

Tonal Complexes and Tonal Alignment

Akin Akinlabi

Mark Liberman

Rutgers University

University of Pennsylvania

1. Introduction

Tone was one of the first areas of generative phonology where constraint-based approaches were proposed (Leben 1973, Goldsmith 1976, Williams 1976). Such works raised the hope that tonal phonology could be explained in terms of universal (if parameterized) constraints operating on appropriate representations. However, derivations based on language-specific rules were later argued to be necessary, in order to handle the wide range of attested tonological phenomena (Hyman and Ngunga 1994).

We propose that a successful account of tonal phonology, constraint based or not, requires enriching tonal representations to include some simple kinds of structures, which organize tones in somewhat the same way that segments, syllables and feet organize non-tonal features. Our proposal is similar in spirit to that of Bamba (1991). Constraints mentioning such tonal structures can motivate deletion, epenthesis, spreading or re-ordering of tonal features, just as constraints on syllable or foot structure may motivate such processes in well-known cases of segmental phonology. Such structures license tones just as syllables license segments.

Specifically, this paper proposes a tonal unit consisting minimally of paired HL or LH tones. We argue that such units, long postulated as underlying elements in accentual systems, also play a crucial role in tonal phonology more generally.

1.1 Contour Formation in Benue Congo

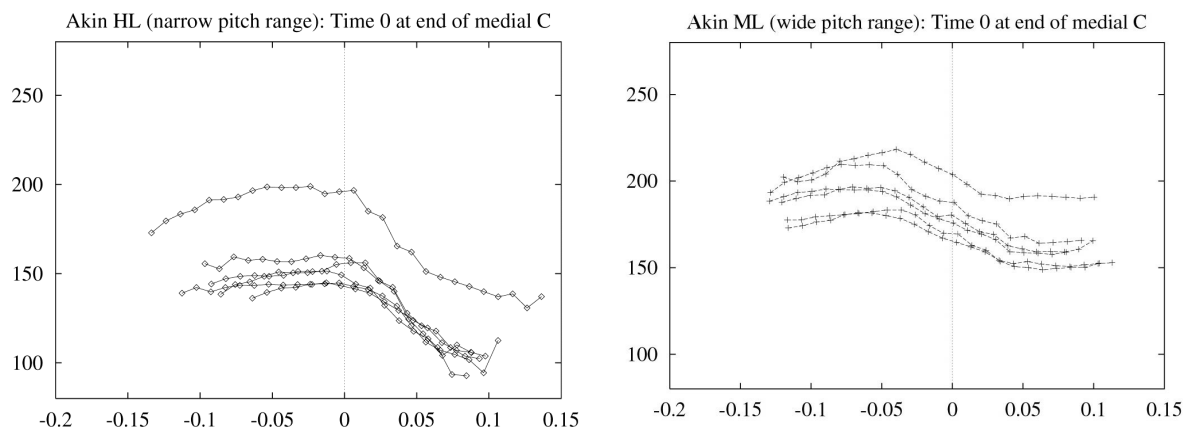
In the Benue Congo languages of West Africa, we often see a process in which the transition from one tonal level to another is delayed so as to create a salient tonal contour

on the syllable beginning the new level. In a notation in which vowels are tagged as "High", "Low", "Rising", "Falling", etc., this leads to the creation of rises and falls, as in the examples below from Yoruba¹ and Edo (Bini):

- (1) (a) Yoruba rising example
 àlá (LH) → àlǎ (L LH) 'dream'
- (b) Yoruba falling example
 rárà (HL) → rárâ (H HL) 'elegy'
- (c) Edo (Bini) falling example
 ékpò (HL) → ékpô (H HL) 'bag'

Phonetically, F0 transitions are always contours, not step functions; however, the effects symbolized by the notations in (1) are not just the necessary phonetic consequence of tonal co-articulation. The most striking evidence for this assertion is the fact that a language may form phonologically-significant contours for some tonal sequences and not others. Thus Edo forms phonologically-significant contours in the case of High Low but not Low High sequences; while Yoruba does so for both Low High and High Low, but not for any case where Mid tone precedes or follows High or Low.

(2) Yoruba HL and ML plots



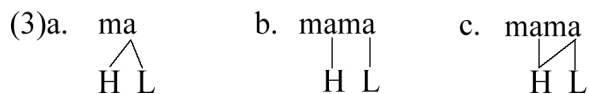
. The difference between the Yoruba H L and M L cases can be seen in the pitch tracks in (2). Each plot shows the F0 tracks for six VCV sequences, taken from the middle of longer phrases. The vertical line marks the point of release of the medial

¹ Note that in Yoruba, the standard orthography shows the input to this process, not the output, so that the resulting contours are not normally indicated in the written forms; therefore the examples in (1) are written as in the left hand column.

consonant, here always a nasal or liquid. The HL cases were spoken in a relatively narrow pitch range, while the ML examples were spoken in a relatively wide pitch range. As a result, the amount of pitch change is similar, but the timing of the pitch change is different: in the HL examples, the pitch fall does not begin until after the release of the medial consonant and continues throughout most of the following vowel, while in the ML examples, the fall is about half complete by the release of the medial consonant, and generally ends early in the following vowel. Informal experimentation suggests that this roughly (half-segment-long) difference in timing can be crucial for native-speaker perception of Yoruba tonal identity.²

In some level-tone languages, for example Igbo and Ijo (Kalabari, Nembe, Kolokuma), phonologically-significant contour tone formation does not appear to occur. Note however that in a language where there is no phonological contrast between the presence and absence of contour formation, it may remain to be determined whether the situation should be described as invariable absence of contour formation, or rather invariable formation of contours.

