African languages and phonological theory

Larry M. Hyman
University of California, Berkeley

1. Introduction

I’ve been asked to write about the mutual influence of African languages and phonological theory, specifically addressing two questions: What and how have African language contributed to phonological theory? What and how have linguistic theories contributed to the understanding African phonology? To treat these questions properly would be a major undertaking, first, because of the large number of African languages, and, second, because of their considerable diversity. To cite Greenberg’s (1963) influential classification, the roughly 2000 languages of Africa fall into four major linguistic phyla: Niger-Congo, Afro-Asiatic, Nilo-Saharan, and Khoisan (see Heine & Nurse 2000 for a more recent overview of African languages and their classification). Except for the click consonants of the last family (which spill over into some neighboring Bantu languages that have “borrowed” them), the phonological phenomena found in African languages are duplicated elsewhere on the globe, though not always in as concentrated a fashion. The vast majority of African languages are tonal, and perhaps most also have vowel harmony (especially the type known as “ATR harmony”). Not surprisingly, then, African languages have figured disproportionately in theoretical treatments of these two phenomena. On the other hand, if there is a phonological property where African languages are underrepresented, it would have to be stress systems—which rarely, if ever, achieve the complexity found in other (mostly non-tonal) languages. However, it should be noted that African languages have contributed significantly to virtually every other aspect of phonological theory.

There are at least two ways one might present the influences of African languages and phonological theory on each other:

First, one could do as I have just done and enumerate the African phenomena which have figured in theoretical work. I would list the following as most prominent: 1) tone; 2) vowel harmony; 3) nasals and nasalization; 4) labial and palatal prosodies; 5) slots and moras; 6) reduplication; 7) syntax-phonology interface.

Second, one could identify issues in the development of phonological theory and specific frameworks where African languages have played a key role. The following come immediately to mind, in approximate chronological order: 1) distinctive feature theory; 2) autosegmental phonology; 3) CV- and moraic phonology; 4) government phonology; 5) lexical phonology and morphology; 6) underspecification theory; 7) feature geometry; 8) prosodic morphology; 9) prosodic domain theory; 10) optimality theory.

Either way, the generalization is that the African continent, which houses approximately 1/3 of the world’s languages, has been central to the study of prosodic phenomena and to the development of formal frameworks to account for them. This is most obvious in the case of tone. Drawing disproportionately from African tone systems, Leben (1973), Goldsmith (1976), Williams (1976) and others showed up the inadequacies of classical (segmental) generative
phonology, as encoded in Chomsky & Halle (1968), henceforth SPE (which, interestingly, cited only one work on tone, Wang (1967), which presents proposals for tone features based on Chinese). The resulting autosegmental revolution then spread from tone to other aspects of phonology and morphology—and, ultimately to syntax and semantics (see Saddock 1991, Yip et al (1987), etc.).

My discussion will be organized as follows. In §2 I discuss two areas where African languages have uniquely contributed to phonological theory: tone and vowel harmony. In §3 I discuss in more rapid succession the remaining five areas mentioned above where African languages have made less unique, but still important contributions. After this, in §4 I conclude by considering the reverse relation, the contribution of phonological theory to African languages.

2. Areas where African languages have uniquely contributed to phonological theory

As mentioned above, African languages have contributed to phonological theory in a number of ways. Here I propose to show this by taking up two of the six phenomena mentioned in §1. I will begin with tone in §2.1, since this is the area that best allows for a focused discussion of both of the questions that I have been asked to address. I shall then turn to vowel harmony in §2.2.

2.1. Tone

As mentioned, tone was central to the autosegmental revolution in generative phonology. Let us recall the basis of this “revolution”. The dominant view within structuralist and early generative phonology was that phonological strings could be subdivided into a succession of discrete segments. Each segment, in turn, consisted of a matrix of simultaneous “distinctive features”, generally claimed to be binary, in the Jakobson-Halle tradition. These features had both a classificatory and phonetic function, being designed to capture the phonological oppositions found in languages as well as their output realizations. While not yet receiving very much attention, the assumption in the 1960s was that tone could be characterized with additional features on vowels, as in (1).

(1) Segmental representation of H and HL falling tone

\[
a. \ [á] = \begin{bmatrix} +\text{syll} \\ -\text{cons} \\ -\text{high} \\ +\text{low} \\ +\text{back} \\ -\text{round} \\ +\text{HIGH} \end{bmatrix} \\
b. \ [â] = \begin{bmatrix} +\text{syll} \\ -\text{cons} \\ -\text{high} \\ +\text{low} \\ +\text{back} \\ -\text{round} \\ +\text{FALLING} \end{bmatrix}
\]

