Phonetica 1993; 50: 147–160

*Mark Liberman*
*J. Michael Schultz*
*Soonhyun Hong*
*Vincent Okeke*

Department of Linguistics,
University of Pennsylvania,
Philadelphia, Pa., USA

# The Phonetic Interpretation of Tone in Igbo

**Abstract**

This article presents a preliminary study of the phonetic interpretation of tone in Igbo, a Kwa language of southeastern Nigeria. The experimental method varies the speaker's pitch range orthogonally with variation in tonal material, and fits a model to measurements of maxima and minima in the resulting $F_0$ contour. A new interpretation of *downstep* is proposed as a result.

## Introduction

Igbo, a language of the Kwa branch of Niger-Congo family, is spoken by about 15 million people in southeastern Nigeria. Its phonology, morphology and syntax have been widely studied [Clark, 1990; Manfredi, 1991], especially with reference to the intricate patterning of lexical tone. This article is a preliminary study of the phonetic interpretation of Igbo tone. We use an experimental method first applied to English [Liberman and Pierrehumbert, 1979, 1984], in which a speaker varies pitch range orthogonally with variation in tonal material, and we compare the success of different models in characterizing the interaction of tone identity, phrasal position, tone sequence, and pitch range in determining patterns of measured $F_0$ values. From the statistical structure of these data, we draw several conclusions about Igbo tone and its phonetic interpretation.

## Lexical Tone in Igbo

In this section, we will introduce the system of lexical tone in Igbo, especially as it relates to the design and interpretation of the experiments we have done. We pass over most of Igbo's intricate and fascinating tonal phonology, since our immediate concern is with the

A shorter version of this article was published as Liberman et al. [1992].

Received: March 16, 1993
Accepted: May 23, 1993

Mark Liberman
Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104–6305 (USA)

**Table 1.** Typical $F_0$ values in downdrift

| Text: | A | ba | no | bi | ọma |
|---|---|---|---|---|---|
| Tones: | H | L | H | L | H |
| $F_0$ | 237 | 178 | 210 | 158 | 181 |

phonetic interpretation of the surface tonal categories. However, the basic nature of these surface categories will emerge as a crucial question.

An Igbo syllable can have one of three tonal categories, known as high (H), low (L), mid (M). The H and L tones occur freely, but the M tone can only occur following an H tone or another M tone. Thus there are five possible tone patterns for two-syllable words: (1) HH, (2) HL, (3) LH, (4) LL, (5) HM. Monosyllabic words are rare, and all of those that can occur in isolation have high tone.

Igbo speakers are of course free to change their pitch range, and do so for many reasons. We may ask whether tonal distinctions are nevertheless maintained – for instance, is HH in a low pitch range distinguished from LL in a higher one? Are HM and HL kept distinct? If so, how? Our experiments provide a tentative answer: contextual effects in the system of tonal interpretation largely maintain the distinctions among bitonal patterns as pitch range varies.

There are several general effects that modify the realization of Igbo tones in phrasal context. One of the most important of these is know as 'downdrift', which progressively lowers H and L tones when they occur in sequence. Thus on one typical pronunciation of the phrase Ábànóbì ọ́ma 'good Abanobi', the successive minimum and maximum $F_0$ values were as shown in table 1.

A model for the phonetic interpretation of Igbo tone must obviously take account of downdrift. The term 'downstep' is used to refer to a different circumstance. Although some phonologists have treated Igbo mid tones as a third tonal category, distinct from high and low [Carrell, 1970; Goldsmith, 1976], others [Clark, 1990] have interpreted the restricted distribution of the mid tone to mean that it is actually just a high tone that happens to be downstepped, i. e. realized at a lower pitch. On this view, HM is actually H$^!$H, where the raised exclamation point marks the downstep location. This situation (or its counterpart on other tone languages) has been given various phonological interpretations:

(1) the downstep marker may be reified as a separate phonological entity, as suggested by the exclamation point, diacritic;

(2) downstep may be viewed as the consequence of a 'floating' low tone between the two high tones, which thus triggers lowering by the more general process of downdrift, but is not otherwise realized [Clements and Ford, 1979; Pulleyblank, 1986];

(3) the downstepped sequence HM may be viewed as the expression of two distinct high tones, whereas an HH sequence is viewed as a single high tone spread over two syllables [Clark, 1990].

One obvious question is whether downdrift and downstep are phonetically the same, as we might expect from the second hypothesis given above. One of our experiments suggests that Igbo downdrift and downstep are phonetically different, in a way that may be interpreted to support a variant of Clark's [1990] theory.

## Design of the Experiments

We will mark Igbo tone using the system of Welmers [1976] and Nwachukwu [1983], which expresses the concept of downstep made explicit in Clark's [1990] theory: high is written with an acute accent; low is written with a grave accent; an unmarked syllable continues the previous tone, and a repeated high tone mark is interpreted as mid. In this system, as in all the similar tonal orthographies known to us, downdrift is not marked, because it is assumed to be automatic (or if variable, only expressively so).