For the past few decades, phonologists have generally followed Hyman and Schuh (1974:88) in treating this process of tonal contour formation as "tone spreading." In this approach, all tonal specifications are built up from a small number of level tone primitives, such as High and Low, with rises and falls treated as Low+High or High+Low sequences respectively. Strings of tones are represented on a separate tier from strings of non-tonal segments, with the alignment of tonal and segmental strings indicated by association lines connecting them. In this perspective, a falling tone is just a tonal sequence HL associated with a single segment, as in (3a) below. When L follows H on successive syllables, as in (3b), such a contour can be created on the second syllable simply by adding an association line, as in (3)c.



1.2 Towards a Solution

Using arbitrary phonological rules or constraints, formal specification of the desired outcome in such cases is easy: in certain contexts, certain association lines are added; in other contexts, they are not.

² Observations of these phenomena go back to the earliest linguistic treatments of the languages in question. Thus in Ida Ward's 1952 work "An Introduction to the Yoruba Language", she remarks that (p. 34) that "[t]he juxtaposition of high and low tones, either high-low or low-high, needs some comment." Citing an example of the form HL, she observes that there is a fall on the second syllable that is "heard as a more or less deliberate glide," and warns that "[u]nless the English speaker makes it, he is apt to give the impression of gliding down on the first syllable ..., which does not satisfy the Yoruba."

As Hyman and Schuh observed in 1974, there is a general pattern to be accounted for. Such tone contour formation is common, but by no means universal. When it happens, the change is always in the direction of a delay rather than an advance of the F0 transition. It is possible to speculate about general articulatory reasons why F0-transition delays might be more likely to happen than F0-transition advances, or general acoustic reasons why delays might be more salient than advances. However, we will argue instead for an account of the asymmetry in terms of formal properties of the phonological representation of tone.

In current approaches to phonology, systematic changes like the one exemplified in (3) are attributed to the existence of general constraints that prefer the output to the input. However, there is no generally-recognized constraint that would motivate the kind of contour formation shown in (3). In fact, as we will see in detail later on, the facts of Yoruba seem to motivate a constraint that is ironically opposite to the well-known "well-formedness constraint" proposed by Goldsmith for tonal association: a tone spreads if and only if the adjacent syllable already has a tone.

Why then does tone contour formation tend to occur? We believe that the answer to this question is tied up with the answer to several other, apparently unrelated questions about tone and accent: Why do tone polarization and polarized tone epenthesis tend to occur? Why do multiple-tone sequences sometimes but not always simplify? Why do high-before-low raising and low-before-high lowering tend to occur?

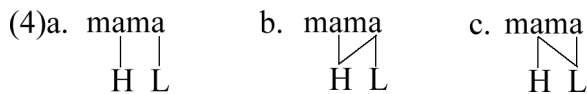
None of these phenomena are inevitable, but all are commonplace, and typologically typical of tonal systems. We believe that all of these cases are symptoms of the same cause, namely the formation of tonal complexes. Tonal complexes are "bound states" of (two or more) unlike tones, such as [high low] or [low high], and we propose that they have a role in organizing tonal features somewhat analogous to that of moras and syllables in organizing segmental features.

On this analogy, tone contour formation is like re-syllabification, in which a coda consonant becomes an onset for a following syllable. Tone polarization and polarized tone epenthesis are like the epenthesis of vowels and consonants in rescuing forbidden or marked syllable structures. And the phonetic dissimilation of tone sequences is like the different phonetic interpretation of high vowels or nasals in onset vs. rhyme positions in syllables.

In such cases, the availability of alternative structural configurations can provide a tool for motivating constraint-based restructuring, or explaining differing phonetic interpretations.

This is enough to suggest why (3c) is sometimes preferred to (3b). After spreading, the adjacent H and L tones form a tonal complex, and a constraint requiring tones to be bound into tonal complexes will then be satisfied.

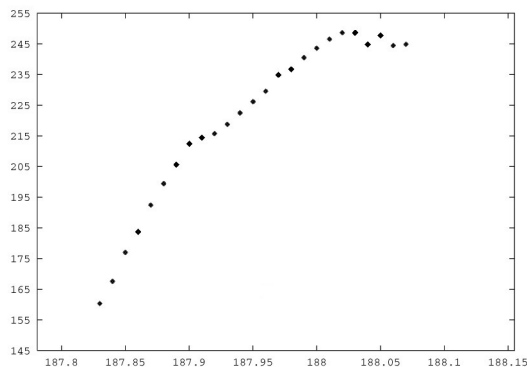
As we have explained things so far, however, we have left a crucial observation of Hyman and Schuh 1974 unexplained. We have shown how to motivate the particular kind of tone-onto-tone spreading that results in tonal contour formation. However, we have not explained their important generalization that this process always delays and never advances the point of pitch fall or rise. For example, HIGH LOW always become HIGH FALLING, never FALLING LOW, even though both outcomes produce a tonal complex in our sense. Using autosegmental notation, the result of an input like (4a) is always as in (4b) below, never as in (4c):



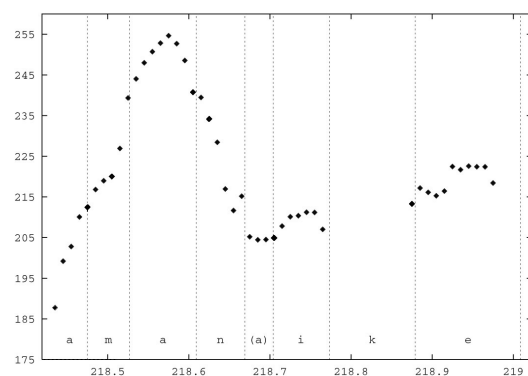
In order to explain this generalization, we will need to look further into the kind of thing that a "tonal complex" is. Hints are provided by two other general observations about the phonetic realization of tone, namely that the F0 target for a single static tone tends to occur at the (temporal) end of the associated phonetic region, and that the simplest cases of "dynamic" or contour tones force us to posit a second, earlier (phonetic) alignment point within the associated time region.