In feature systems such as Wang (1967), based largely on Chinese dialects, high (H) tone could be indicated as \(+\text{HIGH}\), as in (1a), while a high-to-low (HL) falling tone would be \(+\text{FALLING}\), as in (1b). Pike (1948) had split tone systems into what we can refer to as a Chinese- vs. African-type: Whereas Chinese dialects have an abundance of contour tones, contours seem quite secondary in African tone languages, where the primary opposition is between level tones. Many African languages have only two level tones, high and low.
Tonal representations such as the above were shown to run into a number of problems. In many African languages, a falling tone such as in (1b) shows “edge effects”: It appears to be a H tone from the point of view of what precedes it, but a L tone from the point of view of what follows. Thus, if a L is raised to a mid (M) tone before a H tone, we expect also that it will be raised before a HL falling tone. The feature specification [+FALLING] does not capture this fact, and any attempt to represent the fall as a sequencing of [+HIGH][-HIGH] squeezed into a single matrix below the segmental features would be incoherent in a formal framework which otherwise views segments (here, vowels) as a single vertical array of distinctive features.

In establishing autosegmental phonology, Goldsmith’s (1976) proposal was that an /a/ with high- or falling-tone should be represented roughly as in (2).

(2) Autosegmental representation of H and HL falling tone

\[
\begin{align*}
\text{a. } [\acute{a}] & = \begin{bmatrix}
+\text{syl} \\
-\text{con} \\
-\text{high} \\
+\text{low} \\
+\text{back} \\
-\text{round}
\end{bmatrix} \\
\text{b. } [\grave{a}] & = \begin{bmatrix}
+\text{syl} \\
-\text{con} \\
-\text{high} \\
+\text{low} \\
+\text{back} \\
-\text{round}
\end{bmatrix}
\end{align*}
\]

As seen, Goldsmith proposed a distinction between a segmental tier vs. a tonal tier, which are semi-autonomous in the sense that they are separate, but linked by association lines. Among the familiar arguments for a two-tier representation are the three listed in (3).

(3) Three arguments for two tiers (segmental vs. tonal):

a. **non-isomorphism:** F’s of one tier do not line up/synchronize with F’s of the other tier, i.e. overlapping of segmental vs. tonal features

b. **stability:** F’s of one tier may be deleted without affecting (deleting) F’s of the other tier

c. **zero representation:** F’s may be specified on one tier but partially/totally lacking on the other tier

By non-isomorphism is meant that associations of tones to tone-bearing units (TBUs) are often not one-to-one. Two tones may link to a single TBU, as in (2b). In addition, a single tone may link to two TBUs. As a result, an underlying contrast arises in (4a) in the Bantu language Kukuya (Paulian 1975, Hyman 1987):

(4) a. Medial b. Prepausal

\[
\begin{align*}
\text{a. } \text{mé-bá} & \quad \text{wátá} \\
\text{b. } \text{mé-th} & \quad \text{wátá}
\end{align*}
\]

Both má-bá ‘they are oil palms’ and wátá ‘bell’ are pronounced H-H in medial position. Before pause, however, there is a rule that lowers a H to M. As seen in (4b), the H → M rule affects the
last H feature, not just the last TBU. The contrasting representations in (4a), which had no equivalent in pre-autosegmental tonology, provide the structural difference that results in the surface opposition of H-M vs. M-M before pause. These and other facts from Kukuya showed that at least in some cases we must be able to talk about tones in terms of abstract melodies, rather than concrete features on syllables, moras, or vowels. Paulian (1975) thus recognized five “schèmes tonals” (tonal melodies) in Kukuya which can be predictably mapped onto stems of different lengths: L, H, LH, HL, LHL. Leben (1973) had proposed exactly the same for Mende, although not without complications and challenges (Leben 1978, Dwyer 1978, Conteh et al 1983). The formal recognition of such tonal melodies inspired analogous non-linear analyses of Semitic templatic morphology (McCarthy 1981) and partial reduplication (Marantz’ 1982).

Africa also gifted phonology with the floating tone. While there had been early descriptions of tonal morphemes (e.g. the H associative tone of Igbo), other studies showed that floating tones could also be lexical. A particularly persuasive case comes from Aghem (Grassfields Bantu). Although the two nouns kí-fú ‘rat’ and kí-wó ‘hand’ are both pronounced H-H in isolation, the phrases in (6) show that they exhibit quite different tonal behaviors in context:

(6) Floating lexical tone, e.g. Aghem /-fú/ ‘rat’ vs. /-wó`/ ‘hand’
   a. kí-fú kí-m`ì ‘one rat’ fú kí-n ‘this rat’
      H H L L H H
   b. kí-wó kí-m`ì ‘one hand’ wó `kí-n ‘this hand’
      H H L L L H L H

In the forms on the left, each noun is followed by the numeral kí-m`ì ‘one’, which is pronounced L-L in isolation. As seen in (6a), the H of kí-fú spreads onto the prefix of the numeral to derive a H-L realization. H tone spreading does not occur after kí-wó in (6b). The reason is that the root /-wó`/ carries a lexical floating L tone.

The above and other arguments thereby justify the basic premise of autosegmental tonology stated in (7).