In this notation, the materials for our first experiment consisted of the 13 words:

| | | | |
|---|---|---|---|
| HH: | ísi 'head' | óke 'male' | íre 'tongue' |
| HL: | ísì 'odor' | ókè 'boundary' | írè 'to be effective' |
| LH: | ìsí 'six' | òké 'rat' | |
| LL: | ìsi 'blindness' | òke 'share' | ìre effectiveness' |
| HM: | ísí 'to cook' | | íré 'to sell' |

On each trial, the subject was asked to read a word or phrase in one of three modes: addressed quietly to someone seated nearby; addressed to someone seated on the other side of a broad table, a little more than a meter away, or addressed to someone at the other end of a room, about 10 meters away. 195 utterances (five repetitions of each phrase in each mode) were elicited in random order. The subject was a man aged 45, from the village of Awo-Omamma in Imo State, about halfway between Owerri and Onitsha.

This procedure produces a good deal of pitch range variation, of the kind involved in thus 'raising' or 'lowering' the voice. Comparable measurement points in different tokens of the same tonal type have $F_0$ values up to about an octave apart, which is several times larger than the difference between lexically distinct tone categories in a given pitch range. As a result, we often find (for instance) that an initial L tone in a wide-pitch-range utterance is actually higher than an initial H tone in a narrow-pitch-range utterance.

Loudness and duration also vary in an experiment of this type, giving an independent indication of the speaker's level of vocal effort. We would like to point out that there are many other functional dimensions that are often associated with $F_0$ effects that could be described in terms of pitch range variation. Examples include the speaker's overall level of arousal, the topic structure of the discourse, and the relative prominence of particular words and phrases. It should not be assumed that all such $F_0$ effects will show the same patterns as the effects we have studied here, which are linked with the distance of an interlocutor.

We will not model the duration and amplitude effects of the range variation that we have induced, nor will we try to model all aspects of the $F_0$ contour. We will represent the pitch of each syllable by a single value, taken automatically from the $F_0$ time function. We have tried various definitions for this representative value, including the value at the syllable midpoint and the average or median value; we find that the general structure of the data is the same, but the cleanest patterns result when we pick the maximum value for H tones, and the minimum value for L tones. We will discuss this point somewhat further when we consider the question of whether high tones are raised before low tones.

Like many (but apparently not all) tone languages, Igbo shows a pattern in which tones occurring later in a sequence are sometimes lowered relative to the values for the tones occurring earlier. In order to model these downtrends, we need to look at $F_0$ patterns in longer sequences of tones. Thus our second experiment used the 13 phrases shown below, each produced six times in each of the three pitch-range modes. These phrases are personal names, or concatenations of personal names with *nà* 'and', a consctruction chosen because it is semantically flat, and also fails to induce the complex tone changes that occur with many syntactic combinations.

| | | |
|---|---|---|
| 1 | Díké nà Áma | HM L HH |
| 2 | Íke nà Áma | HH L HH |
| 3 | Íbè na Áma | HL L HH |
| 4 | Ńne Ụ́ba nà Íbè | HH MH L HL |
| 5 | Ụ́gwụ nà Íbè | HH L HL |
| 6 | Áma nà Íbè | HH L HL |
| 7 | Áma nà Íke | HH L HH |
| 8 | Íke nà Áma nà Ába | HH L HH L HH |
| 9 | Ọ́nụ́ ọma | HM HH |
| 10 | Ọ́lụ̀ká ọma | HLH HH |
| 11 | Ọ́nwụká ọma | HHM HH |
| 12 | Ábànóbì ọ́ma | HLHL HH |
| 13 | Ábàríkwú ọma | HLHM HH |

The term 'downdrift' applies to the regular lowering of H and L in alternating sequence, while 'downstep' is used for the so-called 'mid' tone, which is thus considered to be a lowered version of H. The materials in experiment 2 are designed to let us characterize and compare these phenomena.

