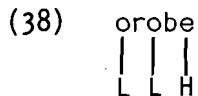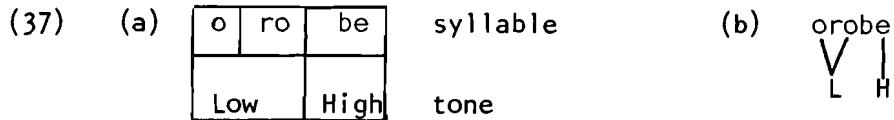
One plausible initial hypothesis might arise from the observation that while English, an accentual language, has only one neutral intonation contour, Igbo, Etung, and Mende appear to have several tone contours which are essentially neutral--they may appear in isolation, for example. We might speculate, then, that an accentual system may not, in principle, have more than one neutral tone melody. (A non-neutral melody is one used to express a question, or in some other way convey a speaker's particular attitude.)

This speculation, while it appears to contain a germ of truth, is wrong. Goldsmith (1975a) and, in much more detail, Haraguchi (1975) show that in many Japanese dialects, accentual systems coexist with more than one neutral tone melody.

I believe we should look in a different direction for a deeper understanding of the interaction of the Obligatory Contour Principle and accentuation. The Obligatory Contour Principle has going for it the fact that it is almost always true: this hardly seems accidental. The examples from Etung, however, disconfirm it in its strongest form.

I would suggest that the Obligatory Contour Principle is not a condition on underlying forms, but a reflex of the way in which underlying forms are learned. In Chapter Four below, we shall consider certain related questions in more depth; for now, let us observe that a word òròbé has as its "tonal projection"-- that part of it that is purely tonal--simply the melody L H; see (37).

(37) (a)

o	ro	be
Low		High

syllable

tone

(b) orobe

V |

L H

(38) orobe

| | |

L L H

we shall suggest that it is a more complex operation for the language--learner to change (37b) to (38) than to leave the tonal melody as it stands in (37).

Nonetheless, in order to map the melodies (LH,LHL, etc.) in Etung onto the syllabic structure, the language-learner constructs a on
this, the more marked--more complex--operation of expanding the one phonetic L into two underlying L's is called for, violating, in the final grammar, the Obligatory Contour Principle.

In short, the Obligatory Contour Principle is a condition not on possible underlying forms, but on simply-learnable grammars; not all grammars, however, need be simple. Furthermore, the Principle arises as an artifact of the way in which tone is processed (though this more important point has not been argued here: we consider this in Chapter Four).

The one-to-one tone-to-syllable mapping, as we observed above, starts from the left. Suppose we had a rule, corresponding to the accentual rule (1) above, that simply attached the first tone to the first syllable:

(39) # CV

| |

| |

| |

T T

(1) repeated $\overset{*}{V}$

.

.

.

$\overset{*}{T}$

They would operate on (40) to give (41).

(40) # e dimba #

| |

H L H

(41) # e dimba #

| |

H L H
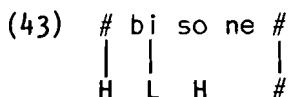
Now there is only one way for the Well-formedness Condition to be satisfied, as defined in Chapter One; (41) must become (42).

(42) # e di mba #

| | | | |

H L H
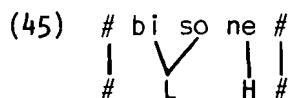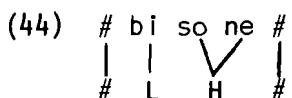
What is missing from this account is a notion of left-to-right.
We have posited the simplest possible non-accentual major
association rule (39) (compare it with (1)), and it says
'left", but it doesn't add "-to-right"; the rest seems to be
taken care of by the Well-formedness Condition. We can see
by inspection that (42) adds two association lines to (41),
that there are no ways of satisfying the Well-formedness
Condition by adding only one, and that there is no other way
of satisfying it by adding only two. Therefore (42) is
unambiguously derived.

The only cases where the Well-formedness Condition as it
stands does not suffice to give the correct output on the
basis of rule (39) is for cases like a tri-syllabic word
associated with a two-tone melody, as in (43), or the parallel
case of a bisyllabic word with a three-tone melody.

(43)   # bi so ne #
        | | |
        H L H   #

(43) could be corrected minimally either to (44) or to (45);
in fact, only (44) is correct.

(44)   # bi so ne #   (45)   # bi so ne #
        | | \/ | | \/ | |
        # L H #   # L   H #

We could take this as evidence either that the language-parti-
cular rule is more complicated than the simple (39), or that
the Well-formedness Condition would, in effect, tend not to
reassociate a starred segment.

(22) archipélago   (23) archipélago   (25) archipélago
          | \/ \/   \/ |
          H L   H   L   H   L
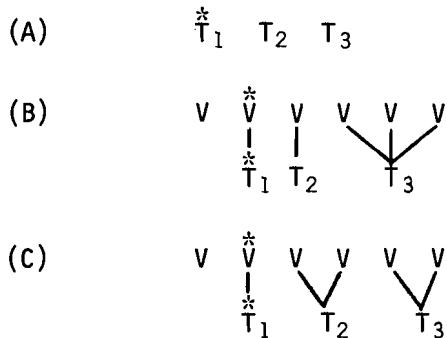
The situation is parallel in (43-45), but there is no
accent. We may simply say (and thereby render the accentual
modification (26) of the Well-formedness Condition unnecessary
in its specificity) that

(46)   Given an ambiguity in ways to fulfill the
        Well-formedness Condition, do not reassociate
        an already associated segment.

Methodology, the attempt to make the Well-formedness
Condition stronger in exchange for simpler language-particular
rules must be preferable to the alternative course of making
a more complex language-particular rule. The first option may
ultimately prove wrong, but it is the only one which offers to
tell us something about the nature of tone languages in
general, and thus about language in general.

This stronger option--which says that "left-to-right" spreading is not a rule-governed operation--predicts further, given our present understanding of the Well-formedness Condition, that there could not be two free unassociated tones on the same side of an associated tone, as in (A)--or if this arose, there would be any number of ways to associate this, as in (B) or (C). Only in the case where there were only three syllables would the arbitrariness disappear, as in (43). Now, in fact it seems to be correct that tone languages implying the so called left-to-right spreading for melodies of three tones are restricted to three syllable words.[3] It is also the case that in the extended analysis of Haraguchi for numerous Japanese dialects (1975) and Liberman for English (1975),[4] there are tone melodies of the form (T_1) T_2 (T_3), but not of the form (T_1) T_2 T_3 T_4--that is, a situation where there would be two free tones on the right side of T_2 after the application of rule (1).

(A) $\overset{*}{T}_1$ T_2 T_3

(B)

$$V \quad \overset{*}{V} \quad V \quad V \quad V \quad V$$
$$| \quad | \quad \quad \bigvee$$
$$\overset{*}{T}_1 \quad T_2 \quad \quad T_3$$

(C)

$$V \quad \overset{*}{V} \quad V \quad V \quad V \quad V$$
$$| \quad \quad \bigvee \quad \bigvee$$
$$\overset{*}{T}_1 \quad \quad T_2 \quad \quad T_3$$

It is not difficult to imagine how these observations might be made to follow in a principled way from languages using major association rules of the type in (39) or (1)-- that is, major association rules without variables.

I shall not, however, persue this here. Such a direction would be extremely attractive if one could claim that (1) and (39) were essentially the only choices a language had in the way of major association rules. The next examples, though, appear to revoke the possiblity.

4.   Ganda

The following paradigm comes from Ganda, or Luganda, a Bantu language (the particular example was called to my tion by Frank Heny). It is one of a series of several "tenses"; they have received a number of analyses in the literature, always essentially in accentual terms, though never, to my knowledge, within a reasonably constrained accentual theory. We shall propose a non-accentual solution. (See 49).)

This is clearly a rather complex series of tonal forms and one might expect that the rules mediating the underlying and the surface forms would need to be rather complex. We shall suggest that a simple solution exists, involving one essentially morphological rule, and a melody association rule.

First, the relevant morphemes. On the syllabic tier, they appear, in the following order:

(47)
$\begin{Bmatrix} a \\ e \end{Bmatrix}$ te $\begin{Bmatrix} y \\ ba \end{Bmatrix}$ ta   a   gu   Verb

rel mar.	Neg M.C.	subj mark	relat. clause neg	far past marker	object marker

We assign the following tones to each morpheme:

(48)   a   te $\begin{Bmatrix} y & ba \\ L & / & H \end{Bmatrix}$ ta   a   gu   Verb
       e   (see text)
                                                                            HH(H)
       L   L   H   H   L

(49) CV

(A) 'Grind' /sa/

(i) yàsâ   yàgúsâ   tèyàsâ   tèyàgúsâ

    he   he ground   he didn't   he didn't
    ground   it   grind   grind it

(ii) báásà   báágúsâ   tèbáásà   tèbáágúsâ

    they   they   they didn't   they didn't
    ground   ground it   grind   grind it

(iii) èyàsâ   èyàgúsâ   àtáásà   àtáágúsâ

    he who   he who   he who   he who
    ground   ground it   didn't grind   didn't grind it

(iv) àbáásà   àbáágúsâ   àbátáásà   àbátáágúsâ

    they who   they who   they who   they who didn't
    ground   ground it   didn't grind   grind it

(B)   Iya 'eat'

(i)   yàIyâ   yàgúIyà   tèyàIyâ   tèyàgúIyà

(ii) báályà   báágùIyǎ   tèbáályà   tèbáágùlyà

(iii) èyàlyâ   èyàgúIyà   àtáályà   àtáágùlyà

(iv) àbáályà   àbáágùlyà   àbátáályà   àbátáágùlyà

CVCV

(A)   gula 'buy'