(7) Tones (T’s) must be represented as semi-autonomous from the tone-bearing units (TBUs) on which they are realized

In African language courses I was taught that words that differ only in tone, e.g. Nupe bá ‘be sour’, ba ‘cut’, and bà ‘count’ (H, M and L tone, respectively), are as distinct from each other as words that have different vowels, e.g. bí ‘crumble’, bé ‘come’, bá ‘be sour’. However, our transcription practices suggest otherwise: Whereas we have IPA symbols to represent different front-back features on low vowels, [æ], [a], [ʌ], we do not have separate IPA vowel symbols for H, M and L tone [a]. While I often exchange email with colleagues in which I use upper case A E I O U for H tone vowels and lower case a e i o u for L tone, no orthography to my knowledge has ever proposed this or anything comparable. Whether based on Eurocentrism or on general linguistic intuition, where tone has been marked, all students of African languages have done so by adding accents or numbers, rather than creating new vowel symbols.
As indicated, the traditional (pre-generative) view that tone is “semi-autonomous” from other vowel features was brilliantly captured by the autosegmental framework. It would be interesting to speculate on the form the subsequent autosegmental revolution might have taken without the impetus of African tone. Would present-day phonologists be talking as readily about autosegmentalized H(igh) and L(ow) features for Chinese tonal contours if it were not for the input from Hausa, Igbo and Mende? Would Pierrehumbert (1980) and subsequent scholars have developed an analogous approach to intonational systems such as in English? And what would our view be of other phonological phenomena to be discussed below, which also have autosegmental properties?

2.2. Vowel harmony

After tone, African languages are well known for providing vast numbers of phonological systems with vowel harmony (VH), particularly of the advanced tongue root (ATR) variety. Stewart (1967), cited in SPE, and Schachter and Fromkin (1968) educated early generations of generative phonologists as to the intricacies of Akan (Kwa) ATR harmony. However, it was Clements (1977a, 1981) who applied the new autosegmental framework both to Akan and to VH in general. Although the existence of transparent neutral vowels had been known from Finnish and Hungarian, Clements provided an autosegmental account of opaque neutral vowels, based on Akan. Since this language has both prefixal and suffixal harmony, as seen in (8),

\[
\text{(8) ATR harmony in Akan} \\
\text{O + fitI + I \ \ \ [o-fiti-i] ‘he punctured (it)’} \\
\text{ [+ATR]}
\]

he also was able to establish the general property of “root-control”. For Clements this meant that the directionality of assimilation in VH need not be stipulated, but rather followed from convention: the root features [+ATR] and [-ATR] spread left and/or right, as needed, so that no vowel would lack a specification and, hence, be ill-formed. For some theorists, root-control is a necessary, definitional property of VH systems. Thus, vowel alternations which are referred to as metaphony, umlaut, or stem ablaut, and which are often triggered by suffixes, would not be identified as VH.

Since Clements, Niger-Congo and Nilo-Saharan ATR harmony systems have figured prominently in the theoretical study of VH. At the same time, they have contributed to theories of vowel features and feature geometry. Government- or dependency-theories involving the vowel elements \{i\}, \{u\}, \{a\} are developed by Kaye, Lowenstamm & Vergnaud (1985) and Rennison (1986) based on Kpokolo (Kru) and Koromfe (Gur), respectively. More recently, African ATR systems have provided the fuel for Archangeli & Pulleyblank’s (1994) grounded phonology and are also featured in a number of optimality theoretic works, including recent dissertations such as Bakovic (2000) and Krämer (2001).

As in the case of tone, African ATR harmony has not only contributed to linguistic theory, but also to the way VH is described in other languages. Hall & Hall (1980), for example, are
explicit in applying their Africanist insights to Nez Perce, whose unusual harmony, they suggest, should be analyzed in terms of ATR. There are striking resemblances between the VH found in the Pacific Northwest and that found on much of the Asian land mass. It is thus not surprising that ATR/RTR has also been recognized in Tungusic languages (Li 1996, Zhang 1996), and may very well be implicated in languages from Tibetan to Chukchee.

While African languages have provided the model for VH based on ATR, or its retracted counterpart, RTR, they also provide the world’s greatest supply of VH systems based on tongue height (F1). Best known are those found in Bantu, exemplified below from Luganda:

(9) Height harmony in Luganda

<table>
<thead>
<tr>
<th>plain stem</th>
<th>stem + causative</th>
<th>stem + applicative</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. lim-a</td>
<td>lim-is-a</td>
<td>lim-ir-a</td>
</tr>
<tr>
<td>tum-a</td>
<td>túm-is-a</td>
<td>túm-ir-a</td>
</tr>
<tr>
<td>land-a</td>
<td>land-is-a</td>
<td>land-ir-a</td>
</tr>
<tr>
<td>b. sek-a</td>
<td>sek-es-a</td>
<td>sek-er-a</td>
</tr>
<tr>
<td>kól-a</td>
<td>kól-es-a</td>
<td>kól-es-a</td>
</tr>
</tbody>
</table>