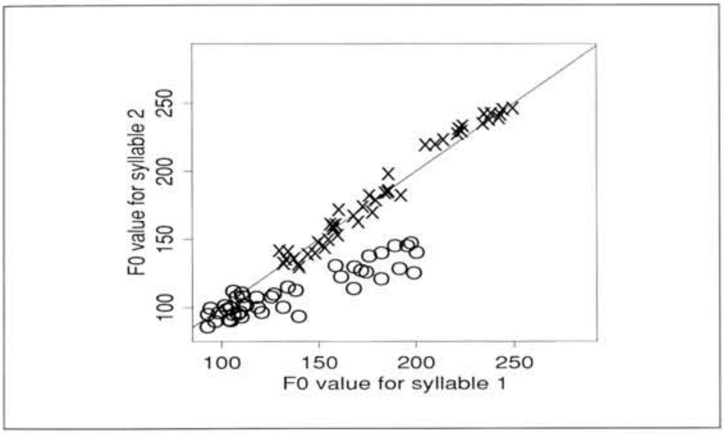


**Fig. 1.** Igbo HH (X) LL (○) disyllables.

### Results: Experiment 1

Figure 1 shows some of the data from experiment 1, as a scatter plot in which the X axis shows the $F_0$ value of the first syllable, and the Y axis shows the $F_0$ value of the second syllable. The $y = x$ line is plotted as well. We can see several things in figure 1. First, although the points span a broad range, their relationship is quite a tight one: the correlation of the HH measurements is 0.988, while the correlation of the LL measurements is 0.905. Second, the two H tones are generally about equal in value, but the second of the two L tones is usually lower than the first. Functionally, this helps distinguish an HH sequence in a low pitch range from an LL sequence in a high pitch range. We can express the same difference between the HH and LL data by saying that the LL trend is best fitted by a line with a slope less than 0.5 and a nonzero intercept, while the HH data is not statistically distinguishable from $y = x$. Regression on the HH data gives a slope of 1.0 (standard error 0.025) and an intercept of –6.8 (S. E. 4.7). Regression on the LL data gives a slope of 0.44 (S. E. 0.032), intercept 51.9 (S. E. 4.5).

Figure 2 shows that the H/L proportion is systematically greater in the HL order than in the LH order, consistent with the fact that L is lower in the second position of LL sequences. Also, the HL data clearly requires a nonzero intercept, while the LH data does not. Thus regression on the HL data gives a slope of 0.25 (S. E. 0.03), with an intercept of 60 (S. E. 6.1), while the LH data gives a slope of 1.3 (S. E. 0.08), with an intercept of –5.7 (S. E. 12). The HM data gives a slope of 0.92 (S. E. 0.03), and an intercept of –5.6 (S. E. 5.4). The fact that the HM slope is so much greater than the LL slope calls some what into question the interesting suggestion of Stewart [1971] that the lowering of final low syllables is to be interpreted as 'another manifestation of key lowering', a cover term in which he includes downstep.

The patterns in the data of experiment 1 suggest a model in which the $F_0$ is predicted to be T+TRFD. Here T is a factor that has one value for H and another for L; F is a factor that is < 1 for L in second position and 1 otherwise; D is a factor that is < 1 for a 'downstepped high' (here mid) tone, and R is 'latent variable' representing the pitch range of a

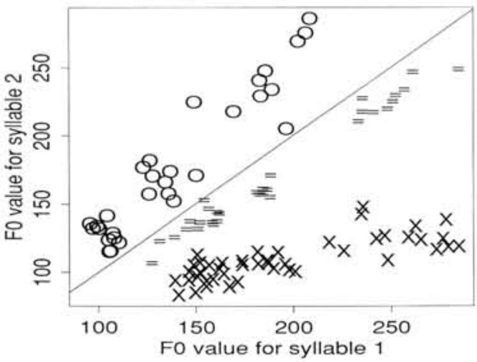


**Fig. 2.** Igbo HL (X) LH (○) HM (=) disyllables.

given utterance. This model yields the following expressions for predicting the syllable 2 pitch (y) from the syllable 1 pitch (x) in a given utterance:

| | |
|---|---|
| HH | $y = x$ |
| HL | $y = (FL/H)x + (1-F)L$ |
| LH | $y = (H/L)x$ |
| LL | $y = Fx + (1-F)L$ |
| HM | $y = Dx + (1-D)H$ |

The model suggests values for the L and H parameters which are quite reasonable, given the speaker's observed $F_0$ in low-pitch-range utterances. The model also predicts significant intercept terms in the LL and HL cases, where we did find them, and no intercepts in the HH and LH cases, where we did not. Unfortunately, it predicts an intercept in the HM case as well, where we did not find one; the predicted intercept is small, since 1–D appears to be small, but this may point to a problem in the model. (T+TRF)D would predict no intercept for the HM case; space does not permit further exploration.

We should point out that this model embodies some probably wrong assumptions about the nature and environment of the lowering conditioned by D, namely that (in disyllables) it applies only to the HM sequence, and not (for instance) to the L in the HL sequence.

Having decided on the structure of our model, we can optimize its parameters with respect to the whole data set. (Treating the R parameters as latent variables, we use the downhill simplex method [Press et al., 1988] to perform the optimization.) Doing this on the data of experiment 1, assuming the T+TRFD model with D applying only to the HM case, we get parameter estimates H=112, L=86, F=0.42, D=0.80, and an RMS error across all the data of 6.2 Hz. Figure 3 plots the result of replacing each data point with its model counterpart. In this plot, the basic features of the data set are well captured. Given space limitations, we leave it to the readers to convince themselves that certain other simple functional forms for the model, such as TRFD, T+R+F+D, and so on, produce qualitatively wrong predictions. Their overall prediction error is significantly higher, but just as important, some of the key qualitative aspects of the

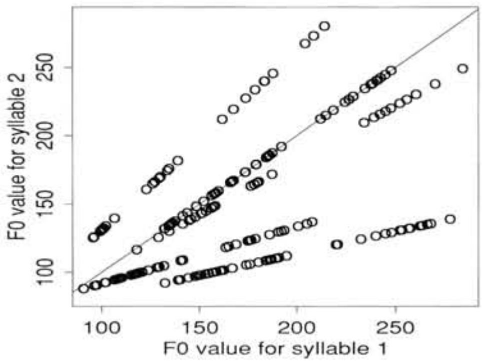


**Fig. 3.** Model values for Igbo disyllables.

data are misrepresented. We find it especially interesting that the all-multiplicative model makes such obviously wrong predictions, since this failure calls into question the popular practice of using a semitone scale for $F_0$ contours.