(i)   yàgúlà   yàgúgúlâ   tèyàgúlà   tèyàgúgúlâ

(ii) báágùlà   báágúgúlâ   tèbáágùlà   tèbáágúgúlâ

(iii) èyàgúlà   èyàgúgúlâ   àtáágùlà   àtáágúgúlâ

(iv) àbáágùlà   àbáágúgúlâ   àbátáágùlà   àbátáágúgúlâ

(B)   Iaba 'see'

(i)   yàlábà   yàgúlàbà   tèyàlábà   tèyàgúlàbà

(ii) báálàbà   báágùlàbà   tèbáálàbà   tèbáágùlàbà

(iii) èyàlábà   èyàgúlàbà   àtáálàbà   àtáágùlàbà

(iv) àbáálàbà   àbáágùlàbà   àbátáálàbà   àbátáágùlàbà

CVVCV

(A)   kweeka 'hide'

(i)   yàkwéékà   yàgúkwéékà   tèyàkwéékà   tèyàgúkwéékà

(ii) báákwèèkà   báágúkwéékà   tèbáákwèèkà   tèbáágúkwéékà

(iii) èyàkwéékà   èyàgúkwéékà   àtáákwèèkà   àtáágúkwéékà

(iv) àbáákwèèkà   àbáágúkwéékà   àbátáákwèèkà   àbátáágúkwéékà

(B)   leeta 'bring'

(i)   yàIéètà   yàgúlèètà   tèyàléètà   tèyàgúlèètà

(ii) báálèètà   báágùlèètà   tèbáálèètà   tèbáágùlèètà

(iii)èyàléètà   èyàgúlèètà   àtáálèètà   àtáágùlèètà

(iv) àbáálèètà   abáágùlèètà   àbátáálèètà   àbátáágùlèètà

First note that (48) is not written with asociation lines; these are meant as lexical entries, before any association rules have applied. Thus "te/L" describes two, as yet unassociated, sets of segments.

The tone assigned to gu requires a further statement. The difference between the "A" verbs and the "B" forms of (49) are attributed, under this analysis, to gu's tones being deleted when a B-verb follows (which will, note, be immediately following). Furthermore, in the other case, of an A-type verb, gu has the tone HHH in relative clauses or with plural subject; elsewhere it is HH in tone. This relative complex specification appears to arise from pressure towards paradigm parallelism (observe the point at which the tone falls in CVCV and CVVCV verbs A-type with gu). Rather than go through all the examples, I shall rest content with some illustrative cases.

(50)   y a gu la

        L H   L

(50) is an underlying form; no associations are present yet. If we associate from left to right, one-to-one--(50) becomes (51); note that y is not syllabic, and so does not associate with a tone, even though there is a tone in the lexical entry for this morpheme.

(51)   y a gu la
          ／ | | yàgúlà
        L   H L

(52) works the same way.

(52)   a ba ta a kweeka   a ba ta a kweeka

        L H H H L   →   L H H H   L

If there are syllables or tones left over after associating one-to-one from the left, the Well-formedness Condition effects their association, as in (53).

(53)   a ba ta a gu gula   a ba ta a gu gula
                     → | | | | | | |
        L H H H HHH L   L H H H H H HL

By WFC   a ba ta a gu gula = àbátáágúgúlâ
     → | | | | | | |\
            L H H H H H HL

This particular analysis is of some interest because it violates the intuition one might have had, upon observing the initial data in (49). There we would observe that all the forms bore the tonal "envelope" L H L; all that varied among the examples was where the tone rose to H, and where it fell. Fixed on this, several analyses have been offered, but all are quite complex, and do not bear out the value of the initial intuition that a single melody LHL is involved.

Within the theory developed so far, such an accentual approach would not be stateable, since both the H and the L would be star-linked, presumably, and we have not needed to have recourse so far to such apparatus. In sum, what may appear to the linguist to be a single accentual melody, as in (49), may turn out to be quite different.

This analysis further requires a way to state formally the procedure: associate tonemes and vowels one-to-one starting from the left. In Goldsmith (1975a), I proposed the following algorithm; although the arguments given there I no longer find very persuasive, it is the best formalization I am aware of for this procedure. To formalize the "one-to-one" mapping, we shall appropriate the notion of rule schema as set forth in Vergnaud and Halle's work on variables in phonology.

We set up the phonological segments at the top of a sequence of columns, and the tonological segments at the left of a sequence of rows, like (D).

(D) $i \rightarrow$

		CV	CV	CV	CV	CV
j	L	✓				
↓	H		✓			
	L			✓		

The assignment rule then takes the form of (E)

(E) $V \simeq T \,/\, \# \, X \, \dfrac{V}{T}$ (where "$\simeq$" means "is associated with")

This is to be interpreted in what I take to be the natural way, given the way Vergnaud and Halle's "real" variables work: we produce an algorithm that checks off a certain number of the boxes in the matrix (D). These boxes then represent pairs of vowels and tonemes (column and row), and we apply the structural change of the rule to these pairs of toneme and phoneme, which means in this case to hook them together.

In particular, the algorithm is to check off the box (i,j) (think of the induces as coordinates as indicated in (D)) where (i,j) is the minimal pair such that $i > i'$ and $j > j'$ for all other check-off boxes of the form (i',j'). This algorithm is reactivated successively until it cannot do anything anymore (Goldsmith,1975a)).

Whether this is correct or not, however, does not force the question considered earlier as to whether Etung or Tiv utilize a rule like (39) or like the one quoted in the passage immediately above with a variable.

5.   Tonga
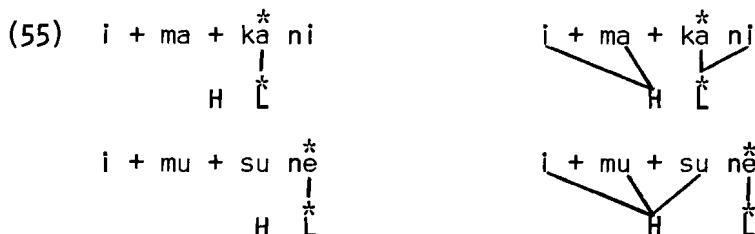
In his article "Some Tonga Tone Rules," James McCawley deals with some problems in the tonal system of Tonga, a Bantu language. In most analyses of Tonga and related languages--McCawley cites Meeussen and Stevick, for example-- the tonal characteristics of particular morphemes or syllables has seemed elusive. One couldn't simply report that such and such a morpheme was High-toned, or Low-toned--whatever inherent tonal characteristics it had seemed to show up elsewhere, frequently on the following syllable. Indeed, McCawley says, "In these languages it is often far from obvious which member of an underlying tone contrast should be designated 'underlying high' and 'underlying low.' Many scholars...have avoided such terms, preferring DETERMINANT and NEUTRAL, here determinant is that underlying tone which (under at least some conditions) causes derivation from the "least marked" tonal contour (which is generally level low pitch)" (p. 140).

Proceeding from the assumption that the Determinant/ Neutral contrast must correspond to either a High/Low or Low/High contrast, McCawley develops a system of tone rules for Tonga, although he recognizes from the start that it is nigh onto arbitrary whether we choose Determinants to be under- lyingly High or underlyingly Low. Although historically the determinant morphemes derive from High tone, McCawley opts in favor of a synchronic analysis which analyzes Meeusen's "derterminants" as underlyingly Low; this choice is made so that global rules can be avoided. At the end of the article, McCawley notes another analysis might be possible which was based essentially on accent or pitch-fall.

In an unpublished paper (1974), Michael Cohen has developed an analysis of Tonga tone which follows this second, accentual, approach, using an accentual autosegmental frame- work. Because his work is not generally accessible, and is highly instructive regarding the functioning of a complex accentual system, I shall review a part of his findings in

some detail. In some of the conclusions drawn I differ slightly
from Cohen, but I shall note explicitly where any parts of
the analysis below are mine and not Cohen's.

Cohen's first tack is to attribute the Determinant/Natural
distinction not to a tonal High/Low distinction but rather to
equate it with an Accented/Non-Accented distinction. Thus
Determinant morphemes will have a star on them. The neutral
tone melody for Tonga, furthermore is H L̃. Thus we get (55).

(55)
i + ma + ká ni

H L̃

i + mu + su né

H L̃
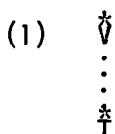
i = noun prefix

ma, mu = class marker

kani = 'news,' sune = 'ox'

Cohen's own analysis of the major association rule is
similar to that described in Williams (1972) and Leben (1973),
but I shall proceed using the principles developed in the
present chapter. Thus the first stage in the two derivations
in (55) results from the application of rule (1); the Well-
formedness Condition unambiguously results in the correct
output in the right.

(1)   V̇
      .
      .
      .
      T̃
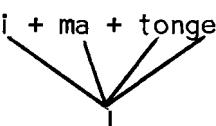
Noun stems that Cohen cites are mono- or bi-syllablic, and
are commonly as illustrated above in (55). McCawley (p.150)
gives trisyllable forms that work according to the same
pattern. A small class of monosyllable nouns can be pronoun-
ced either as in (55), or with an alternate tonal melody, in
which all the syllables are High. For this class we shall
say there is the optional melody variant, H̃. See (56).