The plot in (5a) shows fundamental frequency as a function of time for the Igbo word *yá* meaning "he". Although this monosyllabic word has high tone, the pitch is not a uniform level high. Instead, the pitch rises throughout the syllable, with the peak value found near the end. In languages like Igbo and Yoruba, other things equal, the phonetic target value of a tone – the highest F0 of a High tone, or the lowest F0 of a Low tone – is found at the end of the span of time corresponding to the associated tone-bearing unit.

(5a) Igbo: *yá* ("he")



b. Igbo: *ámá nà íké* ("Ama and Ike")



We need an additional F0 target at the start of the utterance. We can think of this as a junctural value, or as a default value, but in any case, it usually falls in between the target values of an initial High and an initial Low tone, so that a low-tone stretch would be falling, just as this high-tone syllable is rising.

When we look at tone sequences, and at phrases involving a longer stretches of tone-bearing units with the same tone, the same pattern generally holds. For example, consider the pitch track shown in (5b) for the Igbo phrase *ámá nà íké* “Ama and Ike”. The first two syllables are High, but they are neither uniformly high, nor rising followed by high. Instead, there is a rise distributed over the two-syllable high-tone region, with the highest point falling at the end of that region. The next syllable is Low, and the low target is unsurprisingly at its end. The last two syllables are High again, and again the high target (lower than before because of downdrift and final lowering) is at the end of the two-syllable low-tone region.

Thus a crude but roughly correct way to synthesize sentential pitch contours for a language like Igbo is:

1. Divide the utterance into maximal regions of like tone.
2. Place a mid-valued tonal target at the start of the utterance.
3. Place a tonal target at the end of each region, choosing an F0 value determined by the tonal type, downdrift/downstep, and final boundary effects if any.³
4. Interpolate linearly from target to target.

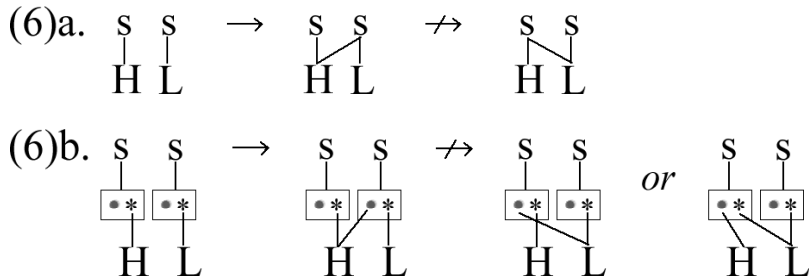
When we look at tone-bearing units that carry a contour tone, as in the Yoruba HL examples in (2), we can see the need to postulate additional phonetic anchor points at the beginning of some phrase-internal tone-bearing regions. We believe that these phonetic modeling practices point the way to an appropriate phonological structure for tone.

Let’s assume that in Igbo or Yoruba every tone-bearing region has two potential tonal targets. The one at the end is obligatory, while the one at the beginning may remain unfilled. We propose to reify this pair of potential tonal targets, and call it a “tonal complex.” In this sense, a tonal complex (TC) is an entity that binds two (or perhaps more) tonal positions to a tone-bearing region (which is one or more moras or syllables). These two tonal positions are a bit like the onset and rhyme of a syllable: whenever one is present, the other “ought” to be there, even though it may often be lacking. And as with the onset and rhyme of a syllable, these structurally differentiated positions are not equally optional.

³ See Liberman et al. (1993) for a simple model of downdrift and downstep in Igbo phrases.

In the simple binary tonal complexes we are considering, the second tonal "hook" is the stronger of the two positions, and therefore is the default. If only one tone is available, it will go there, with the earlier tonal position in the TC remaining empty.

Therefore we can reformulate (6a) below as (6b):



In both (6a) and (6b), the *s* symbols in the top row stand for syllables (or other tone-bearing units). In (6b), we've used boxes to symbolize two-element tonal complexes. Each tonal complex as a whole connects to a tone-bearing unit, as indicated in the diagram by the fact that the upper association line connects to the box. (Although it is not shown in (6b), a tonal complex might also link to a string of adjacent tone-bearing units). Within the tonal complex, we've used an asterisk and a dot to symbolize the primary and secondary tonal association points. The tones link to these association points within the tonal complex.

Given this notation, the diagrams in (6b) make it is easy to see why the "forward-spreading" option should be more natural than the "backward-spreading" one. Forward spreading requires adding one association line while deleting none, as shown in the second configuration in (6b). We can create backward spreading, graphically speaking, by adding a crossing association line, as in the third panel of (6b) – but this is presumably a configuration that a proper formalization of association wouldn't even permit to be stated. We can also do it by deleting one association line and adding two, one of which links a tone to a non-adjacent association point, as in the fourth panel of (6b) – another configuration that should probably be formally impossible. Finally, we could do it by deleting one association line and adding three (not shown, but obvious), thus at least creating a formally plausible configuration, but producing an output that is much less faithful to the input.

Just as not all syllables are CV, so some tonal complexes might be more complicated. Several languages have been cited, for example, in which the maximal tonal pattern allowed for a word is of the form L+HL, where the L if present must align with the start of the word, while the HL constitutes an optional accent. Kyoto Japanese and Mawukakan are both examples of this type.⁴ This would naturally be treated as a three-

⁴ See Poser (1985) for a discussion of Kyoto Japanese, and Bamba (1991) and Liberman (1995) for discussion of Mawukakan.

element tonal complex, analogous to a CVC syllable, associated with a domain such as the accentual phrase or the prosodic word.

In the sections that follow, we will apply this style of analysis to the simple but subtle case of Yoruba, and we will also suggest that pitch accents should in general be analyzed formally in terms of tonal complexes (as has long been the informal analytic practice).