### Results: Experiment 2

The second experiment permits us to explore some other interconnected questions about the phonetics of Igbo tone. For instance, we can ask whether downdrift affects both H and L tones in the same way; whether downstep and downdrift are the same, and whether H is raised in front of L.

*Raising of H before L*

We take up the last point first, since it affects the measurements used in evaluating the other issues. In two independent studies of Yoruba, Connell and Ladd [1990] and Laniran [1992] suggest that H tones are raised before L. They offer various arguments, notably the fact that the last H in sequences such as HHL, HHHL, HHHHL, etc. is markedly higher than the earlier occurrences, and that this effect is not seen in sequences such as HHM, HHHM, HHHHM, etc.

There are no exactly comparable cases in Igbo, since Yoruba has a true, freely distributed M tone, whereas Igbo's M tone is a downstepped H. Still, we ought to be able to see the effect in Igbo, mutatis mutandis. Some examples seem to suggest that raising of H before L also holds in Igbo, as figure 4 clearly exemplifies. This figure shows the phrase *Áma nà Íke* (tone pattern HH L HH) spoken in the version with a distant interlocutor (and thus a high and broad pitch range). We can clearly see that the second syllable of *Áma* is much higher than the first.

However, we are not convinced that this fact should be interpreted in terms of a rule of high-before-low raising. Essentially the same pattern for *Áma* would be possible in the absence of a following L tone. Indeed, the presence of more than one H syllable is by no means required. For instance, the $F_0$ track (shown in fig. 5) for the isolated monosyllable *yá* 'he' shows a rising contour that is rather similar to the rising pattern seen on *Áma* in figure 4, and quite similar rising patterns could

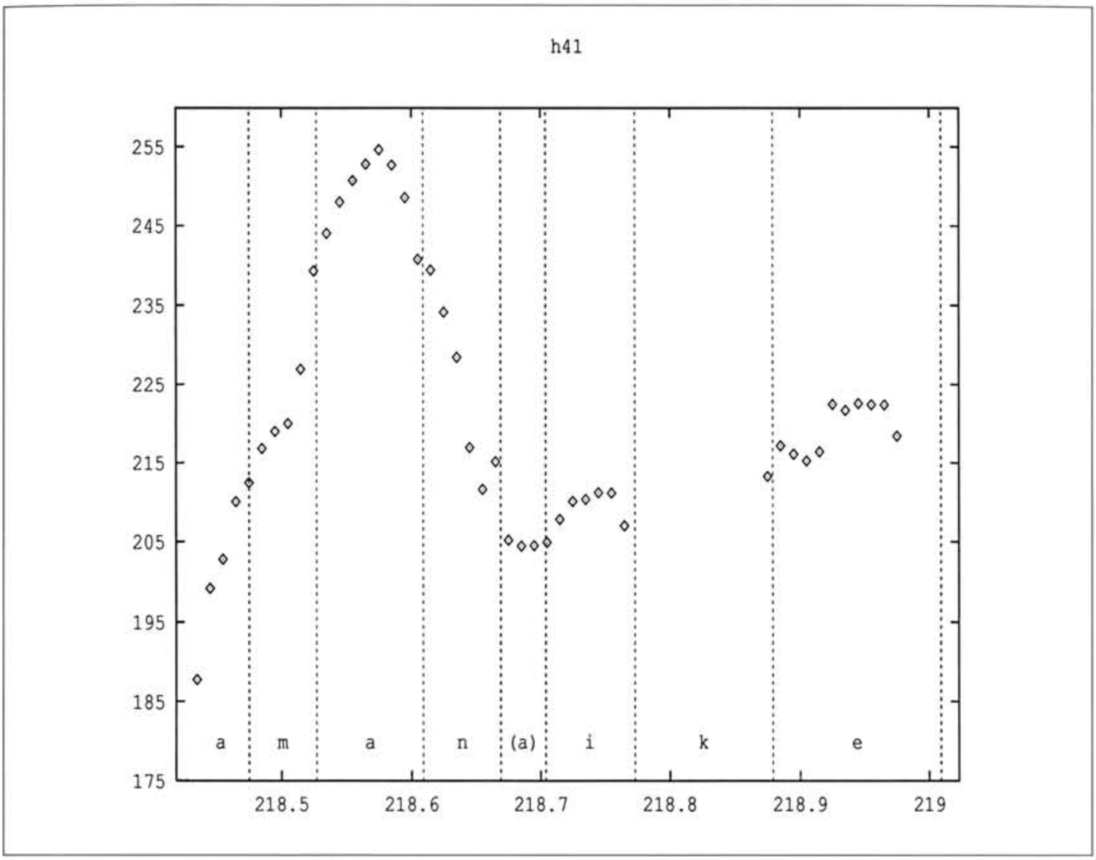


**Fig. 4.** *Áma nà Íke* (far interlocutor).

be seen in utterances that begin with three H syllables.