(56)
i + ma + li í   variant of   i + ma + li í

H̃   H L̃

li = 'money'; ma (concord)

There is a further set of nouns whose tone follows neither the pattern in (55)--accented, with the melody H Ĺ--nor that in (56)--accented with the melody Ĥ. This class is pronounced in isolation with all syllables Low-toned, and is analyzed by Cohen as unaccented. Some process--and about this there is here little detail to specify, unfortunately--will insert a low-tone for a word totally devoid of tonemes. See (57).

(57) tonge 'ruins'   i + ma + tonge

                              \ | /
                               \ | /
                                \|/
                                 V
                                 L
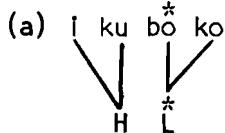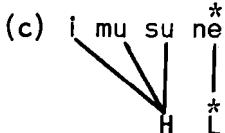
The accentless nature of these forms is illustrated by the paradigm in (58) (a modification of Cohen's (4)). See especially (58e,f). Cohen says:

> Neglecting the second and only considering the behavior of the first word, we see its behavior is quite regular. High tone spreads over the first word until the first accented syllable is encountered. If none is encountered, high tone spreads over the whole word.

(58)   isolation form   genitive

(a) i ku bŏ̌ ko   (b) í kú bŏ̌ kó wá mu kǎ ìn tù

       \ | \ /   \ | \ / /\ \|/
        \ | V   \| V / \ V
         V | H   L H   L
         H L   H   L

     boko 'arm'   'arm   of woman'

(c) i mu su nĕ̌   (d) ímú sú nĕ̌ wá mu nĕ̌ ně̌

       \ \ | | \ \ | | | /\ \/
        \ \| | \ \| | | / \ V
         V | V | |
         H   L   H   L H L H L

       'ox'   'ox   of old man'

(e) i mu li mo   (f) [mú lí mó wǎ̌ mú sím bǐ̌

       \ | /   \ \ \ / | \/ |
        \ | /   \ \ \/ | V |
         \|/   \ \ / |
          V   \ \/ | H L
          L   V |
   H   L   H L
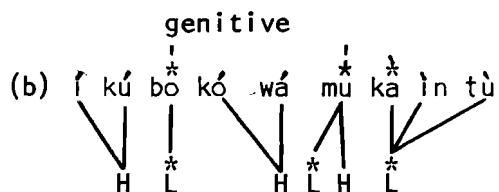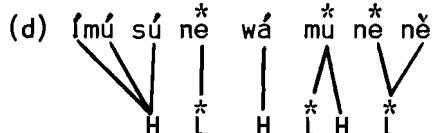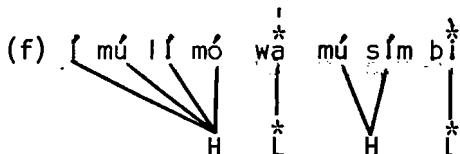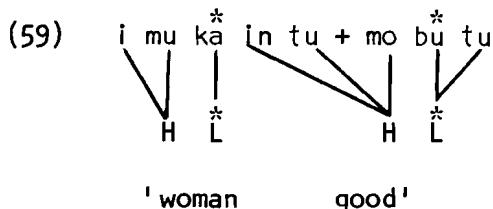
   'work'   'work   of girl'
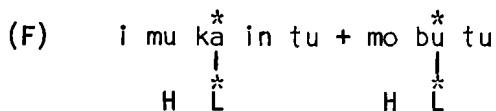
(Note that the derived genitive forms are intermediate forms and not final forms. I have noted the surface forms with diacritic accent marks over the syllables; they will not

correspond exactly to the autosegmental notation because they
result from the application of further rules. We shall return
below to the tonal characteristic of wa, the possessive marker.)
Although the forms in (58) involve some complications, Cohen's
primary point--that the nouns that display Low tone in isolation
are accentless--comes through clearly upon comparison of the
tone in (58b, d, and f), or comparison of (58e) and (58f).
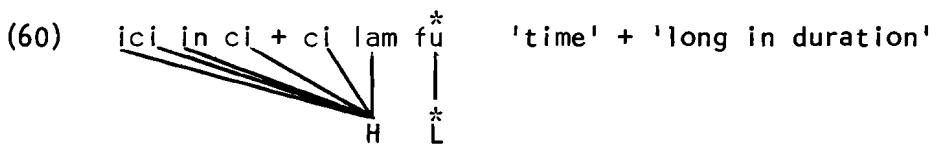
The same argument applies to compounds, Cohen argues. The
facts are much as in (58) except that there is no connecting
particle parallel to the wa in (58). Thus an accented word
like ímúkáìntù has an H L contour in isolation; in a compound,
it is affected as in (59), and uttered ultimately ímúkáìntú
móbùtù.

(59)   i mu ká in tu + mo bu tu

         H L   H L

'woman   good'
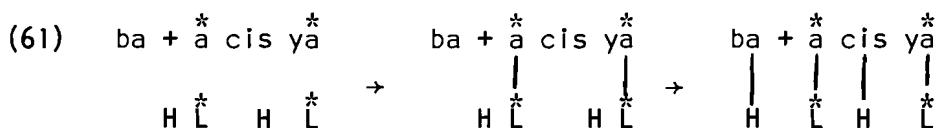
Note that this spreading of the H-tone follows automatically
from the Well-formedness Condition if unassociated tones
have precedence in spreading over already-associated tones.
Thus (59) derives from (F) directly.

(F)   i mu ká in tu + mo bu tu

        H L   H L

I have no independent justification, I might add, for the nature
of the boundary in (F). Accentless forms. like ìcììndì (an
isolation form) are all H in compound; see (60).

(60)   ici in ci + ci lam fú   'time' + 'long in duration'

                  H L
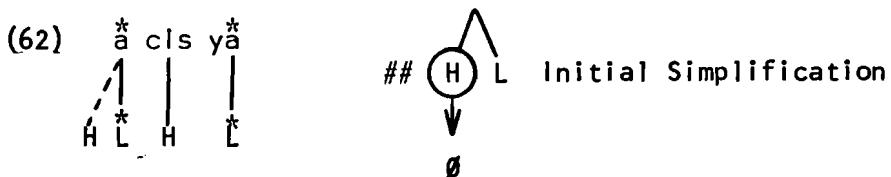
In another connection, Cohen notes another class of
exceptional nouns in Tonga. Rather than being exceptional by
having no accent, as in the preceding forms, this next class
is exceptional in that words in this class have two accents.
Corresponding to each accent is an instance of the tone melody
HL. Thus we get forms like (61).

(61)   ba + á cis yá   ba + á cis yá   ba + á cis yá

                    →   →

    H L   H L   H L   H L   H   L H   L
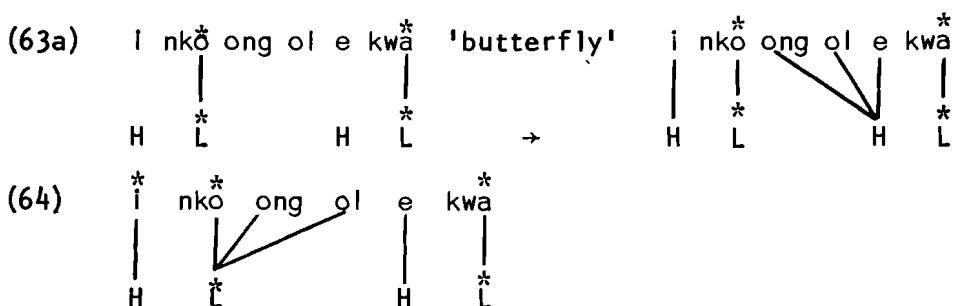
The final form (61c) is later subject to a rule that
realized a single L-toned syllable surrounded by H-toned
syllables as a downstep; thus (61c) is uttered bá à cís yà.[5]

Even if the prefix ba is not present, as it is in (61),
we get a Low tone on the first syllable of à cís yà; an
initial ĤL (falling) tone is simplified by rule to L tone.

(62)   å cís yå

(H) L   Initial Simplification

In Cohen's analysis this deletion of H results from a
more general provision of the major association rule involved
here. The difference is not crucial here.

Pursuing this, Cohen cites the following "polysyllablic
nouns," not indicating any morphological division. Whether or
not there are any morphological boundaries internal to the
form is essentially irrelevant, however, to the point that
the conditions we have lately suggested for the Well-formed-
ness Condition (statement (46) above, or equivalently here
statement (26)) provide for the reassociation between (63a)
and (63b) automatically. (64) is not generated by the Well-
formedness Condition, correctly enough, and all this is
accomplished without recourse to talk about "left-to-right" or
"right-to-left" spreading.

(63a)   I nkŏ ong ol e kwå   'butterfly'   i nkŏ ong ol e kwå

(64)   i nkŏ ong ol e kwå

Cohen's analysis, as I alluded to above, for these forms
depends on a specific right-to-left conception of tone-mapping
rather than the interpretation I am giving.