2. Yoruba

Yoruba has three phonemically distinctive tones-H(igh), M(id), and L(ow). H occurs in word-initial position only in marked consonant-initial words, which reveal an implicit initial vowel when preceded by another word in genitive construction. Most words start with a vowel, which is L or M but not H. Except for this minor tonotactic restriction, tones occur freely in lexical representations, without apparent restrictions on word melodies. So there are three possible tonal patterns for monosyllables, nine possible tonal patterns for disyllables, and so on, as in (7).⁵

(7) Lexical tone contrast:

ra H 'to disappear'	ra M 'to rub'	ra L 'to buy'
ọkọ MH 'hoe'	ọkọ MM 'husband'	ọkọ ML 'vehicle'
ilu LH 'town'	ilu LM 'opener'	ilu LL 'drum'
pako HH 'plank'	kese HM mythological place name	pako HL 'chewing stick'

2.1 Non-specification of the Mid tone

The Yoruba mid tone has been analyzed as underlying tonelessness since Akinlabi (1985) and Pulleyblank (1986). In both Akinlabi's and Pulleyblank's works, several arguments are given for this hypothesis. We will briefly sketch two examples, relating to *tonal stability and tone spreading*.

⁵ The examples in this paper are given in the standard Yoruba orthography. In this orthography, ẹ = [ɛ], ọ = [ɔ], ẹ́ = [ɛ́], ọ́ = [ɔ́], ẹ̀ = [ɛ̀], ọ̀ = [ɔ̀], p = [kp], j = [j]. A nasal vowel is written as an oral vowel followed by “n”, otherwise an “n” before a consonant represents a syllabic nasal. An acute accent on a vowel indicates a (H)igh tone, a grave accent marks a (L)ow tone. (M)id tones are unmarked. Where necessary we indicate the tones with the letters HML in addition to marks on the vowels.

2.1.1 Tonal Stability

When an object noun follows a verb in Yoruba, the two words are combined phonologically by deleting either the final vowel of the verb or the initial vowel of the object. Any High or Low tones of the deleted vowel are retained in the result. However, Mid tones are not “stable” in this sense, but instead behave in various combinations with other tones as if they were simply not there. Thus a Mid tone verb followed by an object whose initial vowel is Low will yield a combined form whose first vowel is simply Low, not some sort of Mid-Low contour, or a Mid with a following downstep, or anything else of the sort.

The crucial cases are exemplified below. The tone patterns in each of the (a) and (b) examples in (8) - (12) are the same; in the (a) examples the vowel of the verb is deleted whereas in the (b) examples the vowel of the noun is deleted.⁶

H verb + L initial noun			
(8)a.	wa (H) + look (for)	ẹkọ (LH) education	→ wẹkọ (H LH) look for education
b.	mu (H) + take	iwe (L H) book	→ muwe (H LH) take a book
(9)a.	wa (H) + look (for)	ọnọ (L L) way	→ wọnọ (H L) look for a way
b.	wa (H) + look (for)	imọ (L L) knowledge	→ wamọ (H L) look for knowledge
(10)a.	ji (H) + steal	ọbẹ (L M) knife	→ jọbẹ (H (L) M) steal a knife
b.	fẹ (H) + want	iwo (L M) horn	→ fẹwo (H (L) M) want a horn
H verb + M initial noun			
(11)a.	wa (H) + look (for)	owo (MH) money	→ wowo (H H) look for money

⁶ We will not discuss vowel deletion, which is a complex question requiring a monograph-sized treatment of its own.

- b. wa (H) + ile (MH) → wale (H H)
 look (for) house look for a house

M verb + L initial noun

- (12)a. jọ (M) ajẹ (L H) → jajẹ(L H)
 resemble witch resemble a witch
- b. sin (M) oku (L H) → sinku (L H)
 bury dead (body) bury the dead

A few remarks are necessary for the motivation behind the selection of the above forms. First, as noted above since V-initial nouns cannot start with H in Yoruba, no examples of the form X+HX can arise. Second, when a L-tone verb precedes its object, the tone always deletes even if the vowel is preserved, so the case L+XX offers no evidence in this matter.

Extracting the tonal input and output alone from the above examples, we have the following:

Summary of Tonal Input and Output in Yoruba V+N combinations:

- (8)a-b H + L H → H L H
- (9)a-b H + L L → H L
- (10)a-b H + L M → H L M⁷
- (11)a-b H + M H → H H
- (12)a-b M + L H → L H

Thus in all the cases that can arise, and whose output is not obscured by the deletion of the verbal L, we can say that H and L always remain when their lexically-associated vowel deletes, while M never does. We assume therefore that Yoruba has privative H and L tones, and that the Mid tone is simply lack of tone.

⁷ Any examples whose output is specified as (HL M) are pronounced exactly as this notation implies in some dialects, but in standard Yoruba, they are pronounced as a raised H followed by an M. In earlier studies (see Bamgbose 1966, Akinlabi 1985, Pulleyblank 1986), this was thought to be an H followed by a tone between M and L, a sort of "downstepped Mid." The essential point is that the L tone is in some sense preserved.

2.1.2 "Tonal Spreading" treats Mid as nonexistent

In Yoruba, a sequence of L H is realized as L LH and a sequence of H L is realized as H HL, that is adjacent H and L tones always spread (rightwards) onto each other, creating LH and HL contours.

- (13) ala (LH) → ala (L LH) 'dream'
 rara (HL) → rara (H HL) 'elegy'

On the contrary, an M L sequence does not become *M ML and an M H sequence does not become *M MH. That is, M never spreads to any other tone; and so there are no MX contours, as indeed there could not be if Mid does not exist.

- (14) ole (ML) → ole (ML) 'thief'
 ile (MH) → ile (MH) 'house'

Furthermore, when H or L is spread onto M, M is (usually) completely erased. Thus in monomorphemic words with initial M, a following H or L can optionally spread backwards onto the first syllable. In this case the initial M is completely supplanted, and the result is homophonous with an underlying HH or LL sequence.

