1. Introduction

The place of Yorùbá prosody (or tonality) within Kwa and Niger-Congo has been a crux of the generative Africanist literature for three decades (Stewart 1965; Williamson 1968; Meussen 1970; Voorhoeve 1971; Welmers 1973, Hombert 1974). The problem boils down to whether typological relationships of this kind are insightfully stateable in terms of rules, à la Pulleyblank (1983), or of representations, as I will argue here.

§2 presents a list of problems, some framed in impressionistic or categorial terms (Ward 1952: 29-41; Abraham 1958: x-xxxi; Bâñgbọ̀se 1966A,B, 1970; Òwọbụlụyì 1970, 1987; Òyèlànà 1970; Welmers 1973; Sterk 1975; Akínlabì 1983A,B), others having been refined, revised or revealed instrumentally (LaVelle 1974; Hombert 1976; Akínlabì and Làníràn 1988; Làníràn 1988, 1992; Connell and Ladd 1990; Akínlabì and Liberman 1995). §3 presents the representational analysis which is sloganized in the title of this note.

If metrical licensing indeed captures prosodic generalizations of Yorùbá and related, prosodically different, languages at the systematic phonetic level, it supports two more general propositions as advanced by Kaye and Lowenstamm (1982) and Giegerich (1985) among others. (i) Prosodic constituents, like other phonological and syntactic domains, are licenced by the government relation. (ii) Construction of government domains in phonology, as in syntax, entails the existence of empty categories.

2. Rule-based puzzles

Early generative fixes on Yorùbá in tone typology held a standard SPE dichotomy between representations and rules, with limited attention to concepts of markedness or possible tone system. In that worldview, the matter posed itself as follows. Several Benue-Kwa languages (Àkàn, Èdó, Ígbọ̀, Èfìík-Ibibíò…) have at most two lexical tones (H and L), plus at least two tone rules (downstep and downdrift).\(^1\) Yorùbá, by contrast, happens to have three lexical tones (H, M, L) but just one of the rules (downstep, the “assimilated L tone” of Bâñgbọ̀se 1966). This picture is more or less explicit in Armstrong (1968), Courtenay (1968), Maddieson (ed. 1970) and many other contemporaneous writings.

For such a viewpoint, Yorùbá posed several intriguing questions, which have remained mysterious:\(^2\)

(a) Did Yorùbá gain a third lexical tone vis-à-vis its Benue-Kwa relatives Àkàn, Èdó, Ígbọ̀ and Èfìík-Ibibíò via a process of “tone-splitting” (Stahlke 1974, Maddieson 1974)? If so, is the ‘new’ tone Yorùbá’s present-day M or L?

(b) Did Yorùbá lose the strong, perseverative downdrift pattern of Àkàn/Èdó/Ígbọ̀/Èfìík-Ibibíò as a consequence of acquiring its third lexical tone (Ward 1952: 37ff.)? How did it nonetheless keep downstep (the “assimilated L tone”), if downstep presupposes downdrift à la Stewart (1965)\(^3\)?

(c) How can Yorùbá be distinguished from Êkóm-Yàlà (in the macro-Èdomà group of eastern Kwa), a language with three tones, downstep and also canonical downdrift (Armstrong 1968)?

---

\(^1\) Gbà (including Ètè, though ‘tonetically’ more complex than Àkàn or Ígbọ̀, like them has just two lexical tones (Stahlke 1971).

\(^2\) Perhaps because they involve ‘local’ typological comparisons within the Benue-Kwa family of Niger-Congo, these matters at the outset were posed diachronically, but soon were translated into synchronic terms, which is how I will treat them here.

\(^3\) Alongside Yorùbá, another Benue-Kwa example of downstep without downdrift is the Bantu language Kipare (Odden 1983).