Cohen proceeds in his analysis from possible noun tone
patterns to the tone patterns displayed by affixes. Perhaps
the most interesting class has the following three character-
istics:

"1.   They are always high even if not followed by an
accented syllable.
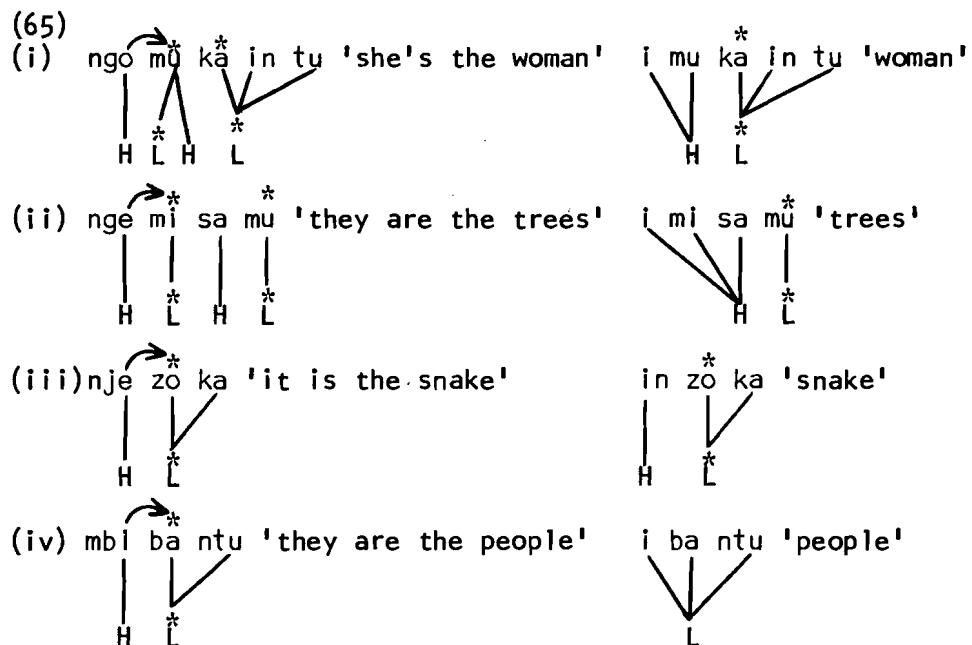
2. If the next syllable would have been realized as a high, then the following syllable becomes a slipped high.

3. Otherwise, the following syllable becomes a low."
    (p. 8)

Cohen continues, "segments with this behavior seem to occur more frequently in Tonga than other exceptional seg-ments." Cohen's own analysis for these forms is that these ⃰ are forms accented and individually specified for the melody HĿ.

We shall propose, to the contrary, that the melody here is H Ĺ, but that these forms are <u>post-accenting</u> forms: that, under specifiable conditions, they place a star on the following syllable prior to the melodic association rule.
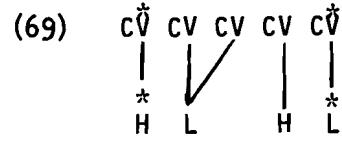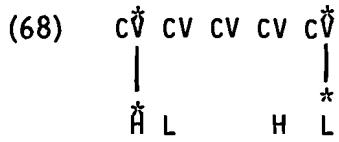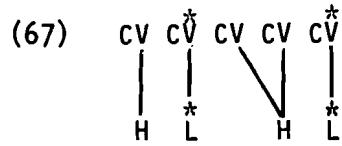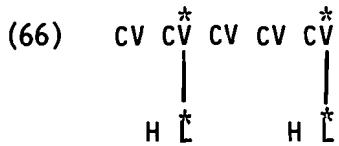
Examples of this behavior are as in (65) = Cohen's (15), slightly modified.

(65)
(i)   ngó mȕ kằ in tu 'she's the woman'   i mu ka in tu 'woman'
       H L H L   H L

(ii)   ngé mì sa mȕ 'they are the trees'   i mi sa mȕ 'trees'
       H L H L   H L

(iii) njé zo ka 'it is the snake'   in zo ka 'snake'
       H L   H L

(iv)   mbí ba ntu 'they are the people'   i ba ntu 'people'
       H L   L

Post-accenting morphemes: (i) ngo; (ii) nge; (iii) nje; (iv) mbi.
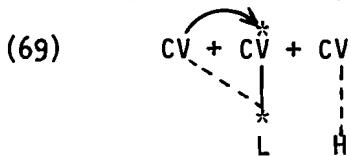
Note that by calling these elements "post-accenting," the theory makes predictions for structures of the form (66); these should become by the Well-formedness Condition (67). If instead of (66) we had the structure (68), the Well-formed-ness Condition would make no prediction as to whether (67), or, say (69).

(66) CV CV̊ CV CV CV̊
      | |
      H L̊   H L̊

(67) CV CV̊ CV CV CV̊
      | | \| |
      H L̊   H L̊

(68) CV̊ CV CV CV CV̊
 | |
      H̊ L   H L̊

(69) CV̊ CV CV CV CV̊
      | /| |
      L̊ H L   H L̊

I do not have examples that meet these specifications, but
they can certainly be found, either in Tonga or in related
languages with similar systems, such as Bangu-Bangu, another
Bantu language, where the underlying tone melody is L H.

The notion "post-accenting" should be a basic one in
any adequate theory of accentuation. We would expect that the
post-accenting property of morphemes, for example, should
be maintained despite a shift in the underlying tone melody.
Thus if HL is the basic tone melody, a post-accenting morpheme
appears to bear an H tone and put a Low-tone on the following
syllable, The same facts would be stateable within a theory
which attributed an exceptional melody H̊L to that morpheme.
The post-accenting theory predicts, however, that should the
normal tone-melody shift to, say, L̊H, post-accenting morphemes
will suddenly no longer look like they did before, but will
rather put a Low-tone on the following syllable and a High
tone on the syllable two syllables to the right, as in (69).

(69)   CV + CV̊ + CV
              | |
              | |
              L̊   H

Although the data there is not sufficient to provide a com-
plete analysis, McCawley's discussion of Bangu-Bangu ("Global
Rules and Bangu-Bangu Tone") suggest very strongly that this
is precisely what has occurred there. The question requires,
of course, further investigation.
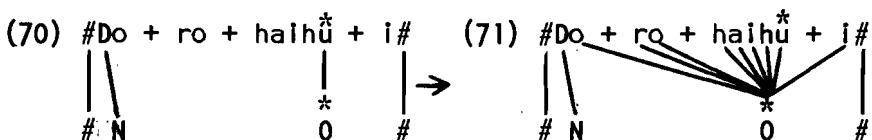
6.   Further Non-Accentual Questions

In the past few sections, we have been trying to separate
out the functioning of the Well-formedness Condition--the
universal contribution of the autosegmental theory--from the
language particular major association rules. The attempt has
been to keep the language-particular rule as simple as possible--
or rather to show that this, the best result *a priori*, is
consistent with a wide range of facts. We have worked with an
interpretation of the Well-formedness Condition (46) (repeated
here):
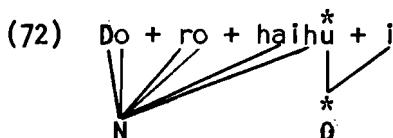
(46) Given an ambiguity in ways to fulfill the
      Well-formedness Condition, do not reassociate
      an already associated segment.

This formulation appears to supercede (26) above, which
gives precedence to non-accented segments in spreading due
to the Well-formedness Condition.
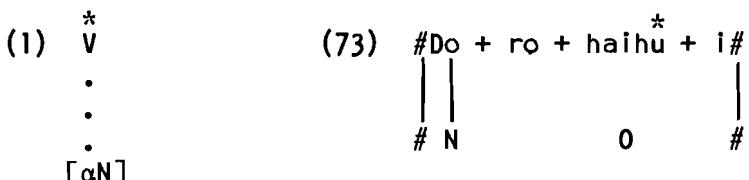
Condition (46) explains what was left unexplained in the
discussion of Guarani in Chapter One, Section 6. There we
derived, for example, (71) from (70).

(70) #Do + ro + haihũ + i#   (71) #Do + ro + haihũ + i#

The association of the nasal segment, we asumed, was listed
in the lexicon, while the association of the starred Oral
melody 0 was rule-governed. The association effected by the
Well-formedness Condition in (71) appears to violate form (26)
of the Condition, which gives precedence in spreading to non-
starred segments. (26) would predict (72) rather than the
correct (71).

(72) Do + ro + haihũ + i

However, if we adopt the more general condition (46) rather
than (26), we see that (71) derives not from (70), which has
putatively undergone the effects of a major-association rule
like (1), but rather has undergone no major association rule.
Thus the stage preceding (71) directly is not (70) but (73).

(1) $\overset{*}{V}$ (73) #Do + ro + haihũ + i#

     •
     •
     •
     [αN]

(modified)

The correct output results, then, from an even simpler
system than that suggested in Chapter One: there is, in fact,
no major association rule in Guarani.

One might ask how forms like (76b) of Chapter One, are
derived. That is, the question might arise how it is insured
that the 0 melody starts off to the right of the N autosegment
if there is no major association rule to govern where the
autosegment is.

This is a fair question, but it is not relevent to the
question as to whether there is a melodic association rule.
Even in the analysis given in Chapter One some means was
necessary to place the Oral or Nasal Melody in the proper place
in the nasality tier, and it was not the melodic association
rule that determined the structure of that tier. Some steps
toward a deeper understanding of this question arise from
G. N. Clement's analysis of vowel harmony, where the simple
answer is given that the melody is entered in the lexicon as
such. We turn to this now.

Vowel Harmony

In its simplest form, vowel harmony is a phenomenon
wherein the vowels of a language can be divided into two
classes, and where within a certain domain--generally the word--
all of the vowels must be chosen from one of these harmony
classes or the other (it may happen that a vowel, like /i/,
occurs in both classes--it is then called "neutral").
In addition to merely constraining the form of possible under-
lying forms to those containing vowels entirely from one
class or the other, vowel harmony also determines the surface
form that an affix will take, When an affix is joined to a
stem, its harmony class is generally determined by the harmony
class of the stem; thus, depending on the particular stem that
the affix is joined to, the affix will have one of two possible
phonetic forms.