In general, the starting $F_0$ for utterance-initial H tone stretches in Igbo is variable, but almost always lower than the peak $F_0$, which is usually not reached until near the end of the sequence of H tone syllables. It is not clear what governs the variability in this matter, but stretches of different numbers of H syllables show qualitatively similar patterns, whether in isolation, before L, or before M.

Furthermore, this behavior seems to be a particular case of a much more general non-equivalence of $F_0$ values realizing adjacent equal tones. Similar phenomena in Yoruba are extensively discussed by Laniran [1992]. In our much more limited examination of the Igbo case, these phenomena seem consistent with a system in which a block of contiguous 'same-toned' syllables (or other tone-bearing units) have just the same number and type of $F_0$ targets that a single syllable in the same context would have. This view is a sort of phonetic version of the phonological principle known as the obligatory contour principle (OCP). The phonological OCP has the consequence that a stretch of segments with the same value of a phonological feature must

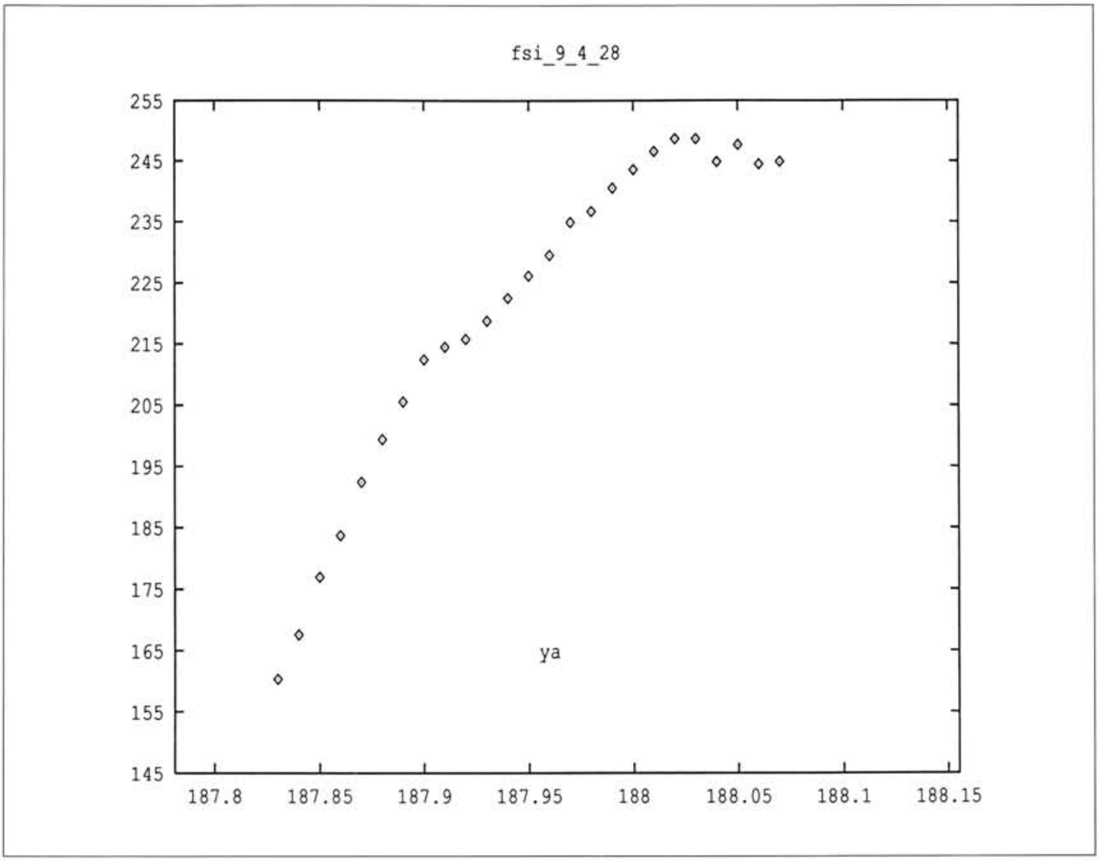


**Fig. 5.** *yà* (FSI 9.4).

represent a single feature spread over the whole sequence. Our phonetic version of this principle suggests that a tone spread over a stretch of tone-bearing units will only be interpreted phonetically one time, regardless of the nature and quantity of the units that it is associated with (abstracting away from coarticulatory effects that may arise when targets are crowded too close together, or physical interactions with other ongoing articulations). This hypothesis contrasts with systems in which each tone-bearing unit in such a block is subject independently to $F_0$ interpretation.

Laniran's [1992] treatment of Yoruba is intermediate between such a 'phonetic OCP' system and a system in which each tone-bearing unit is given independent $F_0$ interpretation. The raising of the last H in a sequence when an L tone follows is one key case in which additional targets seem to be required. It seems to us that the Yoruba examples might be amenable to a treatment in which the H tone targets in the sequences HL, HHL, HHHL, etc. have the same number and the same timing principles regardless of the count of H tones.