_Avec une même structure grammaticale ne peuvent se combiner qu’un nombre limité des systèmes phonologiques._

– Troubetskoy

---

As to (1a), based on language-internal M~L alternations exemplified in (2–3), Stahlke (1974) suggests that Yorùbá M is more L-like than H-like, and hence that proto-Kwa *L ‘split’ into Yorùbá L and M:

(2)a. Mo ra ba tà. (3)a. Mo lọ.
  1S buy(M) shoe 1S(M) go
  ‘I bought (some) shoes’ ‘I went’
b. Kì ni ó rà?
  what COMP 3S buy(L)
  ‘What did s/he buy’
b. Mò ní lọ.
  IS(L) PROG go
  ‘I’m going’ (cf. Báington 1965: 18)

Stahlke cites two other distributional restrictions to make the same point. These are allophony of object clitics between M (after a verb root that bears H) and H (otherwise), cf. (4); and a systematic absence of segmental noun prefixes bearing H, cf. (5). These latter two arguments sound a bit tendentious, however. True enough, the elsewhere case in (4) groups M and L-bearing verb roots together, but only in a negative way which doesn’t require them to share any property. And Yorùbá has nothing against an H-bearing prefix per se: the syntactic nominal prefixes (infinitive, gerund), like other default functional heads of the language, bear H tone, cf. (6) from Awoyalé (1983), thus the gap in (5a) holds only for open-class (lexical) items.

(4)a. Ò ré mi. (5)a. *[N V-CV… ]
  3S see 1S(M)
  ‘S/he saw me’
  h
b. Ò pa mí.
  3S affect 1S(H)
  ‘S/he hit/killed me’
  ‘calabash’
c. Ò wò mí.
  3S regard 1S(H)
  ‘S/he looked at me’
  ‘locust bean, eggplant’
(6)a. Iṣù dára [INF á-jè ].
  yam.AGR good H-eat
  ‘Yam is good to eat’
b. Jímò ó gò [INF ó-só órò].
  AGR stupid H-say talk
  ‘Jímò talks unreasonably’
c. Jímò ó pé ní [GER lí-lọ-ílè ].
  AGR late in H-go-home
  ‘Jímò is late in going home’

External evidence also fails to support the derivation of Yorùbá M from *L. A search for Igbo cognates of Yorùbá MM nouns retrieves six HH and one H!, but none with an L on either prefix or root, cf. (7). For M verb roots, the result is scarcely different: three H (notably all from the ‘strong H’ class) and one L, cf. (8).5

<table>
<thead>
<tr>
<th>Yorùbá nouns</th>
<th>Igbo nouns</th>
<th>Yorùbá verb roots</th>
<th>Igbo verb roots</th>
</tr>
</thead>
<tbody>
<tr>
<td>akin ‘manly one’</td>
<td>òkhe ‘male, large’</td>
<td>(8)a.  jè ‘eat’</td>
<td>rì ‘eat’</td>
</tr>
<tr>
<td>erin ‘elephant’</td>
<td>ẹnyi ‘elephant’</td>
<td>mu ‘drink’</td>
<td>nì ‘drink’</td>
</tr>
<tr>
<td>èni ‘person’</td>
<td>ọnye ‘person’</td>
<td>yi ‘be tough’</td>
<td>shì ‘be tough’</td>
</tr>
<tr>
<td>ènu ‘mouth’</td>
<td>ọnu ‘mouth’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>èran ‘animal’</td>
<td>ọnu ‘animal’</td>
<td>b. bè ‘cut’</td>
<td>bè ‘cut’</td>
</tr>
<tr>
<td>ibi ‘place’</td>
<td>ẹbe ‘place’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>oni ‘water’</td>
<td>miʃin</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As Akinlabí (1985a,b) notes, all Stahlke’s examples are reinterpretable as evidence that M is not a tone in underlying representation, but rather a default pronunciation of the absence of a tonal element drawn from the set {H, L}.6 Specifically, on the default-M view, the effects in (2–4) are more convincingly rule-derived as L-deletion, L-epenthesis and H-deletion respectively. However, Stahlke’s idea was that a single property underlies all four of the above phenomena, so if we reject his L/M architoneme we need to come up with