For example, there is a "tenseness" harmony in Igbo.
Adopting the feature "Advanced Tongue Root" (ATR), this har-
mony system, we may say, insures that all vowels within a
certain domain of the Igbo word agree in the value they sustain
for the feature "ATR." There are four vowels marked +ATR: i,
e, u, and o; and four marked –ATR: i̟, a, u̟, and o̟ (these
correspond to the more traditional "tense" and "lax" groups,
respectively). Thus /mádhù̟/, /ǹòchàrà/ (lax), and /óchíè/,
/ógbù/ (tense) are possible (indeed, existent) word in Igbo,
whereas */madhu/, for example, is impossible.

We noted above, in Chapter Two, that the prefix /a/
that marks the II Incomplete form will appear as /e/ if the
verb stem to which it is attached begins with a tense (=+ATR)
vowel. In general, we can find affixes--prefixes and suffixes
--which illustrate this kind of vowel alternation. Thus,
/kwa/, a suffix, has the form /kwa/ after a lax stem, and the
form /kwe/ after a tense stem. Similarly, /ghi̟/ may also
appear as /ghi/, depending on whether the verb stem is marked
+ or –ATR (depending, as is said, on the "harmony class" of
the stem).

For example, with the tense stem /le/ we get the form /hwe/, as in Lèkwé yà; (see (I)), with the lax stem /dị/, we get dịkwá (see (ii)).

(i) Lè + kwé yá !

      look suf- it   "Look at it!"
      at   fix

(ii)] dị kwà mmá

      you are suffix well   "Are you well?" ('How are you?')

Viewed from the standpoint of vowel harmony, this alternation of the suffix /kwa/ and /kwe/ is caused, we may say, to preserve the harmony of the vowels in the domain governed by the stem vowel. Alternatively, we might view the vowel in this suffix as being an "archivowel," in that it has no specification for the ATR feature underlyingly, and receives this specification from the verb stem.

G. N. Clements has approached the phenomenon of vowel harmony from an autosegmental point of view. He suggests, following Ultan, five properties distinguishing vowel harmony systems:

(i) a phonological feature, or set of features, is determined by the root or stem (with certain qualifications we will consider in a moment);

(ii) the affixes on this root then agree with the stem, and the domain over which the agreement holds is unbounded, that is, not limited to one or two affixes;

(iii) the harmony occurs both to the left and to the right of the stem; it is, that is, bidirectional;

(iv) vowel harmony is not optional;
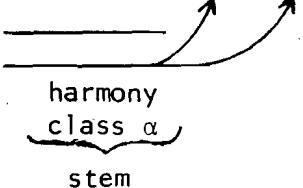
(v) The phenomenon of vowel harmony is "phonetically natural," in that the feature which appears to spread over the harmony domain is phonetically and phonologically motivated, from among the class: back, high, round, advanced tongue root, constricted pharynx.

Clements proceeds with the assumption that the mettle of a theory of vowel harmony is ultimately tested not by how it deals with the non-exceptional case--for the non-exceptional case can be treated, to some extent or other, by several formal approaches, including the "features spreading" conventions (Lightner, Chomsky and Halle), iterative rule application, or general variables (Vergnaud and Halle); rather, the test

comes from the way exceptional cases are treated, and how the exceptional cases are related to the regular forms.

Clements argues that autosegmental theory provides an e tremely nat al way to subsume both the normal case of vowel harmony and the exception "opaque" vowel, the vowel which does not "go along with" the harmony specification on those vowels

(74)   (a)   CV̆ CV̆ CV̆   + CV̆ + CV̆   normal case

harmony
class α
stem
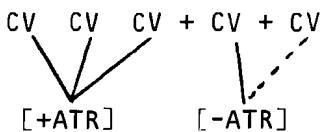
(b)   CV̆ CV̆ CV̆ + CV̆ + CV̆   opaque vowel
                           opaque case

harmony   harmony
class α   class β

This opaqueness, Clements suggests, arises from a vowel being underlyingly associated with a harmony autosegment. Non-exceptional vowel harmony behavior is the surface manifestation of vowels underlyingly unassociated with any harmony autosegmental representation.

(75)   CV CV CV + CV + CV

[+ATR]

CV CV CV + CV + CV

[+ATR]   [-ATR]
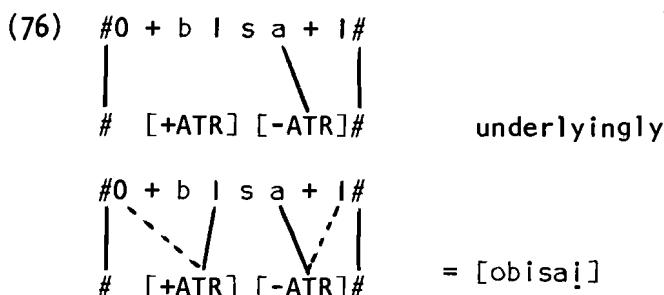
The notion of "deviance" or "amount of exceptionality" of a form then becomes easily quantifiable, for it is, as a first approximation, the number of vowels underlyingly associated with harmony autosegments.
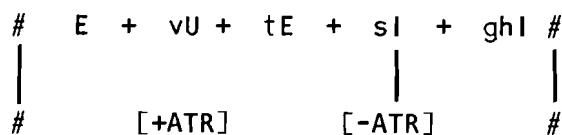
We may remove the provision of "first approximation" if we note that there are some "exceptions" which are, indeed, predictable. Thus, Clements says, the Advanced Tongue Root (ATR) harmony in Akan is consistently blocked by the vowel /a/. "We assume, then," he writes, "that in the phonology of Akan there is a statement to the effect that every lexical occurrence of the low vowel is bound to an occurrence of the feature -ATR." Thus the form /biṣa/, with a +ATR first syllable, causes prefixes to become +ATR, but the second

vowel, /a/, blocks this from spreading to the suffixes; rather, the suffixes become -ATR, as illustrated in (76).

(76)
```
 #0 + b I s a + I#
 |         \   |
 |          \  |
 #  [+ATR] [-ATR]#        underlyingly
```

```
 #0 + b I s a + I#
 | \   /   \ / /
 |  \ /     V /
 |   V      V/
 #  [+ATR] [-ATR]#     = [obisai]
```
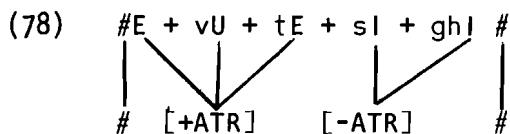
There may be, furthermore, suffixes in a system like that of Igbo described above, whose vowel harmony specification is not determined by the stem, but is, rather, specified underlyingly --once and for all--by the suffix. Such a suffix would not alternate in form depending on context; such a suffix would additionally block vowel harmony from extending past it, by the Well-formedness Condition. Clements pointed out that his is the kind of behavior we do in fact find in a suffix like /si/ in Igbo:

(77)   prefix stem suffix suffix suffix

```
 #   E  +  vU +  tE  +  sI  +  ghI #
 |                     |          |
 |                     |          |
 #       [+ATR]       [-ATR]      #
```

The captial letters stand for vowels unspecified for the feature Advanced Tongue Root; the second level is an autosegmental level consisting of the feature Advanced Tongue Root. The Well-formedness Condition changes (77) to (78).

(78)
```
 #E + vU + tE + sI + ghI #
 |\  |  /      |   / |
 | \ | /       |  /  |
 |  \|/        | /   |
 #  [+ATR]    [-ATR]  #
```

This is the correct output, in fact: /evutesighi/.

We shall not consider in any further detail the autosegmental treatment of vowel harmony; the reader is referred to Clements' work, cited in the bibliography. Two points are relevant to us, though: first, a clear account of certain vowel harmony systems can be obtained through an autosegmental treatment in which the autosegmental level corresponds to the feature being harmonized for; and second, such an account is able to account for exceptions of vowel harmony without recourse to any notion of "rule feature" or "rule exception"; what makes "exceptional" segments that, within this framework,

is additional underlying specification.

7.   Conclusions

In Chapter Two, we used the tonal system of Igbo to get
a clearer idea of how an autosegmental system operates.   In the
the present chapter, we have used autosegmental techniques
to gain new insights into traditional questions, beginning
with the nature of stress subordination, and proceeding to
other questions of accentuation and the nature of underlying
forms.   By attempting to make the mapping of tone melodies
to syllable structures as simple as possible in any particular
language, we were led, step by step, to a view where the very
simplest system had precisely the properties of the well-known
phenomenon of vowel harmony, in the case where the autosegmental
level was the level corresponding to the features being harmon-
ized for (advanced tongue root, backness, and so on).

What we have is insidious autosegmentalism.

When we began this study, we seemed to be showing that
certain paradoxes for the standard segmental theory that tonal
systems manifested were resolvable by making a slight change
in our assumptions of the "geometrical shape" of underlying
forms.   What we are finding now is that this modification
extends, at least potentially, to virtually all the familiar
features, not just the tonal ones: nasality, backness, ATR,
rounding, and height all can participate in vowel harmony sys-
tems.   Why should this be?   Why doesn't autosegmentalism stick
to the tonal domain?   Why does it arise in the first place?
This is the subject of the final chapter, Chapter Four.

FOOTNOTES FOR CHAPTER THREE

1. The caveat comes from Mark Liberman's (1975) discussion of of the spreading of tonemes after the melodic association rules of English. He argues that in spoken English, a form like (i) (his 3.1/10) does not become (ii); rather, the shift from the L toneme on *spec* rises gradually to the High toneme on *cau.*

(i) *especially elaborate precautions*
     | | |
     L   H L

(ii) *especially elaborate precautions*
     \\\///_____ | |
      \\V/_____ | |
       L   H L

Such a structure as (ii), however, would be correct, he continues (p. 102) for changed English.