- (15)a. MLL LLL / MLL
 eřrà èèrà / eèrà 'ants'
 erùpè èèpè / eèpè 'sand'
- (15)b. MHH HHH / MHH
 egúngún éégún / eégún 'masquerade'
 orúkọ óókọ / oókọ 'name'

2.2 The paradox of Yoruba tone (non-)spreading and (non-)relinking

The Yoruba cases of tone non-spreading and non-relinking present a paradox. In configurations (16a) and (16b) below (illustrated in 17a), where each syllable has its own tone, the first syllable's tone insists on crowding onto the second syllable. In configurations (16c) and (16d), where the second syllable is unspecified for tone, the first syllable's tone stays home, leaving its neighbor tonally empty (i.e. Mid). (See examples in (17b))

- (16) a. σ σ b. σ σ c. σ σ d. σ σ
 H L L H H L
- (17)a. ala (LH) → ala (L LH) 'dream'
 rara (HL) → rara (H HL) 'elegy'

- (17)b. $kese (H\emptyset) \rightarrow kese (H\emptyset)$ ‘mythological place name’
 $ilu (L\emptyset) \rightarrow ilu (L\emptyset)$ ‘opener’

Furthermore, in vowel deletion a Low tone “relinks” if the final syllable has a high tone (18a), but again not if the final syllable is tonally empty (18b). Examples are in (19a) and (19b) respectively.

- (18) a. $\begin{array}{ccc} \sigma & \sigma & \sigma \\ H & L & H \end{array} \rightarrow \begin{array}{cc} \sigma & \sigma \\ H & LH \end{array}$ b. $\begin{array}{ccc} \sigma & \sigma & \sigma \\ H & L & \end{array} \rightarrow \begin{array}{cc} \sigma & \sigma \\ & HL \end{array}$

- (19)a. Input \rightarrow Output (19)b. Input \rightarrow Output
 $m \ u \ + \ i \ w \ e$ $m \ u \ w \ e$ $f\grave{e} \ + \ i \ w \ o$ $f\grave{e} \ w \ o$
 $\begin{array}{ccc} | & | & | \\ H & L & H \end{array}$ $\begin{array}{ccc} | & / & | \\ H & L & H \end{array}$ $\begin{array}{cc} | & | \\ H & L \end{array}$ $\begin{array}{c} | \\ H \end{array}$

Yoruba thus presents a case ironically antithetical to the original Williams/Goldsmith Well-Formedness Condition: the rule seems to be that a tone “spreads” or “relinks” if and only if the target syllable already has its own tone!

Another three-tone language with similar facts to Yoruba is Ghotuo (Elugbe 1985, 1995).

2.3 Sketch of an analysis of Yoruba

Following the suggestions that we made in section 1.2, we propose that Yoruba tones do not spread or relink in order to satisfy the needs of adjacent tonally unspecified vowels (or moras or syllables), but rather do so in order to fill available positions in adjacent tonal complexes. These are prosodic combinations of HL or LH tones, roughly in the way that moras and syllables are prosodic combinations of segments. On this analysis, such tonal units, long postulated as underlying elements in accentual systems, also play a crucial role in this aspect of Yoruba tonal phonology.

There are several approaches we might take to formalizing this idea. One obvious one is to assume that every underlying Yoruba tone is connected to the primary association point of its own tonal complex, which has an “iambic” [w s] pattern. When two tonal complexes happen to be adjacent (i.e. associated with adjacent syllables), the first tone will spread in order to link to the empty secondary association point of the second tonal complex. This is basically the same thing that happens when a consonantal segment links to an empty onset position in a following syllable.

It is gratifying that following this approach, we can resurrect the Williams/Goldsmith Well-Formedness Condition as a workable treatment of Yoruba.

Echoing the Williams/Goldsmith language for old times' sake, the constraints might be something like:

1. Association lines can't cross (and can't skip either).
2. Don't delete underlying association lines.
3. Every tone must be associated with some tonal complex.
4. Every tonal complex must be associated with some tone-bearing unit.
5. Every strong position in a tonal complex must be associated with some tone.
6. Every weak position in a tonal complex must be associated with some tone.
7. Don't add association lines.
8. Every tone-bearing unit must be associated with some tonal complex.

We share with Williams/Goldsmith the prohibition against crossing associations. The original Williams/Goldsmith WFC had two other parts: every tone must be associated with some TBU, and every TBU must be associated with some tone. Since we mediate these associations through tonal complexes, which have two elements on their tone-linkage side, an analogous statement logically requires at least five constraints: 3, 5, 6, 4 and 8 in the list above.

These constraints are violable and ranked in the currently usual style. Constraint 2 was implicit in Williams/Goldsmith. We need to add constraint 7, and to rank it higher than constraint 8, so as to permit Mid tones to survive, but we need to rank it lower than the other constraints, so as to permit spreading of High onto Low and Low onto High.

A full treatment requires a more careful formalization of the postulated structures, and some revisions and additions to handle aspects of Yoruba tonology that we haven't discussed here. One might also try to merge the set of constraints with those used in current theories of syllable structure. Appropriately formalized and revised, we believe that such an approach can restore to tonal phonology the attractive simplicity and elegance of constraint-based theories.

3.0 Phonetic and phonological similarities to Japanese pitch accent

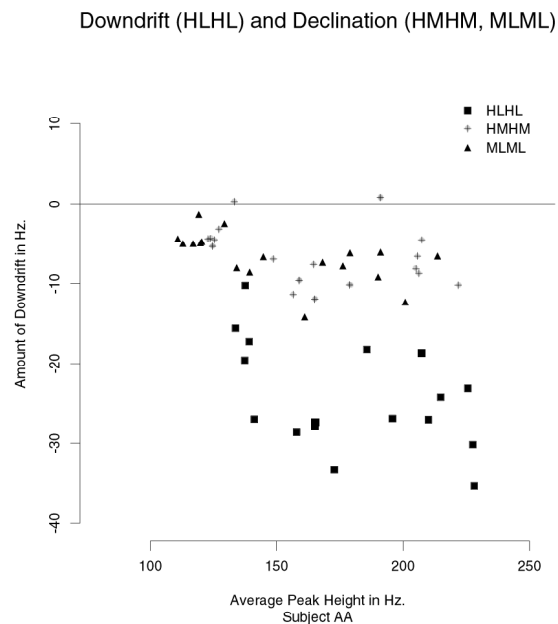
The tonal complexes postulated as derived structures in Yoruba have both phonetic and phonological similarities to (underlying) pitch accents in languages such as Japanese and English. Many descriptions, going back at least a century, have treated pitch accents (and also dynamic tones) as somehow singular entities, even if they are also sequences of tone levels. In the recent literature, Pierrehumbert (1980) argued that certain English pitch accents should be treated as sequences of two tones, and therefore should be called "bitonal", but that the two tones should share a single linkage to an accented syllable. Pierrehumbert and Steele (1987) proposed that alignment differences in such bitonal accents should be treated in terms of [s w] vs. [w s] labeling. Yip (1989) argued for branching structures of tones in the treatment of various Han dialects. Pierrehumbert and