Note that we accept Laniran's [1992] argument that the H tone target is raised in Yoruba

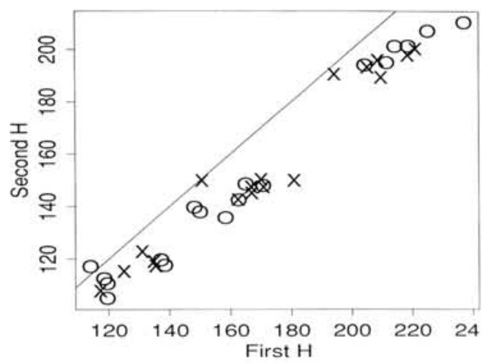


**Fig. 6.** H/H relationship in HLHL (○) vs. HLHM (X).

before L vs. M. However, our Igbo data provide quite a strong argument against the view that a similar process operates there. This argument arises from comparing the relationship of the two H tone targets in HLHL sequences and in HLHM sequences. If H were raised before L to anything like the same extent that Laniran [1992] suggests for the case of Yoruba, the relationships should be significantly different in the two cases.

Figure 6 presents a scatter plot of this relationship in our data. As this figure suggests, the 2 cases are not statistically distinguishable. This makes it rather unlikely that Igbo has a general rule of H raising before L and suggests that the apparent existence of such an effect in examples like figure 4 must be explained in terms of other $F_0$ interpretation principles.

Applying this assumption, we will treat patterns such as HLH, HHLHH, and HLLHH in the same way for purposes of evaluating the scaling of downdrift, choosing the $F_0$ minima and maxima (wherever they occur) as the points of reference.

*Effect of Downdrift on H and L*

Does downdrift affect H and L tones in the same way? The answer appears to be 'yes' in Igbo. Figure 7 shows that successive L tones in an HLHL sequence are clearly lowered. Furthermore, the relationship between pairs of downdrifting H tones in an HLHL sequence and the relationship between pairs of downdrifting L tones further along in the same sequence appear to be the same. Naturally H tones are higher than L tones, other things equal, and so the L tone pairs tend to be drawn from a lower region of the distribution – but the distribution seems to be pretty much the same for both kinds of tone pairs.

*Downstep vs. Downdrift*

Is downstep (the lowering of M tones after H) the same as downdrift (the lowering of the second H in an HLḤ sequence)? Our data suggest that it is not. We take downdrift data from phrases 2, 3, 5, 6, 7, 8, 10, 12 and 13 of experiment 2, using only the first downdrift in phrase 8. We take downstep data from the HM words

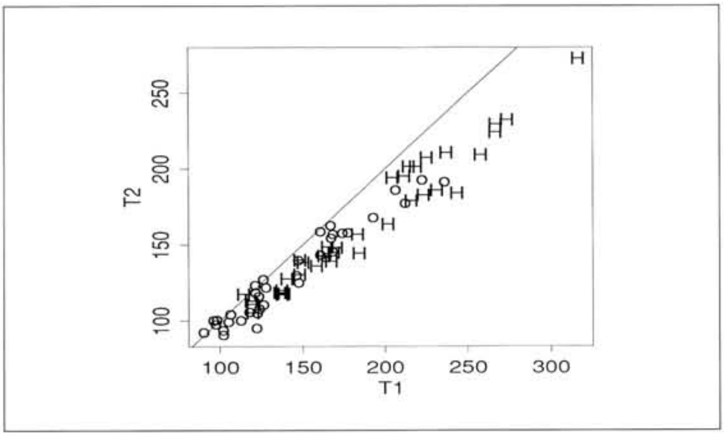


**Fig. 7.** Downdrift on L (o) and H (H).

**Table 2.** Regression coefficients for downstep and downdrift

| | Intercept | SE | Slope | SE | n |
|---|---|---|---|---|---|
| HLH | 7.65 | 2.74 | 0.824 | 0.014 | 162 |
| HM | −2.30 | 4.14 | 0.924 | 0.021 | 85 |

of experiment 1, and from phrases 1, 9 and 11 of experiment 2. As before, the $F_0$ relations are remarkably homogeneous and well controlled: r = 0.978 for the downdrifts, and 0.980 for the downsteps. The slopes and intercepts derived from regression, and the standard errors of the estimates, are given in table 2. It is pretty clear that the slopes are different in the two cases, and indeed it seems that downdrift imposes roughly twice as much lowering as downstep does. Thus we can provisionally reject the hypothesis that (from a phonetic point of view) downstep is 'the same thing' as downdrift.

This is a surprise, since there have been argued to be good typological, historical and phonological reasons to equate the two processes. We see three possible explanations: to accept that the processes are distinct; to salvage their phonological equivalence by excusing their phonetic distinctness on some independent grounds; or to take the view that downdrift is two units of downstep. This last move strikes us as the most interesting one. To make its content clearer, we return to our simple-minded model T+TRFD. If the D parameter is to be used to model either downstep or downdrift, it will have to take on a sequence of successively lower values as we accumulate lowering in a phrase like HLHLHLH or HMMMM. The obvious way to do this is to write the formula as $T+TRFD^N$, where N starts at 0 and increments by 1 for each unit of lowering. Then we might increment N in several different ways, among them those shown in table 3.