---

4This gap also holds outside Yorùbá, e.g. for neighboring Gbè (Stahlke 1971, Capo 1991, Gbeto 1993).
5Transcriptions are orthographic: no tonemark denotes M in Yorùbá, but in Igbo it denotes “same as preceding marked tone”. Supporting evidence for the paired cognates in (7) and (8) is given in Manfredi (to appear); in particular, the consonantal sound correspondences belong to quasi-regular sets. The Igbo distinction between strong and weak H, first observed by Swift et al. (1962), systematically cross-classifies the non-L verb roots in southern districts such as Owere and Mbas. It is wholly absent in northern communities like Ònicha and has a vestigial role in the copula system in Òneèwi, cf. Nwachukwu (1976), Émenanjo (1981), Manfredi (1984, 1991, 1993), Clark (1989) and the Appendix of Déchaine (1993: 497–520).
6Kaye (1981) analyzes M as absence of tone in the ‘four-tone’ systems of Kru (at the western extremity of Kwa).
an alternative or else ascribe the clustering all four alternations in one language to coincidence. Accepting Kaye and Akinlabi’s default–M analysis as a necessary step toward an alternative, representational account, it clearly does not answer the typological question in a holistic way. Admitting that additional hypotheses are required, in order to escape the stipulative character of rule-based theory, we can explore the consequences of the following: (i) a metrical representation of tone domains, and (ii) a fonosyntactic grammatical architecture which constrains metrical constituency.

Going back to the typological issues, consider now (1b): how to allow downstep in a language without requiring downdrift, and still maintain Stewart’s (1965) claim of an implicational relationship between the two phenomena. This problem was already stated by Schachter (1961): in a language with downstep but not downdrift, the downdrift rule would have to be derivationally global, applying just to strings containing an abstract L, i.e. just in contexts where L-deletion will subsequently apply. Thus for Schachter (1965), the theoretically defensible option is to reject Stewart’s inherent relationship between downdrift and downstep. But this may be unduly pessimistic, if current theory offers more possibilities, and in any case the empirical question implied by Schachter’s qualms still needs an answer: how far does Yorùbá have downstep and yet lack downdrift? On this point, much relevant instrumental evidence has been collected in the past decade.

Courtenay effectively dismissed the problem when she called Yorùbá an “orthodox… ‘terraced-level’ language” (1971: 254), but lab phoneticians haven’t gone along. Connell and Ladd (1990) replicate Akinlabi and Lánirán’s (1988) finding that the lowering of H values across a sentence is not systematic lowering of H after L, of the sort familiar from Igbo etc. but (at least partly) the cumulative effect of events of local H-raising before L. This suggests a new way to frame the Igbo/Yorùbá contrast, which is at the heart of the typological problem at hand. At least one assumption of the older literature survives instrumental scrutiny: in both languages, the declination countour of a sentence is computed with respect to tonal context⁷, but then the differences begin. In Igbo (and Èdó), the relevant tonal context is the number of HL feet; but in Yorùbá, only HL sequences increase the differential (Lánirán 1992: 240, 270). Instrumental studies find that a sentence-initial L in Yorùbá is neither raised nor does it contribute to declination (Lánirán 1992: 219): this would follow if L does not participate directly in the creation of higher-level prosodic constituents. In Igbo (and Èdó) by contrast, an initial L is phonetically cliticized to the following H, and hence is sharply raised, to about the pitch of initial M before H in Yorùbá (Ìhiónù 1988); this would follow if L is a foot-final element, Igbo (and Èdó) so that a domain-initial L must attach in a special way. A third difference between the two types is that in Igbo (and Èdó), L as well as H declines across the downdrift domain (= the sentence), but this is not true in Yorùbá, where L maintains a more nearly level value even if the H value dramatically declines. An abstract way to integrate all these observations into one parameter is to say that Igbo/Èdó has a nonlocal downdrift domain, but Yorùbá doesn’t. To make this parameter explicit and concrete, a first approximation is Lánirán’s idea that tonal feet are LH in Yorùbá — versus the HL feet standardly assumed for Igbo-type systems in the Huang-Clements-Inkelas tradition of ‘register tone’ representations.