‘cultivate’
‘send’
‘climb’
‘laugh’
‘work’

As seen, the causative and applicative suffixes surface with the vowel [i] when preceded by /i, u, a/, but with [e] when preceded by /e/ or /o/. Sometimes called “mid harmony”, variants of height harmony are found in most Bantu languages (Hyman 1999). The relation to ATR has not been missed by Africanists or theoreticians such as Clements (1991), who provides a geometric model of vowel aperture designed to capture both types of VH. Other views of aperture and its relation to other vowel features also draw heavily from African languages, e.g. Hyman (1988), Odden (1991). (See also Pulleyblank 1988 for an important contribution to vowel underspecification, based on Yoruba.) Finally, within optimality theory, Beckman’s (1997) notion of positional faithfulness is based on Shona height harmony, which has the same properties as in Luganda, Swahili, Chichewa, Cibemba, and so forth.

3. Areas where African languages have significantly contributed to phonological theory

In the two areas of tone and VH, it can be argued that African languages have unique properties to offer the field, or at least, that the richness that is provided by the majority of 2,000 languages provides varied and unusual opportunities for theoretically-minded phonologists. In both cases what we have learned from African languages has provided major insight into the nature of phonology. In this section I will briefly survey five areas where African languages have clearly played an important role, but perhaps in a somewhat less “unique” fashion. I have ordered them from least to greatest impact.

3.1. Nasals and nasalization

African languages have contributed to our understanding of nasals and nasalization in several ways. Nasality is of course another feature that often takes on suprasegmental or prosodic character. A large number of West African languages have nasalized vowels, and in many of these nasal consonants are in total or near complementary distribution with voiced oral counterparts.
Thus, quite early in generative phonology, Schachter & Fromkin (1968) had proposed derivations such as following for their dialects of Akan:

(10) No underlying nasal consonants in Akan

a. /bã/ → [mã] ‘give’
b. /dã/ → [nã] ‘and’
c. /yã/ → [ỹã] ‘receive’
d. /wãdi/ → [wãni] ‘scrape’
e. /hũ/ → [hũ] ‘fear’

Rather than representing nasality on vowels, where it is contrastive, they also could quite easily have abstracted the feature away as a prosody, [+NAS], quite in keeping with the Firthian tradition. This was subsequently proposed by Leben (1973) for Terena and Goldsmith (1976) for Guarani, two non-African languages, and by Hyman (1982) for the Lower Cross language, Gokana.

Concerning other issues of nasality, African languages figured prominently in Herbert’s (1986) work on prenasalization and N+C sequencing. Many if not most African languages allow NC segments or clusters of some type, and it is not surprising to see African cases contributing to their analysis: Are they one or two segments? If one, what is their feature geometry? If two, is the nasal moraic or not? Finally, I should mention the theoretical significance of on-going research on long-distance nasal consonant harmony, e.g. Yaka /miituk-idi/ → [miituk-ini] ‘pout + perfective’, found also in Kikongo and other Western Bantu languages (Ao 1991, Hyman 1995, Piggott 1996, Rose & Walker 2003, Walker 2000, Hansson 2001).

3.2. Labial and palatal prosodies

Another feature which received early attention in autosegmental phonology was palatalization—and once again Clements (1977b) was there building on African examples: Chimwiini (Bantu), Fe’fe’-Bamileke (Grassfields Bantu), Ewe (Kwa). The real African contribution, however, occurred a few years later, when McCarthy (1983) introduced Afro-Asiatic prosodies into current phonology. Autosegmentalists had become accustomed to tonal morphemes, but had not yet appreciated the type of labial and palatal morphemes found in Chaha (Ethiopian Semitic):

(11) Perfective 3 m. sg. in Chaha

<table>
<thead>
<tr>
<th>Without object</th>
<th>With 3 m.sg. object</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. dænæg</td>
<td>dænægw</td>
</tr>
<tr>
<td>nædæf</td>
<td>nædæfw</td>
</tr>
<tr>
<td>nækæb</td>
<td>nækæbw</td>
</tr>
</tbody>
</table>

‘hit’
’sting’
‘find’
b. s\textsuperscript{y}æfær  s\textsuperscript{y}æf\textsuperscript{wær} ‘cover’
  nækæs   næk\textsuperscript{wæs} ‘bite’
  kæfaet   kæf\textsuperscript{wæt} ‘open’

c. qætær   q\textsuperscript{wætær} ‘kill’
  mæsær   m\textsuperscript{wæsær} ‘seem’
  mæk\textsuperscript{y}ær   m\textsuperscript{wæk\textsuperscript{y}ær} ‘burn’

As seen, the third person masculine singular object is marked by a labialization feature which links to the right-most labializable consonant—which means a non-coronal. Although this phenomenon is known mostly from Ethiopian Semitic, there are comparable cases found in Chadic languages as well as templatic phenomena reminiscent of Arabic. For further theoretical work on “featural affixes” involving African languages see especially Akinlabi (1996) and Zoll (1998).