> "Thus we may attribute the following property
> to chants, but not to unchanted intonation
> in English: <u>Every syllable must be associated
> with some tone.</u>"

The other conditions of the Well-formedness Condition, as presented here, are explicitly adopted on p. 120:

3.2.3/1a All tones must be associated with some syllable.

1b Association lines may not cross.

Liberman is clearly right in his analysis of (i) and (ii)--the full Well-formedness Condition would be too strong, and thus empirically wrong, in requiring (i) over (ii). This raises serious questions, but of detail, I believe, rather than principle. As Liberman himself says, in considering the relation of speech and chant, "It is hardly necessary to point out that the introduction of unchanged speech...into a primarily changed song-structure means that the child is prepared to accept speech as the (so to speak) aesthetic equivalent of the chant. We will argue shortly that this equivalence is in some ways a very deep one" (p. 44). I take it that his reference here (to "shortly") is to the citation above from his p. 102. The relation between chant and speech remains, then, a baffling one, if chant should be a mode that imposes a three-part Well-formedness Condition, and speech a mode that imposes only a two-part Condition, leaving open whether syllables associate with a toneme. Furthermore,

languages like Igbo, which we considered in detail in Chapter
Two, obey the "chanting" principles--why? I believe that an
answer may emerge out of the considerations in Chapter Four.
For the moment, observe that in all of the discussion so far--
except in the Introduction--not even passing reference has
been made to the fact that consonants intervening between
the vowels, or more generally, syllabics, do not associate
with tonemes. There is thus an important sense of "opaqueness
to the Well-formedness Condition" that has been implicitly
brought in from the beginning. The provision of the Well-
formedness Condition might read, more exactly:

> All segments which are tone-bearing (by some criterion
> independent of the structure at hand) must be associated
> with some toneme.

This notion of "tone-bearing" is, clearly, the operative
one, although we won't have much to say about it here.
In general, however, the string of segments on ther upper,
syllabic tier, appears to partition into disjoint domains,
each containing exactly one tone-bearing segment; the
other segments are then unspecified for the feature on
the other autosegmental level. We shall have more to say
about unspecifiedness in Chapter Four.

2. Morris Halle points out that there may well be a
   generalization here to be observed, if instead of listing
   the melodies as in (i), we list them as in (ii).

	i.		ii.	
	H		HHH	
	HL		HLL	
	LH		LHH	
	L		LLL	
	LHL		LHL	
	LLH		LLH	
	HHL		HHL	
	HLH		LHL	

The generalization is then clear: all melodies expres-
sible as $T_1 T_2 T_3$ appear in (ii). Whether this is a
linguistically real generalization, we can only conjecture
at this point; the observation, however, is striking.

3. The exceptions I know of are borrowings.

4. We ignore here what Liberman calls "boundary tones,"
   which are irrelevant since they are not free, unassocia-
   ted tones.

5. The fact that a single noun may be associated with two
   different copies of the H L melody distinguishes
   Tonga from Indo-European and Japanese systems, where the
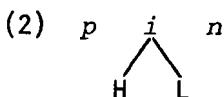
domain of a single tone melody is always at least as
large as a single word. Thus if a single word contains
more than one accent, some language-paricular principle
must select which accent associates with the starred toneme
in the tone melody. Tonga, as we see, is not restricted
in this way; in fact, there are as many copies of the
tone melody H L inserted as there are accents. For dis-
cussion of the Indo-European system, see Halle (1975),
Kiparsky (1973), for the Japanese, see Haraguchi (1975).

CHAPTER FOUR

THE AUTOSEGMENTAL INDEX

We have managed, with only a momentary lapse or two, to overlook a certain basic question for most of the development so far: just _why_ are there autosegmental tiers? Not all questions have answers, of course, and why-questions are notoriously the worst in this regard. Nonetheless, we shall pose this question, and answer it, after a fashion, with a hypothesis.

This hypothesis harks back to the discussion at the beginning of Chapter One that likened the act of speech to an orchestral production, a harmonized score of several independent musical "voices." These "voices" of the analogy correspond to the distinctive features of phonology. At the most superficial level, the speech signal is broken down into a large number of independent linear parts--autosegmental tiers--with at least as many of these tiers as there are independent articulators. Thus there will minimally be such a tier for the velum, for the laryngeal gesture corresponding to pitch, and so forth. In short, the phonetic level looks like the orchestration sketched in (3) at the beginning of Chapter One, reproduced here as (1). As before, we shall call the number of autosegmental tiers in a particular representation the _autosegmental index_ of that representation. Thus the autosegmental index of (2) is 2.

(1)　Lips　　　....Close up....Open.........................
　　　Tongue　....High and front............touch the palate
　　　Velum　　...Raise............Lower....................
　　　Larynx　...High Pitch....Low Pitch...................

(2)　p　　i　　n

Yet even if (1) is an adequate sketch of the phonetic representation, it is unquestionably the wrong picture for the more abstract--that is, psychologically real--level, where there are atomic units like /p/, /i/, and so forth--in a word, phonemes. The overwhelming weight of phonological work makes it for us axiomatic that the phonological system of a language is composed of a small number of atomic segments; each segment is defined by several feature-specifications, but these segments act, in general, like discrete, atomic units. Thus an /i/--a front, high syllabic segment--may be inserted epenthetically; or an /n/, a coronal, non-syllabic nasal, may be deleted. Such operations insert or delete whole units comprising more than one feature. The assumption that such things may occur I shall call the _Abstract Segment Hypothesis_.
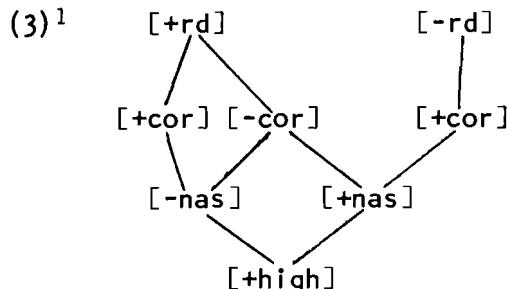
It is intimately connected, it can be seen, with the (Absolute) Slicing Hypothesis of Chapter One.

In a sense, it is ironic to try to make this point here. These atomic segments are what makes phonology possible (especially a phonology which contains rules relating stages in a derivation), and any alternative which denies it has a very weighty burden of proof resting on it. But if one adopts autosegmental theory, then although the Abstract Segment Hypothesis is not thereby put in doubt, the question of why it is true arises. Permit me to emphasize this: without an autosegmental theory, the idea of segments (whether composed of feature specifications or not) is an untestable, virtually unquestionable part of phonology. With the introduction of autosegmental tiers, one may view the abstract segment hypothesis as a limiting case of autosegmental representation. The Abstract Segment Hypothesis says in effect that there will necessarily be a very small number of autosegmental tiers at the most abstact level, and in the case where the autoseg- mental index is 1, all the features will be on one tier. But even if the autosegmental index is 2 or 3, the number of features on some tier will be relatively large.

And we may ask why this is so. Why, that is, do feature- specifications cluster together at the deepest level to form segments incorporating many feature-specifications? This is our question--how, and why.

Our hypothesis shall be that the peripheral devices of the articulatory apparatus--and correspondingly, of the receiv- ing end--are essentially parallel processing devices, as the orchestral score image suggests.

A representation of the phonetic level, under such a con- ception, is extremely rich in information, and correspondingly poor in redundancy. A low-level representation might look like the one in (3).

(3)[1]

[+rd]   [-rd]

[+cor] [-cor]   [+cor]

[-nas]   [+nas]

[+high]

If this were what underlying representation looked like too-- a structure with four autosegmental levels--consider just how many underlying forms there could be. We may even assume that there are a limited number of feature combinations that are

allowed to occur on any level that has more than one feature (some tiers, like that of nasality, will contain only one feature; others could conceivabley contain more than one).

We can calculate first how many possibilites there are even if there is always a one-to-one association between the levels, a wildly unrealistic assumption. Then in a language in which there were p_k possible segments on the kth tier, and t tiers, there would be

$$\prod_{k=1}^{t} (p_k)^n = (p_1)^n (p_2)^n \ldots (p_t)^n$$

possible underlying structures with a length of n segments. If we permit, realistically, a different number of segments on each tier, then this number rises quickly. To give a rough feeling for the size of this number, note that if <u>all</u> the features were on separate tiers, and were binary, then $p_k = 2$ for all k, and then there would be 2^{nt} underlying forms--morphemes--of length n, in a system with t tiers. This figure still assumes that there are the same number of segments on each tier, as I said; dropping this assumption, the figure becomes larger by orders of magnitude. The point to note, however, is that this rough estimate of the number of possible underlying forms, 2^{nt}, is exponentially related to the autosegmental tiers. The more possible underlying forms there are, the less redundancy there is, and the more difficult the task is of remembering the underlying form. Our hypothesis is, then, that the process of language-acquisition includes a process of "de-autosegmentalization," of merging the levels to form a minimal number of levels so that an alphabet, or phonemic inventory, can be realized by the language-acquirer. Once the child discovers what the underlying "alphabet" is, then, --even before such morpheme-structure constraints have been discovered, such as what the syllable can induce-- possible words are simply linear sequences of these "phonemes." These phonemes, furthermore, are the psychologically real elements of the phonological system that the child derives for the language.