Similar considerations bear on question (1a). Yorùbá M neither triggers nor undergoes downdrift, and yet the parametric presence of M in Yorùbá correlates with the restriction of downdrift to H, and to local contexts, as just mentioned. A further point is that Yorùbá M does not get downdrifted systematically, although it does get downstepped, i.e. its pitch is affected by covert L but not (necessarily) by overt L.⁸ If we want to relate these observations to the preceding ones, it is logically impossible to do so in terms of rules, since there is no element in common: one rule will talk about what happens with H and L, the other (non)rule about what doesn’t happen with M. Evidently, a necessary step in linking the two effects together is to posit some representation in which H, M and L all participate in a predictable manner.

---


⁸The context-dependency of the downdrift of Yorùbá M was observed already by Welmers (1971: 107, cited by Elugbe 1995: 69). He notices that the negative aux kọ in (ii) has no effect on the (derived) M of ra ‘buy’, but the L-bearing prefix of ảga chair does lower the M of the root syllable, so that sentences (i) and (ii) end on the same phonetic value for M. (As a control, it would also be nice to know that sentence (iii) has the same final pitch, but that goes beyond the point at hand.)

- **í.** Ọ ra ǎga. 3S buy chair
- **ii.** pro Kọ ra ǎga. 3S NEG buy chair
- **iii.** Mo ra ǎga. 1S buy chair

‘S/he bought [a] chair’ ‘S/he didn’t buy [a] chair’ ‘I bought [a] chair’

Welmers made this remarkable observation without the benefit of any acoustic technology besides unusually big ears.
This brings us to (1c) and Ìkòm-Yàlà: a language with M but with nonlocal downdrift and therefore with downdrift of L and M (as well as H) across the sentence (Armstrong 1968, 1972). Ìkòm-Yàlà is precisely what Stewart and Courtenay suspected Yorùbá to be. The undeniable difference between Ìkòm-Yàlà and Yorùbá shows that while M may be an inert tone with respect to prosodic structure, such is not necessarily the case. In effect, UG has more than one way to integrate M into a prosodic structure which includes just H.

Abraham (1958: 438) describes a loss of H affecting just the item (9)

The difference with respect to Stahlke’s case is that the syntactic context in (3) is before a H-bearing aux, but the lexical context in (9) is before an L-bearing root-initial syllable.

Akinlabí (1985) and ‘S. Oyelárán (p.c.) draw attention to the suppression of H before L in the domain of a prefix that also bears L. The context is either lexical (10a) or derivational/inflectional (10b).

Abraham (1958: 438) describes a loss of H affecting just the item ní ‘have’ in agentive synthetic compounds whose noun contains only L tone, cf. (11); and he records another suppression of the H of ní in (12b).11

(10a) ọpọyà ‘fear’ ← ọ-pá-àyà ‘NOM-cringle-chest’
ike.ìrin ‘fourth’ ← ì-kó-èrin ‘NOM-take-four’ (where [] = downstep)
ìbì ‘twins’ ← ì-bí-èjì ‘NOM-give.birth-two’
òfàlè ‘lover (m.)’ ← ò-(fì-)è-àlè ‘agent-NOM-love-girlfriend’, cf. a-fà-è lè MHL

b. ọ̀nàkòrò ‘any/bad talk’ ← ọ̀nà-kì-ŋò ‘word-C0-word’
òkó ọgbòdò ‘needn’t’ ← ọkó ‘NEG’ ọgbè-ọdò ‘lite.V-immediate.presence’, cf. ọgbè, ọdò ‘3s must’