Beckman (1988) proposed that Japanese accents should be treated as branching units of H L tones with [s w] labeling and a single shared link to the accented mora (p. 124-5).

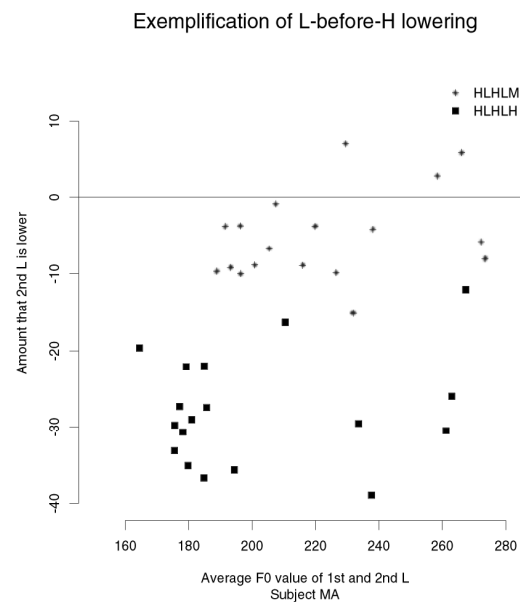
Two specific phonetic parallels between Yoruba and Japanese are particularly suggestive: the limitation of “catathesis” or “downdrift” to bound sequences of HL or LH; and the dissimilation of H and L levels in bound sequences.

Poser (1985) showed that “catathesis” (his neologism for a strong local lowering and compression of pitch range) is triggered in Japanese by a pitch accent, but not by the otherwise similar sequence of higher and lower pitch created by an accentless word, which shows only a relatively weak downtrend that he called “declination.” This result is a robust one, and has been replicated many times. In Yoruba, “downdrift” (the traditional Africanist term for the same phenomenon as catathesis) occurs in sequences of adjacent H and L tones, but not in HMHM, MLML, or HMLMHML sequences (LaVelle 1974). The scatterplot in (20a) shows this effect. Each point plots the relationship between two successive F0 maxima, with the average of the two maxima plotted on the x-axis, and the difference between the two maxima plotted on the y-axis. The black squares (the HLHL case) are uniformly lower than the black triangles (MLML) and the gray stars (HMHM).

(20)a. Yoruba downdrift vs. declination



b. Yoruba L-before-H lowering



In Yoruba, H is well known to be raised before L as compared to before M (Akinlabi and Laniran (1987), Connell and Ladd (1990), Laniran (1992)). This is a general phenomenon of F0 dissimilation in the cases that we are characterizing as sequences bound into tonal complexes. The scatterplot in (20b) exemplifies L-before-H

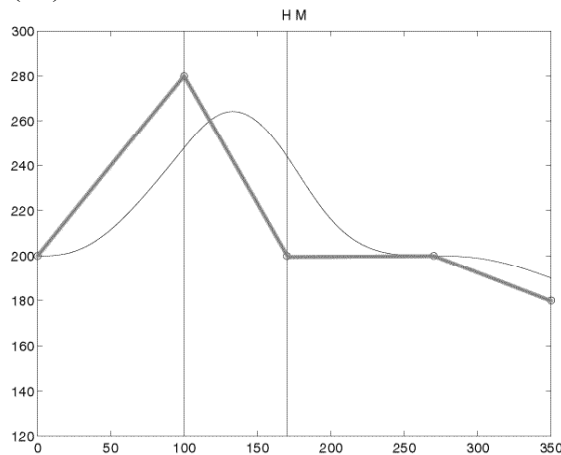
lowering in Yoruba: each point plots the relationship between two successive F0 minima, with the x-axis showing the average value of the two minima, and the y-axis showing the difference between them. The black squares (minima in HLHLH sequences) are uniformly below the gray stars (minima in HLHLM sequences).

Yoruba H-before-L raising is reminiscent of the long-observed fact that Japanese accentual H is higher than non-accentual H (called "accentual boost" by Kubozono (1993), who was the first to provide careful phonetic documentation). This is true even though Japanese accent is not stress-like, in that it does not cause greater segment durations, is not considered a strong position for alignment with musical meter, etc.

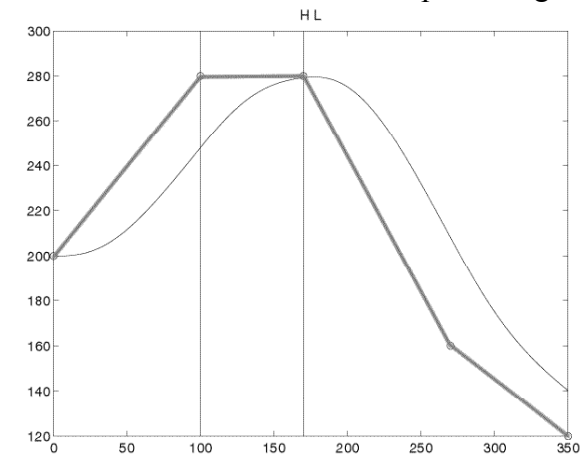
3.1 A speculative connection between tonal spreading and tonal dissimilation

There are a number of plausible reasons for the dissimilatory phenomena just described. For example, we know that the same segment will be phonetically different in different syllabic positions, often in a predictable way. Onset position is somehow more consonantal, so that tract closing gestures will be longer and stronger there than in the coda, whereas by contrast a nasal in coda position will have a longer and stronger velum opening gesture than in onset position. Perhaps something similar is happening in tonal complexes. Perhaps being in a [s w] or [w s] relationship with another tone always provides extra oomph. However, at least in the case of Yoruba, there is an interesting possibility that connects tone spreading with tone dissimilation.