The first two are traditional 'downstep equals downdrift' theories, differing in whether the lowering occurs on the L tone or on the H tone. Phonologists who have explicitly considered the issue have generally opted

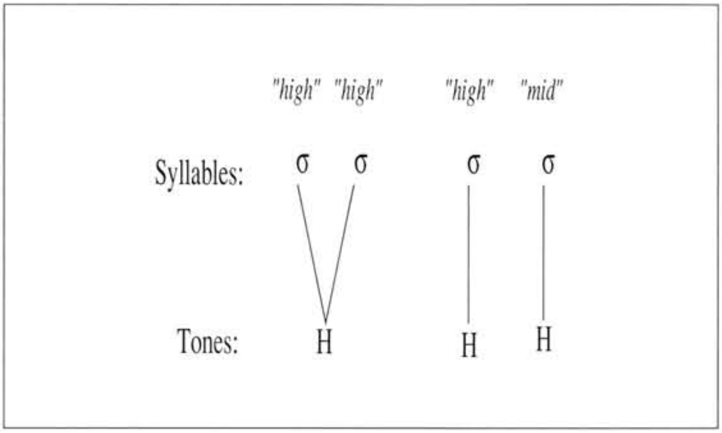


**Fig. 8.** Clark's [1990] theory of the Igbo mid tone.

**Table 3.** Ways to increment N in downstep and downdrift

| | Tones | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | H | L | H | L | H | L | H | H | M |
| N. 1 | 0 | 1 | 1 | 2 | 2 | 3 | 3 | 0 | 1 |
| N. 2 | 0 | 0 | 1 | 1 | 2 | 2 | 3 | 0 | 1 |
| N. 3 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 |

for N. 1 [Clark, 1990; Goldsmith, 1976]. The last idea (labelled N. 3) is an example of a way of counting that makes a downdrift worth two downsteps – it says, basically, that N counts the distinct (in the OCP sense) tones in the string. As far as we know, it has not previously been suggested.

Note that all three of these ways of counting N predict that H and L tones 'see' downdrift in the same way. They differ in the predictions they make about the relationship between downdrift and downstep, and the specific relationship between H and L tones (although the last point can only be explored given independent constraints on the functional form of the model, and on the other model parameters). Production scaling experiments, like those discussed in this article, offer a good opportunity to compare such hypotheses in a quantitative way. Space does not permit a consideration of this comparison here, but it represents an interesting example of how phonetic evidence can be brought to bear on what has been taken to be a phonological issue.

As we noted earlier, there are several different ideas in the literature about the nature of downstep (i. e. 'mid' tone) in Igbo. Some authors treat downstep as a case of downdrift in which the L tone between two H tones is 'floating' (i. e. unassociated with any segmental material), and as a result is not realized phonetically except by virtue of causing the H tone that follows to be lowered. Others argue against this treatment and suggest that the 'mid' tone is either a third phonological category, or else simply an independent H tone, which is interpreted phonetically at a lower pitch value than the H tone that precedes. This point of view [suggested notably in Clark, 1990] is represented graphically in figure 8.

Clark's [1990] arguments against the 'floating low tone' theory of Igbo downstep are entirely phonological in character. She points to the regular appearance of a downstep at the boundary between cyclic domains when we would expect the tones on each side to be simply H, and also the regular disappearance of a downstep in environments that can be simply characterized in terms of a rule 'delete the middle of three adjacent H tones'. She argues that both these classes of phenomena require unmotivated complexities to state if downstep in Igbo is treated in terms of the presence of a floating L tone.

Clark [1990] is still able to maintain the identity of downstep and downdrift:[1] she expresses them as a rule of register lowering, which says: 'Lower the high (and low) pitch registers at the juncture between a high tone and any following tone within the same phrase.'

Our evidence with respect to this question is purely phonetic and somewhat indirect. First, we observe that downstep and downdrift do not lower an H tone by the same amount; thus the floating L tone theory, which treats downstep as downdrift of H tones across an invisible L, must provide some excuse for this difference in values. Second, we observe that L tones are downdrifted in exactly the same manner, quantitatively, as H tones. Finally, we observe that downstep lowers successive H tones by about half as much as downdrift does. All of this comes together nicely if each successive new tone, whether H or L, causes the whole tonal system to 'deflate' by a fixed amount. Then both downstep and downdrift are just the tick of passing tones.

[1] The main reason to maintain this equivalence seems to be a typological one. Downstep and downdrift, both very common but not ubiquitous, apparently have an implicational relationship – two-tone languages with downstep always have downdrift, but not vice versa.

Our version of Clark's register-lowering rule says something like: 'Increment the exponent of the D parameter whenever a new tone occurs.' It is equivalent, but more palatable psychologically, to maintain the current D value by successive multiplications, and reset it (partly or completely) at phrasal boundaries. Note that this formulation of the 'deflation' principle differs from Clark's [1990], in that hers still maintains the phonetic equivalence of downstep and downdrift. However, our formulation is consistent with Clark's phonological analysis, and inconsistent with the floating-L-tone analysis. In that sense, our findings support her position.