(11) alàgbà ‘respected person’ ← o-ní-àgbà ‘ER-have-seniority’
ìlàyé ‘living creature’ ← o-ní-àjì ‘ER-have-space’
ołóṣí ‘poor person’ ← o-ní-òjì ‘ER-have-destitution’

(12a) Ta ni ò ni [kp ò [dp ò [np åjà]]]? b. Ta ni ò ni [kp Ø [dp [d àjá] [np tì]]]?
‘Who has/owns a dog?’ who COMP 3S V dog
[talòlajá] = M H H H
‘Who has/owns the dog in question?’ who COMP 3S V dog

(6) above showed some null functional heads—possibly D0—spelling out as H (cf. also Manfredi 1992A). H-spellout also affects T0, C0 and K0, cf. (13) and Déchaine (1992, 1993). In relative and resumptive contexts as in (13b), default H is supplemented by segmental material (tì, ò) recovering chain (agreement) features.

(13a) Ògbè ì lọ. 
farmer T go ‘A/the farmer went’

b. Ògbè tí ò lọ
c. Ògbè ò lọ. 
farmer C.AGR T.AGR go is cat K
‘the farmer who went’ ‘I ate it’
The irrealis modals \textit{kọ} ‘doesn’t/won’t’, \textit{yóó} ‘will’, \textit{máá} ‘Don’t!’ bear a final \( L \) which might be motivated compositionally as the left prosodic boundary of a head-movement domain, cf. Déchaine (this volume).

The spreading of \( H \) onto \( L \) and conversely (Ward 1952) fails to cross either \( M \) (Láníran 1992: 1999, Akinlabí and Liberman 1995) or covert \( L \) (Bámígbóṣé 1966).

Given the learnability problem, one may be excused for hoping that all the above tonal phenomena converge as evidence of a simple representational parameter. Here goes.

3. **Metrical analysis**

Long before Bamba’s (1992) metrical framework of tone analysis, several scholars of Kwa languages had proposed that ‘latent’ or zero tones underly downstep (Stewart 1965; Bámígbóṣé 1966, 1970). In Benue-Congo, Williamson (1968, now also 1988) reported the phrasal determination of pitch accent in Izón. Soon thereafter, analyses appeared of Kwa languages where tone structure seems to condition segmental deletion (Dakubu 1975; Kaye 1981; Chumbow 1982), though this causality was impossible in terms of the reigning, autosegmental theory. Considerations of typology suggested that tones are licensed in metrical positions (Kimenyí 1978, 1987; Clements and Ford 1979; Huang 1980; Odden 1984; Ladd 1991) and that metrical structure motivates abstract or zero tones (Koopman 1982; Manfredi 1992B, 1993).

I assume with Bamba that ‘two-tones plus downstep’ systems like Igbo reflect the constituent in (14):\(^{13}\)

\begin{equation}
(14)
\end{equation}

Metrical structure, binary by design, can accommodate a third tone (\( M \)) only if \( s \) or \( w \) branches and \( s \) or \( w \) is the attachment site. If the two if’s coincide on one node, then the direction of attachment matters, and if they don’t, a reattachment is unavoidable. Of the six possible outcomes, only two are well-formed.

\begin{equation}
(15)
\end{equation}

\begin{equation}
(16a)
\end{equation}

\begin{equation}
*(16b)
\end{equation}

\begin{equation}
*(17a)
\end{equation}

\begin{equation}
*(17b)
\end{equation}

\begin{equation}
*(18)
\end{equation}

In (15), \( M \) attaches to \( w \); \( s \) branches, not \( w \), so \( L \) must reattach to \( s \); right reattachment would violate locality (19a) as the governor \( H \) would be nonadjacent to \( w \), so \( L \) goes left, becoming strong (ungoverned). In (16b), \( M \) on the right side of branching \( w \) is ungoverned, again by (19a), but left attachment (16a) is fine, assuming that \( M \) as an un(der)specified element does not invoke (19a), i.e. \( M \) can ‘transmit’ government from \( H \), analogous to a governed preposition in a verb phrase.\(^{14}\) (17) shows the two subcases where \textit{attach} and \textit{branch} coincide on \( s \); both are out by (19b). (18) ignores all three of the licensing conditions in (19).