2.   Deautosegmentalization

The way in which the origin of autosegmental tiers would best be explained would be if our account for the origin of segmentation, in the more traditional sense, itself led to the possiblity of autosegmental levels. The best situation, that is, would be to show that although autosegmental representation was different from the traditional view, it was nonetheless a natural consequence of the very process--in this case, of language acquisition--necessary to arrive at the traditional segmental view.

This, as I have indicated, is the approach we shall take. The idea can be made nost clear with a simple example.
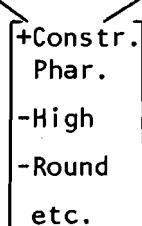
Suppose that a child learning English hears the (perhaps English) word "ha." The child, endowed with the abilities necessary to plot distinctive feature values, hears a mess of sound in the "ha." But this mess has some structure; first, it begins with a distinctive breathiness arising from a spread glottis, and then this breathing disappears. The rest of the word "ha" is homogeneous: it consists of a vowel-type quality we would identify as an /a/. Graphically, we have (4), then:

(4)   "ha"

breathiness (spread glottis)	breathy	non-breathy
mouth position	a	

In a representation more suggestive of autosegmental structure, we could write (4) as (5).

(5)   [+Spr Gl]   [-Spr Gl]

$$\begin{bmatrix} +Constr. \\ Phar. \\ -High \\ -Round \\ etc. \end{bmatrix}$$

I take (4) and (5) to be essentially equivalent.

If the task of deautosegmentalization is the task of the language-learner, then (5) must be converted into a non-autosegmental representation. We can imaginge at least three ways in which the child could do this (there are, the reader may note, others, less obvious):

(a) $\begin{bmatrix} + Spr\ Gl \\ - CP \\ - High \\ - Round \end{bmatrix}$ →   [-Spr Glottis]

Th first segment here is a breathy vowel /a̰/; the second is a segment specified only to bring the glottis back to its normal position.

(b)　　[+Spread Glottis]
$$\begin{bmatrix} \text{-Spr Glottis} \\ \text{+CP} \\ \text{-High} \\ \text{-Round} \end{bmatrix}$$

(b) is, of course, essentially the right solution; it consists of a linear sequence of an /h/--a segment whose nature is to spread the glottis but to have no influence on the gestures made in the mouth--and then a vowel /a/, specified as non-breathy.

(c)　A third possibility is:

$$\begin{bmatrix} \text{+Spread Glottis} \\ \text{+CP} \\ \text{-High} \\ \text{-Round} \end{bmatrix} \quad \begin{bmatrix} \text{-Spread Glottis} \\ \text{+CP} \\ \text{-High} \\ \text{-Round} \end{bmatrix}$$

This analysis consists, we see, of a breathey /ə̰/ followed by a non-breathy (normal) /a/.
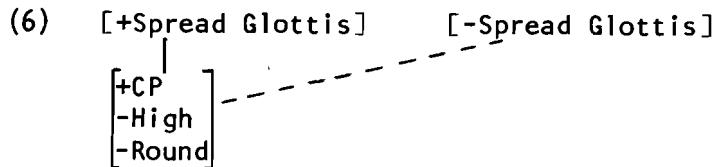
These three analyses for the underlying structure all give rise to the same superficial form, as in (4). Why this should be so is perhaps not obvious.

I shall assume the phonological system of an adult speaker to be faithful enough to reality to include, at the superficial level, a division of the features that compose underlying segments into the channels going to the separate articulatory components or the feature detectors. In this sense, even a competence grammar for English would contain a representation essentially like (4) or, equivalently, (5). Now if we are considering a competence grammar, it is a grammar which is equally a grammar for production and a grammar for comprehension. From this perspective, we may ask what the representation on the phonetic level of (a), (b) and (c) would be.
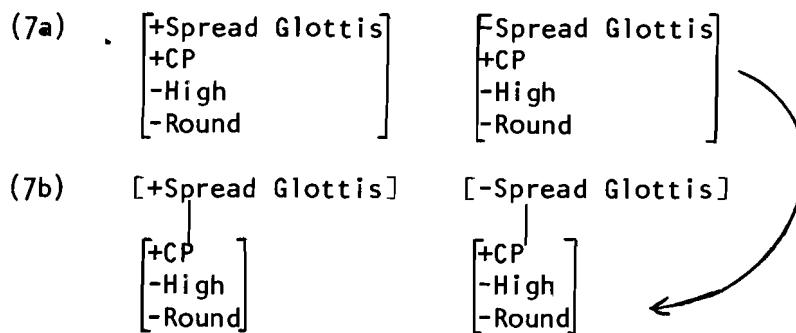
For purposes of exposition only, I shall assume that even at the phonetic level, the features "constricted Pharynx," "high," and "round" are on the same tier (as I already have in (4) and (5)). The transition from representations of the form (a), (b), and (c) to phonetic representation then consists of autosegmentalizing the feature "Spread Glottis." (a) then will become (6), which becomes (5) by the Well-formedness Condition, as I have indicated.

(a)　$$\begin{bmatrix} \text{+Spread Glottis} \\ \text{+Constr Pharynx} \\ \text{-High} \\ \text{-Round} \end{bmatrix} \quad [\text{-Spread Glottis}]$$

↓

(6)   [+Spread Glottis]   [-Spread Glottis]

$$\begin{bmatrix} +CP \\ -High \\ -Round \end{bmatrix}$$

(b) above will undergo the same change, *mutatis mutandis*, producing (5) also.   (c), on the other hand, will undergo the autosegmentalization illustrated in (7).

(7a) . $\begin{bmatrix} +Spread\ Glottis \\ +CP \\ -High \\ -Round \end{bmatrix}$ $\begin{bmatrix} -Spread\ Glottis \\ +CP \\ -High \\ -Round \end{bmatrix}$

(7b)   [+Spread Glottis]   [-Spread Glottis]

$$\begin{bmatrix} +CP \\ -High \\ -Round \end{bmatrix}$$ $$\begin{bmatrix} +CP \\ -High \\ -Round \end{bmatrix}$$

The question arises now whether the derived form in (7) is to be a permissible representation at the phonetic level. For if this level is to characterize both perception and production (7b) is *prima facie* an implausible candidate for the phonetic level.   There is no reason to say that two occurrences of the complex $\begin{bmatrix} +CP \\ -High \\ -Round \end{bmatrix}$, or /a/, are present.
From the perception end of things, there is no reason to posit other than one occurence of it, and surely the same is true of the production side from the phonetic point of view.   In sum, we are led back to the "obligatory contour principle" of Leben's, discussed in Chapters One and Three, and dismissed as false at the underlying level.   We return to it, now, as a realistic hypothesis for the phonetic level:

> Obligatory Contour Level (Revised): At the phonetic
> level, any continguous identical (auto)segments must
> be collapsed into each other.

The effect of this principle, then, is to change (7b) into (5). In so doing, we find that (a), (b) and (c) at the phonological level all correspond to (5) at the phonetic level.

What more can we say about the process of deautosegmentalization that takes (5) into one of (a), (b) or (c)?   Certain constraints are clearly going to be necessary in aiding the language-learner in this task.   First of all, notice what results if we "limit" the use of the revised Obligatory Contour Principle.   When we viewed it from the point of view of

acquisition, we see that for the learner to invoke the Obligatory Contour Principle is for him or her to reduplicate the autosegment, as in (7b), before merging the two levels together to form one sequence of segments. If we postulate this as a marked or psychologically more complex task than the alternatives (yielding, here (a) or (b)), we derive, as a consequence, a form of Leben's original Obligatory Contour Principle. That is, the contour's being obligatory at the phonetic level carries through to being the unmarked or expected situation at the deeper level. Thus it is no accident that Leben's formulation had a lot going for it; it described the linguistically normal situation. It was, however, due to a condition on simple grammars to learn, not on (as Leben suggested) knowable grammars.

Another consequence of this is that there will be, in the unmarked case, segments unspecified for some features. In (b) above, for example, the segment /h/ is specified only for the feature Spread Glottis, and unspecified for all other features.

This conclusion is only a step (though a real one) past the conclusion that we had reached already in Chapter One. There it was shown that morphemes could be unspecified for tonal features. This could happen if tone was autosegmental, and the morpheme had only a syllabic (non-tonal) entry in the lexicon. Since a morpheme could also be specified with a tonal autosegment, this left three categories for a morpheme to fall into with respect to the tonal features: marked +, marked -, or simply unmarked. We are drawing a different conclusion now, formally speaking, but it varies from the same conditions of autosegmental representation. The process of de-autosegmentalization, along with our hypothesis of the obligatory contour principle application being a marked state, leads directly to non-fully specified segments.

The process of deautosegmentalization, as a task for the language-learner, then, consists of learning which sets of feature-specifications on separate tiers may be merged together to form an acceptable segment in that language. The child learning English will eventually learn that there are no breathy vowels, but this is a fact about English, not language in general.

Should the child not find any consistent system by which to merge some autosegmental tier with the other sets of features, or should the separate tier of segments warrant status as an entry in the lexicon, this leaves the grammar with more than one linear sequence of segments at the underlying level. And in such a way, autosegmental phonology, as we have seen it in the first three chapters of this thesis, would arise.

The separate linear sequences of segments would then each constitute items that would have a place in the lexicon of that language; and the phonology would develop melodic association rules to lihk these two (or more) lexical entries together, as we have seen in Chapter Three.

3. Conclusion

Throughout this thesis has run one tension: that between the analysis of phonological phenomena in terms of distinctive features, and seeing phonological representations as composed of atomic, indivisible units. The motive force has been the belief that both of these are correct, and therefore ultimately reconcilable. But such occurrences as contour tones on short vowels seemed to threaten these assumptions; and the auto-segmental hypothesis arose as a way to save these two basic principles and also save the appearances.