3.3. Slots and moras

In the early 1980s, African languages provided important contributions to the development both of skeletal (CV) and moraic phonology. A good case in point is Luganda compensatory lengthening, by which sequences such as /Cia/ and /Cua/ are realized [Cya:] and [Cwa:], respectfully. Clements’ (1986) insight was that the high vowel reassociates to the preceding C slot, delinking from its V slot, which is in turn filled by spreading of the following vowel, as in (12).

\[(12) \text{CV account of Luganda compensatory lengthening} \]
\[\begin{array}{ccc}
C & V & V \\
\rightarrow & /\backslash & | \\
k & i & a \\
\end{array} \rightarrow \begin{array}{ccc}
C & V & V \\
\rightarrow & /\backslash & \backslash \backslash \\
k & i & a \\
\end{array}\]

In the moraic account (Hyman 1985), the /a/ spreads right to left onto the first mora:

\[(13) \text{Moraic account of Luganda compensatory lengthening} \]
\[\begin{array}{c}
\mu \\
k \\
| \\
i \\
a \\
\end{array}\]

Clements also considers the lengthening of a vowel before a NC sequence, e.g. /genda/ → [ge\textsuperscript{nda}] ‘go!’. This time, if the nasal leaves its V to join the following C slot, the preceding vowel can lengthen to take its place. Numerous Bantuists have worked on the problem of NC clusters (see Downing 2003 for a recent statement and references). Finally, most of the arguments in favor of moras developed in Hyman (1985) were based on African languages, particularly Gokana (Lower Cross), where there is no evidence for syllable structure above the moras.

Among the other successes of of the CV tier was its ability to characterize geminates and their “inalterabilty”. Among the major examples were two Afroasiatic languages: Hausa (Chadic) (Hayes 1986) and Tigrinya (Ethiopian Semitic) (Schein & Steriade 1986, Kenstowicz 1982).
Some readers may also be familiar with the contribution of another Afro-Asiatic language, Imdlawn Tashlhiyt (Berber), to the study of syllabification, particularly of consonantal nuclei (Dell & Elmedlaoui 1985, 1988), which provided one of the centerpieces in the development of optimality theory (Prince & Smolensky 1993).

3.4. Reduplication

The vast majority of African languages belong to the Niger-Congo family, and virtually all of these exploit partial reduplication as a morphological process. In West African languages such as Akan (Kwa) (Schachter & Fromkin 1968), the reduplicant consists of a CV copy of the base verb, except that the vowel must be [+high]. I cite examples from Nupe in (14).

(14) CV reduplication in Nupe

a. /gí/ ‘eat’ → gi-gí ‘eating’
   /ge/ ‘be good’ → gi-ge ‘goodness’
   /gà/ ‘separate’ → gi-gà ‘separating’

b. /gú/ ‘puncture’ → gu-gú ‘puncturing’
   /gó/ ‘receive’ → gu-gó ‘receiving’

The Nupe data comes up in the context of an argument in favor of abstract phonological representations (Hyman 1970), whereas corresponding Akan forms are cited both by Wilbur (1974) and Marantz (1982) for their “overapplication” property. Again, this can be illustrated from Nupe, where underlying /ts, dz, s, z/ are palatalized to [c, j, s, z] before front vowels. The issue is that a verb like /tsà/ ‘choose’ reduplicates as tsi-tsà, not as *ci-tsà. The above-cited authors revert to rule ordering (palatalization precedes reduplication), whereas others have used this kind of African data to argue for a global “identity constraint” (Wilbur 1974), which is easily implemented as a base-reduplicant identity correspondence within optimality theory (McCarthy & Prince 1999).

While the above gives some idea of how West African CV reduplication has contributed to phonology, Bantu CVCV verb stem-reduplication has also contributed to the development of prosodic morphology. The verb stem is a constituent consisting of a root plus one or more suffixes. In a number of Bantu languages, but not all, the preposed reduplicant may or must be exactly two syllables in length. Thus, in Kinande, túm-ir-an-a ‘send to each other’ (root-applicative-reciprocal-final vowel) obligatorily reduplicates as túm-a + tum-ir-an-a ‘send to each other here and there’ (Mutaka & Hyman 1990). Interestingly, sw-a ‘grind’ reduplicates as swa-sw-a + swa ‘grind here and there’, where the bisyllabic reduplicant, created by double reduplication, is actually longer than the base verb stem. Odden (1996) shows that in Kerewe there is some choice in how long the reduplicant can be. Hence, lim-il-an-a ‘cultivate for each other’ (applicative -il-, reciprocal -an-) may reduplicate aslim-il-an-a,lim-il-a,lim-il-an-a,lim-il-an-a,lim-il-an-a,lim-il-an-a.