\begin{equation}
(19)a.
\end{equation}

\begin{equation}
(19)b.
\end{equation}

\begin{equation}
(19)c.
\end{equation}

I claim that the two well-formed options are attested: (15) by Yorùbá, and (16a) by Òkóm-Yàlà.

\footnote{Alone of these, Clements and Ford did not explore metrical tone representations, which however they did recognize—at least rhetorically (1979:198)—as a means by which cross-Niger-Congo generalizations could be expressed.}

\footnote{See also Bamba (this volume). I preserve Bamba’s notation whereby the non-head of a metrical constituent—the \( w \) branch—is parenthesized; and for all constituents, tonal or metrical, the head projects to a higher domain by a vertical association line.}

\footnote{This replaces my (1979) stipulation of two flavors of \( M \): \([-H, -L]\) as in Yorùbá and \([+H, +L]\) as in Òkóm-Yàlà.}

3.1 Yorùbá

Inspection of (15) shows that it is the only three-tone metrical constituent consistent with Lánírán’s abovementioned LH foot for Yorùbá. A second useful feature is its consistency with the strength hierarchy \( H > L > M \) posited by Bámígbọ̀se (1965) and others to describe the tonal consequences of vowel contraction.\(^{15}\)

Tone spreading is hypothesized to be an effect of government (Manfredi 1993). Yorùbá \( H \) and \( L \) spread onto each other. If \( H \) is the metrical head, the spread of \( L \) onto \( H \) forces the learner to let \( s \) branch, as in (15). Spread of \( H \) onto \( L \) is less expected; the syntactic analog is perhaps exceptional case marking, cf. (20). The blocking of spread by intervening \( M \) translates easily into the blocking of government, cf. (21).

\[(20)\]
\[
\begin{array}{c}
\text{s (w)} \\
\text{s (w)}
\end{array}
\]

\[L \quad H \quad \ldots\]

\[\text{x} \quad \text{x} \quad \text{x}\]

In this foot structure, domain-initial \( M \) is anomalous. If it adjoins to the subsequent constituent, and that constituent has a zero head, i.e. does not contain \( H \), this might pose a problem of locality which is resolved by forming the adjunction one level down, to \( L \). This is a plausible story for the ML…~LL… cases in (9).

Yorùbá downstep as in \( alá.gbo \) ‘owner/seller of herb infusions (\( āgbọ̀ \))’ also has a straightforward source in the representation in (15). I assume prosodic structure is preserved across elision, which affects the second tone-bearing unit (rime) in (22a). The downstep trigger is the null head of the second foot in (22b).

\[(22)a.\]
\[
\begin{array}{c}
\text{s (w)} \\
\text{s (w)}
\end{array}
\]

\[M \quad H \quad L \quad \emptyset \quad M\]

\[o \quad n \quad i \quad a \quad āgbọ̀\]

\[b.\]
\[
\begin{array}{c}
\text{s (w)} \\
\text{s (w)}
\end{array}
\]

\[M \quad H \quad \emptyset \quad M\]

\[o \quad n \quad a \quad āgbọ̀\]

More challenging is why covert \( L \) should block \( H \) spread. Bámígbọ̀se (1972: 26) cites a minimal pair \( ọlọ̀nọ̀ \) ‘wealthy person, owner of wealth (\( ọ̀rọ̀ \))’ with spread, \( ọ̀rọ̀.rọ̀ \) ‘person in question, owner of discussion (\( ọ̀rọ̀ \))’ with no spread. By the above assumptions, the respective underlying forms are (23a-b). The problem is why elision and reassociation don’t merge the two cases, obliterating the reported phonetic difference.