(21)a. Undershoot in HM



b. Less undershoot in HL with spread target



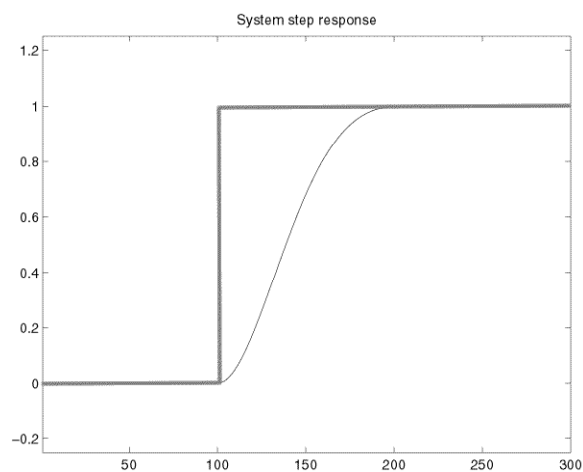
This idea depends on the observation, going back at least to Ohman (1967), that a simple linear model of co-articulation as smoothing shows a somewhat lifelike undershoot of articulatory targets, as long as the time constant of the low-pass filter is large enough relative to the time scale of sequence of target values in the input. We can be sure that the inertia of speech articulators creates a low-pass effect, and that the time constant involved is large enough to create such undershoot. In this situation, spreading

out the target value in the input will certainly decrease or even eliminate the amount of undershoot.

The two plots in (21) demonstrate this effect. Each show a VCV sequence, with vertical lines marking the start and end of the medial C, which has a duration of 70 milliseconds. In both plots, we start with a mid-valued tonal target at the beginning of the first vowel. (21a) shows a HM sequence: in this case, the H target is at the end of the first V, whereas the M is implemented by supplying mid-valued targets for both the initial and the final target positions of the second vowel. (21b) shows a HL sequence: in this case, the H target spreads to occupy both the final position of the first vowel, and the initial position of the second vowel, while the L target applies to the final position of the second vowel. Straight-line interpolation among these input values creates the thick lines, representing the inputs. These inputs are thus a literal phonetic translation of the phonological representations we have proposed.

The inputs are then smoothed with a low-pass filter whose step response, shown in (22) below, requires roughly 100 milliseconds. The thin lines in (21a) and (21b) are the results. Unsurprisingly, the input with the spread-out target is released from undershoot, and thus appears to show dissimilation (here H-before-L raising).

(22) Step response of the low-pass filter

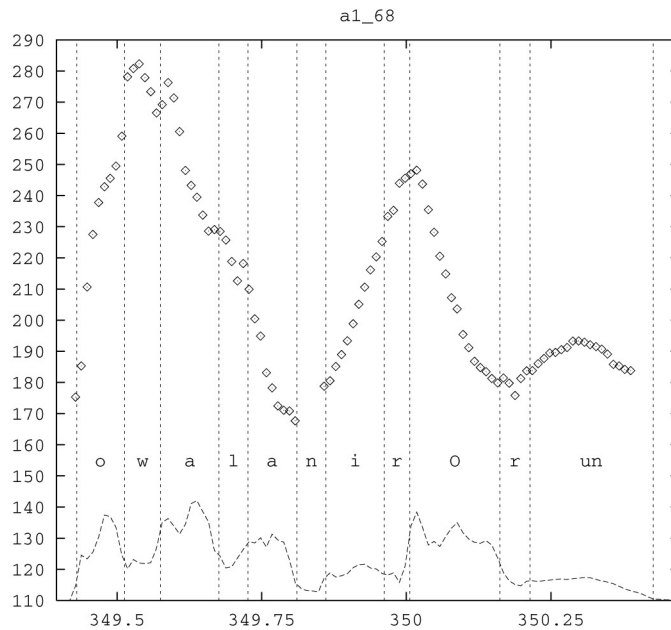


3.2 A remark on sparse tone association

In addition to the phonetic similarities that we have cited between Yoruba tonal complexes and Japanese accent, we might remark in passing on the issue of sparse vs. dense tonal association. Pierrehumbert and Beckman (1988) argue at some length that tonal association in Japanese is sparse, in the sense that there may be arbitrarily long stretches of moras without any tonal association. Their primary argument is a phonetic one: they show that these unspecified regions tend to show gradual transitions between

the F0 target created by one associated tone to the F0 target created by the next one, with the slope of the transition depending on its duration.

(23)

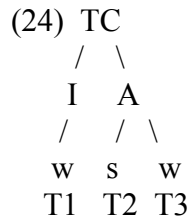


ó wà là ní rọ̀ run 'he drove Alani to heaven'
H (L) H L M

Our analysis of Yoruba requires tonally unspecified regions for phonological reasons – that is our analysis of the Mid tone. However, we want to point out that stretches of material that is all tonally specified, in languages like Igbo and Yoruba, may also show exactly the kind of gradient transitions that Pierrehumbert and Beckman found in Japanese. This is because (as a factual matter) the tonal targets are placed at the ends of each stretch of like-toned material: according to our analysis, a single tonal complex, with one or two tonal target points, is associated with the whole stretch. We showed this before for Igbo in (5); for a simple example in Yoruba, see the F0 track in (23), where a fall is distributed over the two L syllables wà là.