The apparent truncation observed in reduplicated forms such as túm-a + tum-ir-an-a has also raised interesting morphological questions. Downing (1999ab, 2000) sees the final -a as an
indicator that the reduplicant is a morphological constituent, while Hyman, Inkelas & Sibanda (1998) present evidence from Ndebele that the reduplicant is obtained by morphosyntactic doubling, but is subject to additional prosodic restrictions (cf. Inkelas 2002 and Zoll 2002). (See also §4.1.)

3.5. Syntax-phonology interface

The preceding subsection indicates the importance of African languages in the study of the morphology-phonology interface. In this context we can also mention Pulleyblank 1986 for extensive treatment of African tone within the framework of lexical morphology and phonology and Myers 1987 for original contributions on integrating morphology, phonology and syntax, based on Shona. From early generative phonology to the present, African languages had already been central in the study of the syntax-phonology interface. Among the earliest and most informative documentations of this interaction are Kisseberth & Abasheikh’s (1974) treatment of syntactically conditioned vowel length alternations in Chimwiini and Clement’s (1978) treatment of syntactically conditioned tonal alternations in Ewe. Studies such as these informed Chen’s (1987) analysis of Xiamen tone sandhi, from which Selkirk (1986) generalized her end-based theory of derived domains, based largely on Chimwiini. Several of the contributions in Phonology Yearbook 4 (1987) and Inkelas & Zec (1990) deal with the syntax-phonology interface in African languages. Both Kaisse (1985) and Hayes (1987) cite earlier manuscript versions of Odden (1987) on the phrasal phonology of Kimatumbi to support their views on this interface. More recently, Truckenbrodt (1995, 1999) attempts to bring together phonological phrasing, syntactic representation and focus drawing extensively Bantu languages such as Chimwiini, Kimatumbi, and Chichewa—the Chichewa facts deriving from Kanerva’s (1990ab) important work on focus phrasing in that language.

This completes my survey of phonological issues to which African languages have contributed in a significant way. There are doubtless others, and perhaps some phonologists or Africanists will take issue with the choice of issues or specific omissions. As stated in the outset, it is hard to cover the diverse phonological properties of African languages in a short article (see Creissels’ 1994 monograph and Clements 2000 for a recent survey of the phonology of African languages. From the sampling just seen, it is safe to say that African languages have been prominent in almost all of the major phonological developments over the past half century—as I suspect they will continue to be in future developments in the field.

4. Contributions of phonological theory to African languages

The second part of my assignment was to discuss how phonological theory has contributed to the study of African languages. This is a harder task. While it is easy to show how languages provide grist for theoretical mills, showing how theories benefit the study of languages is at best more complicated and perhaps even controversial. There are, of course, language specialists who may only be sympathetic to some general notion of theoretical linguistics, but not to the kind of formalizing for which phonological modeling has been known.

The position I will take here is that phonological theory has made important contributions to the study of African languages, but sometimes in a roundabout way. I give examples below,
but to anticipate, linguistic theory has provided part of the training that is necessary to do good work. In the most successful situations, the result has been a give-and-take between theory and description and between general and African linguistics.

To show this, let us first consider what an Africanist (or other language specialist) might expect from linguistic theory in general—that is, what are the potential contributions of linguistic theory to language-specific work? In response, I believe that theory can serve both as a guide to discovery and as a tool for expressing insights. That is, it can help us deal both with the unknown as well as the known, as I develop in the following two subsections.

**4.1. Theory as a guide to discovery**

As mentioned, a theory can help guide the research towards important discoveries. It might lead us to ask questions we might not have otherwise asked and seek data we might otherwise not have considered. In addition, theoretical awareness can help the researcher see connections that might otherwise not be made, as well as pinpointing problems that might have been overlooked.

Concerning the areas discussed above, I think that this kind of theoretical contribution is most visible in the case of prosodic morphology, which has had a strong and positive effect on the study of reduplication, especially in Bantu. While Prince & McCarthy’s (1986) seminal work dealt relatively little with African languages, the discovery that reduplicants tend to be characterizable in terms of prosodic units (moras, syllables, feet) reverberated among Africanists. Although Bantuists already knew that both nouns and verbs can typically be reduplicated, few realized the richness and complexity that would soon be documented in such languages as Kikuyu (Peng 1993), Kihehe (Odden & Odden 1985), Kinande (Mutaka & Hyman 1990), SeSwati (Downing 1994), Kikerewe (Odden 1996), Chichewa (Myers & Carleton 1996, Hyman & Mtenje 1999), and so forth. In fact, it is rare to find richly descriptive accounts of reduplication in Bantu languages prior to this period. With attention focused elsewhere, descriptivists did not typically go into the kind of detail that is necessary to sort out the parameters which determine, for example, how a verb stem can be reduplicated in different Bantu languages.