\[(23)a.\]
\[
\begin{array}{c}
\text{s (w)} \\
\text{s (w)}
\end{array}
\]

\[M \quad H \quad M \quad L \quad \emptyset\]

\[o \quad n \quad i \quad o \quad rọ̀\]

\[b.\]
\[
\begin{array}{c}
\text{s (w)} \\
\text{s (w)}
\end{array}
\]

\[M \quad H \quad L \quad \emptyset\]

\[o \quad n \quad \emptyset \quad o \quad rọ̀\]

The solution again denies resyllabification and appeals to structure preservation: the first syllable of ‘wealth’ (23a) is in the same foot as the \( H \) of \( ní \) ‘have’, but the same is not true for ‘matter’ (23b). If \( ϕ \) is the foot boundary, it suffices to say that \( H \) does not govern \( L \) in (24b) because it’s not fully external to \( ϕ \).\(^{16}\)

\[(24)a.\]
\[
\begin{array}{c}
\text{s (w)} \\
\text{s (w)}
\end{array}
\]

\[M \quad H \quad L \quad \emptyset\]

\[o \quad l \quad \emptyset \quad rọ̀\]

\[b.\]
\[
\begin{array}{c}
\text{s (w)} \\
\text{s (w)}
\end{array}
\]

\[M \quad H \quad L \quad \emptyset\]

\[o \quad l \quad \emptyset \quad rọ̀\]

\(^{15}\)Chumbow (1982) applies the same strength hierarchy to Ògòrí (probably a macro-Nupe language, ‘Y. Awóyalé, p.c.). The difference of (15) vs. (16a) predicts that non-\( H \) tones will resolve differently under vowel contraction in Yorùbá vs. Íkọ́m-Yálà.

\(^{16}\)A comparable definition of ‘external’ with respect to governing category is given by Chomsky (1986: 9).
Needless to say, the category \( \phi \) in (24) is not a prosodic diacritic, just a visual cue for the constituency relations which are carried over from underlying structures in (23); the same relations could be read off a fuller notation which included syllabic (onset-rime) structure, assuming that this links to prosody. One must acknowledge, however, that \( \phi \) does threaten the prosodic linking parameter: the hypothesis that prosodic structure links to either tones or rimes but implicitly not to both at the same time (Manfredi 1993: 177).

\( \phi \) also says something—though not everything—about the suppression of H exemplified in (10). This differs minimally from the glide-blocking context in (24), in that the domain is L-initial. Here the suggestion is that, if H reassociates to the third TBU due to the elision of the second, not only is its structural position within the second foot incomplete for the reason cited in (24b), but it is equally no longer completely within the first foot, and hence the initial L is unlicensed, cf. (25b).

One can imagine two ways to ‘fix’ (25b): suppress the H, or delete the initial L whose licensing is put at risk by H’s reassociation to the third TBU. Why the former occurs and not the latter, is suggested by the forms like *ike rin ‘fourth’ in which H suppression cooccurs with downstep, cf. (26).

The null head in (26b), whose presence is witnessed by the phonetic downstep in this example, is required by projection (19b): M can’t project a foot. We can speculate that reassociation brings the null head into conflict with a filled head of the first foot, so structure preservation requires suppression of the H.

The list of suppressed Hs, as recounted in §2, includes cases restricted to the morpheme ni, cf. data (11) and (12) above. A form like alagbá in (11) shows a different outcome from that of *òbò in (23b), which is surprising because they have the same tonal input. The possibility of *òbò itself does not undermine the suggestion of Manfredi (1995) that the sequence MH\(L\) is unable to support two feet, since we have argued that the two feet of *òbò are also supported by additional structure: the category \( \phi \). If the *òbò-type is productive but the cases like alagbá in (11) constitute a finite list, one can deny that elision takes place in the latter, perhaps because ni ‘have’ is no longer synchronically analyzable, so there is no basis to postulate \( \phi \). It is also relevant that the H tone of ni ‘have’ is uniquely unstable in Yorùbá morphosyntax, as argued in the same paper from the other instance of H suppression, cf. (12).