4.0 The shape of a Tonal Complex

We speculate that the true form of a tonal complex may actually be something like (24):



which is analogous to a CVC syllable.

Continuing the analogy, the units we have been looking at are like CV syllables. It would make sense that these are apparently the common and default case in Benue-Congo languages. The fuller structure gives us a way to deal with two-tone contours, which seems to be the most complex that are ever found, and also to explain why such contours are more marked.

There is some evidence for two-tone contours with a [s w] pattern – in fact this is the normal case for pitch accents, it seems. This would be analogous to a VC syllable, or perhaps better to a diphthongal vowel with an off-glide. It would be interesting to know if there are any level-tone languages where this is the default pattern for tonal complexes, thus leading to normal placement of F0 targets at the start of the tone-bearing region, and also to “anti Benue Congo” contour shifting, where F0 rises and falls shift to happen earlier in time) instead of begin delayed. We speculate that such cases will not be found, if only because the syllable-initial placement of F0 targets would probably have been noticed.

Finally, this structure gives us a way to deal with the “cumulative” tonal/accental pattern of languages like Japanese and Mawukakan, where a maximally three-tone pattern is associated with the whole word. It remains to be explained why in those two cases the tone sequence is limited to L+HL: are there otherwise similar languages with a wider range of cumulative tonal patterns?

References

- Akinlabi, A. M. 1985. *Tonal Underspecification and Yoruba Tone*. Ph.D. dissertation, University of Ibadan.
- Akinlabi, A. M. and Y. Laniran 1988. "Tone and Intonation in Yoruba Declarative Sentences". 19th Annual Conference on African Linguistics, Boston University.
- Akinlabi, A. and M. Liberman 1995. "On the Phonetic Interpretation of the Yoruba Tonal System," *Proceedings of the International Congress of Phonetic Sciences 1995*, Stockholm, Sweden. pp. 42-45.
- Akinlabi, A. and Liberman, M. In progress. "The sound structure of Yoruba tones". Ms. Rutgers University and University of Pennsylvania.
- Bamba, Moussa 1991. *De l'interaction entre tons et accent*. Ph.D. dissertation, Université de Québec a Montreal.
- Bamgbose, A. 1966. "The assimilated low tone in Yoruba," *Lingua* 16, 1-13.
- Connell, Bruce and Robert Ladd 1990. "Aspects of pitch realisation in Yoruba," *Phonology* 7, 1-29.
- Elugbe, B. O. 1985. "The tone system of Ghotuo," *Cambridge papers in Phonetics and Experimental Linguistics* 4, 1-21.
- Elugbe, B. O. 1995. "The assimilated low tone in Ghotuo," in *Language in Nigeria: Essays in honour of Ayo Bamgbose*, ed. by Kola Owolabi. Ibadan: Group Publishers. 68-74.
- Goldsmith, John 1976. *Autosegmental phonology*. IULC and Garland Press, New York.
- Hyman, Larry and Armindo Ngunga 1994. On the non-universality of tonal association conventions: evidence from Ciyao. *Phonology* 11. 25-68.
- Hyman, Larry and Russel Schuh 1974. "Universals of tone rules: evidence from West Africa". *Linguistic Inquiry* 5, 81-115.
- Kubozono, Haruo 1993. *The organization of Japanese prosody*. Tokyo: Kuroshio Publishers.
- Laniran, Yetunde 1992. *Intonation in Tone Languages: The Phonetic Implementation of Tones in Yoruba*. Ph.D dissertation, Cornell University.
- LaVelle, C. R. 1974. "An experimental study of Yoruba tone," *Studies in African Linguistics*. Supplement 2, 183-200.
- Leben, William 1973. *Suprasegmental phonology*. MIT Ph.D. Dissertation. Published by IULC.
- Liberman, Mark, J. M. Schultz et al. 1993. "The phonetic interpretation of tone in Igbo". *Phonetica* 50(3), 147-160.
- Liberman, Mark. 1995. "The sound structure of Mawu words". In Gleitman and Liberman, eds., *Invitation to Cognitive Science*, MIT Press. Pp. 55-86.
- Ohman, Sven. 1967. "Numerical Model of Coarticulation". *JASA* 41: 310-320.
- Pierrehumbert, Janet. *The Phonology and Phonetics of English Intonation*. MIT Ph.D. Dissertation. Published by IULC.

- Pierrehumbert, Janet and Shirley Steele. 1987. "How Many Rise-Fall-Rise Contours?" In Proceedings of the 11th International Congress of Phonetic Sciences, Vol. 3, 145-147. Estonian Academy of Sciences, Tallinn.
- Pierrehumbert, Janet and Mary Beckman. 1988. Japanese tone structure. MIT Press.
- Poser, W.J. 1985. The Phonetics and Phonology of Tone and Intonation in Japanese. Doctoral dissertation, MIT.
- Pulleyblank, D. 1986. Tone in Lexical Phonology, Reidel, Dordrecht.
- Ward, Ida 1952. An Introduction to the Yoruba Language, W. Heffer and Sons Ltd., Cambridge.
- Williams, Edwin 1976. "Underlying tone in Margi and Igbo". Linguistic Inquiry 7, 463-484.
- Yip, Moira. 1989. "Contour Tones". Phonology 6: 149-174.

Acknowledgements

We would like to thank participants at the following conferences: the 2000 Tone Symposium in Tromsø, the 30th Colloquium on African Languages and Linguistics in Leiden, NELS 31, and the Rutgers Optimality Research Group; especially Roger Blench, Paul DeLacy, Chip Gerfen, Markus Hiller, Larry Hyman, Young-mee Yu Cho, David Odden, Alan Prince, Hubert Truckenbrodt, and Moira Yip.

Akin Akinlabi
Rutgers University
18 Seminary Place
New Brunswick, NJ 08901-1184

akinlabi@rci.rutgers.edu

Mark Liberman
609 Williams Hall
University of Pennsylvania
Philadelphia PA 19104-6305

myl@cis.upenn.edu