There are some conclusions to be-drawn from the discussion in the first three chapters. It is an interesting realization that the formalism of generative phonology is insufficient, and that a multi-linear geometry is needed to deal with what have traditionally been called "suprasegmentals." Still, it is more the tradition--intellectually speaking--than any meta-physical consideration that has kept generative phonologists' representations on a straight line.

The conclusions that may carry some real interest are those that lead to an understanding of the ontogeny and phylogeny of language; and I have suggested in this last chapter how we may interpret autosegmental phenomena as very particular evidence for a human capacity--indeed, drive--to restructure sensations to fit a linear geometry of perception. This condensation, which I called "deautosegmentalization," is motivated not out of some physical or psychological in-ability, we may speculate, but rather in order to restructure phonetics, on both the perception and the production side, to fit in with the extraordinary human invention of language. This process of deautosegmentalization permits the human speaker to develop a lexicon built out of items that are linear sequences of elements from small inventory. Such a task is not difficult for the human, and so the endeavor of learning and using the other components of the grammar may proceed smoothly, each component--morphology, syntax, and semantics--free to develop the kind of geometry uniquely ap-propriate to it, unconstrained by the particular structure of the vocal apparatus, the very organs through which these com-ponents, and by them, thought, ultimately find expression.

FOOTNOTE

1.   This diagram is illustrative of the point in the text, but
what it makes equally clear is that the geometry of multi-
linear associations has not been dealt with. One could
approach this problem from several points of view. Clements
(1976) has suggested--as has Morris Halle in conversation--
following essentially Firthian lines and thus designating
one tier as basic with all other (prosodic) tiers associated
with it. No prosodic tiers would then be associated with
each other. This basic tier would look much like the non-
tonal tier used thoughout most of this thesis.

        Alternatively--following the lines sketched in the
preface, or those developed more seriously in Kahn (1976)--
tone and perhaps other features may be associated with a
tier of syllables. This approach is consistent with, but
not logically connected with, the first. Thus the 'basic'
tier may be chosen as that of syllables, but dealing with
syllables does not entail adopting the 'basic'/'prosodic'
distinction among autosegmental tiers.

        These are--frankly, and obviously--issues for future
research. As I indicated at the begining of Chapter Two
the success of our endeavor is measured both by the old
questions that are answered, and by the new questions
that are asked (vacuity arises, I might add, only in case
the answers to the old questions presuppose the but-yet-
available answers to the new). In this final chapter,
my intent is simply to indicate where I think interesting
questions lie.

BIBLIOGRAPHY

Arnott, D. W. 1964. "Downstep in the Tiv Verbal System." *African Language Studies* 5, 34-51.

Bird, Charles. 1966. "Determination in Bambara," *Journal of West African Languages* 3, 5-11.

Bloch, Bernard. 1948. "A Set of Postulates for Phonemic Analysis," *Language* 24, 3-46.

Bloch, Bernard and George L. Trager. 1942. *Outline of Linguistic Analysis*. Linguistic Society of America Special Publication.

Bolinger, Dwight. 1955. "Intersection of stress and intonation." *Word* 11, 195-203.

_____. 1958. "A Theory of Pitch Accent in English," in Bolinger (1965).

_____. 1965. *Forms of English; Accent, Morpheme, Order*. Harvard University Press, Cambridge, Mass.

_____. 1972. "Accent is Predictable--if you're a Mind Reader," *Language* 48, 633-645.

Bolinger, D. and Louis Gerstman. 1957. "Disjuncture as a Cue to Constructs," in Bolinger (1965).

Bresnan, Joan. 1972a. "On Sentence Stress and Syntactic Transformations," in Brame, ed., *Contributions to Generative Phonology*, University of Texas Press.

_____. 1972b. *Theory of Complementation in English Syntax*. MIT PhD dissertation.

_____. 1976. "Towards a Realistic Model of Transformational Grammar," presented at 1976 MIT-AT&T session on Communication, MIT.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.

_____. 1970. "Remarks on Nominalizations," reprinted in Chomsky, *Studies on Semantics in Generative Grammar*, Mounton.

Chomsky, N. and Morris Halle. 1968. *The Sound Pattern of English*. Harper and Row, New York.

Clark, Mary. 1975. "Evidence from Igbo for a Theory of Tone."
    Unpublished paper, University of Massachusetts at Amherst.

Clements, G. N. 1976. "Vowel Harmony in Non-linear Generative
    Phonology: the Autosegmental Theory (I)." Unpublished
    paper, Harvard University.

Cohen, Michael. 1974. "Topics in Tonga Tone." Unpublished
    paper, MIT.

Crystal, David. 1969. *Prosodic Systems and Intonation in
    English.* Cambridge University Press.

Dubuisson, Colette and John Goldsmith. (in press) "A Propos
    de l'Inversion du Sujet Clitique en Francais," in *Montreal
    Working Papers in Linguistics* vol. 7.

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

Emonds, Joseph. 1970. *Root and Structure-Preserving Trans-
    formations.* MIT PhD dissertation. (forthcoming from
    Academic Press).

Goldsmith, John. 1974. "English as a Tone Language." MIT
    ditto. To appear in *Communication and Cognition* 1977.

_____. 1975a. "An Autosegmental Typology of Tone." In
    *Proceedings from the Fifth North Eastern Linguistics
    Society,* ed. Ellen Kaisse, Harvard University.

_____. 1975b. "Tone Melodies and the Autosegment." In
    Proceedings from the Sixth Conference on African Linguistics,
    in *Ohio State Working Papers in Linguistics* 20, ed. R. Herbert.

_____. In press. "An Overview of Autosegmental Phonology."
    Linguistic Analysis Vol. 2, No. 1, 1976.

Green, M.M. and G. E. Ifwe. 1963. *A Descriptive Grammar of
    Igbo.* Oxford University Press.

Halle, Morris. 1973. "Stress Rules in English: A New Version."
    Linguistic Inquiry 4,4.

_____. 1975. "On Russion Accentuation." *Slavic and East
    European Journal* 19, No. 1.

Haraguchi, Shosuke. 1975. *The Tone Pattern of Japanese: An
    Autosegmental Theory of Tonology.* MIT PhD dissertation.

Harris, Zellig. 1941. Review of Trubetzkoy, *Grundzüge der Phonologie*. *Language* 17, 345-349.

_____. 1944. "Simultaneous Components in Phonology." *Language* 20, 181-205.

Haugen, Einar. "Phoneme or Prosodeme." *Language* 25, No.3.

Higgins, F. R. 1974. "On the Use of Idioms as Evidence for Movement: A Cautionary Note." Presented at the Winter 1974 LSA Conference.

Kahn, Daniel. 1976. *Syllable-Based Generalizations in English Phonology*. MIT PhD dissertation, Indiana University Linguistics Club.

Kiparsky, Paul. 1963. "The Inflectional Accent in Indo-European." *Language* 49, No. 4.

Leben, William. 1973. *Suprasegmental Phonology*. Unpublished MIT PhD dissertation.

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

Lovins, Julie. 1971a. "Melodic Conspiracies in Lomongo Tonology." *Papers from the Seventh Regional Meeting of the Chicago Linguistics Society*.

_____. 1971b. "The tonology of Lomongo Reduplication." *Studies in African Linguistics* 2, No. 3, 257-270.

Lunt, Horace. 1973. "Remarks on Nasality: The Case of Guarani." In Anderson and Kiparsky, eds., *A Festschrift for Morris Halle*, 1973, Holt, Rinehart and Winston, New York.

May, Robert and John Goldsmith. 1975. "Tone Sandhi in Sanskrit." Unpublished paper, MIT.

McCawley, James. 1970. "A Note on Tiv Conjugation." *Studies in African Linguistics*, 1, 123-239.

_____. 1973. "Some Tonga Tone Rules," In *Festschrift for Morris Halle*.

_____. 1974. "Global Rules and Bangubangu Tone," in C. Kisseberth, ed., *Studies in Generative Phonology*.

Palmer, F. R., ed. 1970. *Prosodic Analysis*. Oxford University Press.

Rivas, Alberto 1974. "Nasalization in Guaraní," unpublished MIT paper.

_____. 1975. "Nasalization in Guaraní," in E. Kaisse, ed., *Papers from the Fifth North Eastern Linguistics Society Meeting*.

Selkirk, Elizabeth. 1972. *The Phrase Phonology of English and French*. MIT PhD dissertation.

Spa, Jaap. 1973. *Traits et Tons en Enya, Phonologie Générative d'une Langue Bantoue*. Musée Royale de L'Afrique Centrale, Tervuren, Belgium.

Trager, G. L. and H. L. Smith. 1951. *An Outline of English Structure*. Washington, American Council of Learned Societies, Studies in Linguistics Occasional Papers #3.

Trougetzkoy, N. S. 1939. *Principes de Phonologie*. Editions Klincksieck, Paris.

Vanvik, Arne. 1961. *Stress in Present Day English*. Norwegian Universities Press. Bergen, Oslo.

Voorhoeve, J., A. E. Meussen, K. F. DeBlois. 1969. "New Proposals for the Description of Tone Sequences in the Igbo Completive Phrase." *Journal of West African Languages* 6, 79-84.

Williams, Edwin. 1971. "Underlying Tone in Margi and Igbo." Unpublished MIT paper. To appear in *Linguistic Inquiry*, 1976.

Woo, Nancy. 1969. *Prosody and Phonology*. Indiana University Linguistics Club publication; MIT PhD dissertation.