Another issue where prosodic licensing gives some insight is the complementarity/asymmetry of lexical prefixes, which bear M or L, cf. (5), vs. syntactic prefixes, which bear H, cf. (6) and (13). If it is correct that the open-class lexicon does not have access to the closed-class or ‘functional’ categories (C, T, K, D), then the metrical framework offers a simple way to compel the latter to spell out as H, namely require them to be prosodically strong. A converse requirement—that lexical prefixes are weak—is inconsistent with (15), since L is licensed only in s; however, the desired effect is obtained by restricting the spellout of null heads to phrasal syntax, i.e. by limiting H-spellout to functional categories. The remaining tone effects in §2 are H deletion on an object clitic after H as in (4a), and L-deletion before a direct object as in (2a). The former is plausibly related to the H-spellout as just described, i.e. H-spellout is blocked in the immediate domain of H; intuitively, two H’s are unable to coexist in one foot, but syntax insists that one foot is all there can be, cf. (27a). L-deletion looks like the converse problem: too much structure and not enough tone, cf. (27b): one could argue that H-spellout is not possible there because null K/D is properly governed by a lexical head.
Summarizing, the discussion has moved from relatively phonetic issues (spreading and downstep) to syntax (closed-class affixation, clitics, Case). If specific analyses of individual problems along the way are revised or rejected, my point here has been different: that a list of such problems can be brought face-to-face through the lens of prosody, specifically through the claim that Yorùbá belongs to a specific metrical type, as represented in (15).

3.2 Ìkòm–Yàlà versus Yorùbá

Ìkòm–Yàlà, I claim, is an example of the other way in which a third lexical tone can be integrated into binary metrical structure. The prediction is that Ìkòm–Yàlà and Yorùbá do not differ randomly in phonetic tone effects, but that the properties of Ìkòm–Yàlà converge on a different foot structure, the one in (16a).

Armstrong observed that Ìkòm–Yàlà L does not spread onto H, although Ìkòm–Yàlà M has the property of blocking H-spread onto L (1972: 425). Unlike Yorùbá, Ìkòm–Yàlà has nonlocal downdrift. Summarizing these parametrically, one can say that M is strong and L is weak in Ìkòm–Yàlà, while Yorùbá reverses the two values. All these properties follow by inspection of (16a).

Manfredi (1979) proposed to represent Ìkòm–Yàlà in terms of hierarchical tone features, as in (28a); this translates into the metrical-tone framework as in (28b)

Note, finally, that the constraint set in (19) remains constant across all the languages surveyed.

4. Conclusion: from prosody to fonosyntax

§3.1 proposed explicit answers to the longstanding West-Africanist questions in (1), and provided a coherent account of the list of observations in §2. Whether or not these answers and this account are correct, they are not value-neutral: that is, they contradict a theory of syntax-phonology ‘interface’ in which anything and everything is possible. Rather, they require us to believe that tones are integrated into prosodic structure, and phrase structure, with surprisingly few options. In other words, they insist that Yorùbá phonology is learnable, but only if UG has a constrained architecture and phonology is part of UG.
References


_____ (1992A) A typology of Yorùbá nominalizations. MITWPL 17, 205-17.


_____ (1995) Syntactic (de)composition of Yorùbá ‘be’ and ‘have’. In L. Nash & G. Tsoulas (eds), Langues et Grammaire-1. Département des Sciences du Langage, Université de Paris-8, Vincennes à Saint-Denis.


African Studies Center
Boston University
270 Bay State RD
Boston MA 02215