There is a small remaining problem: is H lowered after a phrase-initial L? Emenanjo [1978] says that it is not, and Hyman and Schuh [1974] claim that a dialect difference exists on this point. Since the amount of lowering involved would not be very great, and pitch values are also affected by overall pitch range, vowel type, etc., it is not clear how to tell if lowering occurs. Note that according to our model, the predicted effect (the change in $F_0$ value of the first H tone in a phrase due to the presence or absence of a preceding L tone) is about 20% of the first H tone's difference from the base of the H register. In conversational speech for a typical male speaker, this difference might be about 20 or 30 Hz, so that the predicted effect will be in the range of 4–6 Hz, which could easily be missed. We do not belive that our data provides any way to check this question. If it turns out that H is not 'deflated' at all following initial L, then our rule would have to be modified to exclude this case.

A more fundamental question is: why are there no downstepped low tones? Perhaps the lowered final L tones should be analyzed as a kind of downstep, but even if this is true, our treatment of L and H is not at all parallel. The floating-low-tone account of downstep offers a reason for this asymmetry, but at too great a phonological and phonetic cost. In effect, our account (in common with Clark's) requires that high tones are exempt from OCP restrictions in certain cases, while low tones never are.

There are many treatments for this state of affairs, among them a metrical account such as that offered in Manfredi [1991, 1992]: H tones might be exempt from OCP violations just in case they are the heads of separate tonal 'feet'. Our experimental evidence does not bear directly on such explanations, but just helps to pose more clearly the problem that they aim to solve.

## Summary of Experimental Conclusions

From the intricate and well-controlled patterns that emerge when pitch range is varied against tonal sequence in Igbo, we derive a number of tentative conclusions. First, the scaling of Igbo tones requires a model that is neither purely multiplicative nor purely additive; we can describe it by saying that increasing pitch range adds to a basic H or L tone value, in units that are proportional to the basic value of the tone type in question. This produces relationships among tone values in sequence that are qualitatively different from those that would arise in purely multiplicative or additive models. Second, Igbo L tones appear to be lowered in final position. Third, Igbo H tone does not appear to be raised before L (as opposed to before M). Fourth, Igbo downdrifted H and L tones seem to behave identically. Fifth, downstep and downdrift are quantitatively distinct: downdrift imposes a significantly greater degree of lowering. This difference is consistent with treating downstep and downdrift as two symptoms of a process that deflates both high and low pitch registers by a constant amount every time a phonologically distinct tone is encountered.

These conclusions must be considered tentative for a number of reasons. We have looked at data from only 1 speaker; the task represents only one (artificial) style of speech; there are other ways to describe and explain each bit of evidence we have presented, specifically alternative statements of the environments and alternative functional forms for modeling. Still, we feel that our approach reveals things about the tonal system that would not come out without pitch range variation. We try to maximize (rather than minimize) such variation, and then use the rich statistical structure of the resulting data to distinguish among alternative hypotheses about the nature of the underlying system.

## References


Carrell, P. L.: A transformational grammar of Igbo. West African Language Monographs, No. 8 (University Press, Cambridge 1970).

Clark, M. M.: The tonal system of Igbo (Foris Publications, Dordrecht 1990).

Clements, G. N.; Ford, K. C.: Kukuyu tone shift and its synchronic consequences. Linguistic Inquiry *10:* 179–210 (1979).

Connell, B.; Ladd, D. R.: Aspects of pitch realization in Yoruba. Phonology *7:* 1–30 (1990).

Emananjọ, E. N.: Elements of modern Igbo grammar (Oxford University Press, Ibadan 1978).

Goldsmith, J.: Autosegmental phonology; MIT PhD thesis (1976).

Hyman, L. M.; Schuh, R. G.: Universals of tone rules: evidence from West Africa. Linguistic Inquiry *5:* 81–115 (1974).

Laniran, Y.: Intonation in tone languages: the phonetic implementation of tones in Yorùbá; PhD thesis Cornell University (1992).

Liberman, M. Y.; Pierrehumbert, J. B.: A metric for the height of certain pitch peaks in English. J. acoust. Soc. Am. *66:* suppl. 1, p. 5130 (1979).

Liberman, M. Y.; Pierrehumbert, J. B.: Intonational invariance under changes in pitch range and length; in Aronoff, Oehrle, Language sound structure, pp. 157–234 (MIT Press, Cambridge 1984).

Liberman, M. Y.; Schultz, J. M.; Hong, S.; Okeke, V.: The phonetics of Igbo tone, pp. 743–764, ICSLP 92, Banff, 1992.

Manfredi, V.: Agbo and Ehugbo: Igbo linguistic consciousness, its origins and limits; PhD thesis Harvard University (1991).

Manfredi, V.: The limits of downstep in Agbo sentence-prosody. IRCS Workshop on Prosody in Natural Speech, August 1992.

Nwachukwu, P. A.: Towards an Igbo literary standard (Kegan Paul, London 1983).

Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T.: Numerical recipes (Cambridge University Press, London 1988).

Pulleyblank, D.: Tone in lexical phonology (Reidel, Dordrecht 1986).

Stewart, J. M.: Niger-Congo, Kwa; Sebeok, Current trends in linguistics, vol. 7 (Mouton, The Hague 1971).

Welmers, W. E.: A grammar of Vai. University of California Publications, vol. 84 (University of California Press, Berkeley 1976).