As an example, consider reduplication in Ndebele (Hyman, Inkelas & Sibanda 1998). As seen in (15), the reduplicant is preposed to the base in Ndebele and is limited to two syllables:

(15) Plain verb stem Reduplicated verb stem

<table>
<thead>
<tr>
<th></th>
<th>Plain verb stem</th>
<th>Reduplicated verb stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>lim-a ‘cultivate’</td>
<td>lim-a+lim-a ‘cultivate here &amp; there’</td>
</tr>
<tr>
<td></td>
<td>thum-a ‘send’</td>
<td>thum-a+thum-a ‘send here &amp; there’</td>
</tr>
<tr>
<td>b.</td>
<td>nambith-a ‘taste’</td>
<td>nambi+nambith-a ‘taste here &amp; there’</td>
</tr>
<tr>
<td></td>
<td>thembuz-a ‘go from wife to wife’</td>
<td>thembu+thembuz-a ‘go from wife to wife here &amp; there’</td>
</tr>
</tbody>
</table>

As in most other Bantu languages, verb stems must end in an inflectional suffix, e.g. the final vowel, -a. While the first two syllables of a verb root must be reduplicated in (15b) and forms such as *namba+nambith-a and *thembu+thembuz-a are ill-formed, the derivational suffixes -el-
(applicative) and -is- (causative) may optionally be truncated and replaced by -a in the reduplicants in (16).

\[\begin{align*}
(16) & \quad \text{a. lim-el-a} & \quad \text{lim-e+lim-el-a} & \quad \text{‘cultivate for/at’} \\
& & \quad \text{lim-a+lim-el-a} \\
& \quad \text{b. lim-is-a} & \quad \text{lim-i+lim-is-a} & \quad \text{‘make cultivate’} \\
& & \quad \text{lim-a+lim-is-a}
\end{align*}\]

Again, the reduplicant must consist of exactly two syllables. As seen in (17), when the verb stem is monosyllabic, consisting of the subminimal root dl- ‘eat’ and the final vowel -a, the bisyllabic required on the reduplicant is fulfilled by the appearance of the default syllable [yi]:

\[\begin{align*}
(17) & \quad \text{dl-a} & \quad \text{‘eat’} \\
& & \quad \text{dla-yi+dla} & \quad \text{‘eat here & there’}
\end{align*}\]

A brief account might stop here, but the question naturally arises as to how such subminimal verbs reduplicate when they are extended by an applicative or causative suffix. As seen in (18), there are several options:

\[\begin{align*}
(18) & \quad \text{a. dl-el-a} & \quad \text{dl-el-a+dl-el-a} & \quad \text{‘eat for/at’} \\
& & \quad \text{dl-a-yi+dl-el-a} \\
& & \quad \text{dl-e-yi+dl-el-a} \\
& \quad \text{b. dl-is-a} & \quad \text{dl-is-a+dl-is-a} & \quad \text{‘make eat’} \\
& & \quad \text{dl-a-yi+dl-is-a} \\
& & \quad \text{dl-i-yi+dl-is-a}
\end{align*}\]

As in the first option in (18a,b), one can have total reduplication of the verb stem. Alternatively, one can truncate -el- or -is-, and use a final vowel -a followed by [yi] in the reduplicant. As seen in the third examples, the reduplicant can alternatively repeat the [e] or [i] of the verb suffix followed by [yi]. If one puts together the facts from (15), (16) and (18), the generalization which emerges is that Ndebele requires that the verb root be as exhaustively parsed in the reduplicant as possible. (Within optimality theory this is nicely captured by high ranking of the constraint MAX(Root); cf. McCarthy & Prince 1993, Urbanczyk 1995, Futagi 1997.) In contrast, (19) shows that the inflectional suffixes -e (subjunctive), -i (negative), and -ile (perfective) may not appear in the reduplicant at all:

\[\begin{align*}
(19) & \quad \text{a. lim-e} & \rightarrow & \quad \text{lim-a+lim-e} \\
& & & \quad *\text{lim-e+lim-e} \\
& \quad \text{b. lim-i} & \rightarrow & \quad \text{lim-a+lim-i} \\
& & & \quad *\text{lim-i+lim-i} \\
& \quad \text{c. lim-ile} & \rightarrow & \quad \text{lim-a+lim-ile} \\
& & & \quad *\text{lim-i+lim-ile}
\end{align*}\]

We thus can conclude the following about the reduplicant in Ndebele verb-stem reduplication: Root material is obligatory, derivational material is optional, and inflectional material is prohibited. Hyman, Inkelas & Sibanda (1998) further show that morphosyntactic structure are
implicated in Ndebele verb stem reduplication—it matters whether something is an applicativized passive or a passivized applicative! The important point here is that theoretical questions have inspired us to investigate reduplication in a systematic and rigorous way. The result has not only been to contribute back to theory, but in fact we get a “thicker” description and better understanding of the languages in question.

4.2. Theory as a tool for expressing insights

In the preceding subsection I have indicated how phonological theory has led to new insights and discoveries concerning how African languages work. The second contribution of phonological theory has been to provide a tool for expressing insights in a revealing way. A particularly clear example of this is how autosegmental phonology provided an explicit framework for expressing the intuitions field researchers had about tone.

Consider, for example, the representation of tone spreading. In Aghem, a prefixal H tone spreads onto a following L tone root, thereby creating a HL falling tone: /é-zù/ → [é-zú] ‘to skin’. In pre-autosegmental phonology, Hyman & Schuh (1974) expressed such a rule roughly as in (20a).

\[(20) \text{Two views of tone spreading} \]
\[\begin{align*}
\text{a. } & \quad \text{H-L } \rightarrow \text{ H-HL} \\
\text{b. } & \quad \text{V C V} \\
& \quad \text{} \\
& \quad \text{H L}
\end{align*}\]

Goldsmith’s (1976) autosegmental representation, on the other hand, is shown in (20b). Whereas Goldsmith’s representation clearly indicates that there is a single H feature involved in tone spreading, Hyman & Schuh’s formulation implies that a H feature is being copied onto the following vowel as in the SPE conception of assimilation. However, consider Hyman & Schuh’s (1974) prose statement about what they feel is going on:

“Spreading is an assimilatory process of the progressive or perseverative type, rather than of the regressive or anticipatory type. That is, the earlier tone appears to last too long, rather than the later tone starting too early. This in fact is the way that we should like to view this phenomenon. There is no process of tone copying or tone addition in the second syllable. Rather, the earlier tone simply enlarges its domain.” (p. 88)

Clearly Hyman & Schuh had something in mind that they could not formalize, but which is conceptually identical to the autosegmental representation of tone spreading in (20b).

As a second example, consider Leben’s (1973) proposal that tone is suprasegmental in Mende. Rather than being a property of (syllabic) segments, both Leben and Goldsmith developed models in which Mende words of varying sizes could be characterized in terms of five tonal melodies: L, H, LH, HL, LHL. In §2.1 I mentioned Paulian’s (1975) independent and converging research on Kukuya, in which stems of one, two or three syllables clearly are restricted to one of these same tonal melodies. However, one can go back at least as far as
Welmers (1962) to find the same insight, this time concerning Kpelle, a Southwest Mande language closely related to Mende: “tonemes must be analyzed in terms of segments between two open transitions” (p.85). Welmers describes “the five types of forms” in Kpelle as follows (presented with his transcriptions):

(21) Five tonal melodies of Kpelle (Welmers 1962:86)

a. High throughout

| pá  | ‘come’       |
| bóa | ‘knife’      |
| láa | ‘lie down’   |
| píli| ‘jump’       |

b. Mid throughout

| kpɔŋ | ‘help’       |
| sua  | ‘animal’     |
| see  | ‘sit down’   |
| kali | ‘snake’      |

c. High followed by low (low begins on the next vowel if there is one)

| yë  | ‘for you’   |
| tôa  | ‘pygmy antelope’ |
| kpɔŋ | ‘door’      |
| kâli| ‘hoe’       |

d. Mid with first vowel, then high followed by low

| tɛë | ‘black duiker’ |
| konã| ‘mortar’      |
| yuð | ‘axe’         |
| kpanâŋ | ‘village’ |

a. Low throughout

| kpòo | ‘padlock’ |
| kpàki| ‘loom’    |
| tɔnɔ | ‘chisel’  |
| tőloŋ| ‘dove’    |

Note, first, that Welmers uses only one tone mark per word. He thus writes /kâli/ for what is pronounced [kálì] ‘hoe’, i.e. H-L. Second, there is no difficulty reducing Kpelle to an underlying two-level system: The M that occurs in the MHL melody in (21d) can be analyzed as a L which is raised before H, and the “M throughout” melody in (21b) is underlyingly /L-H/, as is seen when two “Mid throughout” words occur in sequence:

“In mid-mid, for the dialect being described here, the first mid has a slightly rising allotone... In some areas, the first mid is level, but the second mid begins a little higher and drops quickly to the level of the first. In still other areas, both phenomena occur: the first mid ends a little higher, and the second begins a little higher. In all cases, the conjunction of two mids is accompanied by an upward pressure.” (p.87, note 2)

Welmers goes to considerable trouble to justify his suprasegmental analysis, with one tonal melody per word, or, in his terminology “one toneme between two open transitions” (p.86). What is clear is that he had the same insight as Leben, Goldsmith, and Paulian concerning the semi-autonomy of the five tonal patterns. Like Hyman & Schuh, he did not have an adequate framework such as autosegmental phonology to express this insight.
4.3. Interaction of the two contributions of theory to African languages

While one result of autosegmental phonology was to provide a notation that allowed Africanists to express their insights, it is the interaction of the two contributions of phonological theory that accounts for the particularly strong impact of the latter. For this demonstration I will stay with the issue of tone.

Recall the discussion of floating tones in §2.1. It was seen in (6), that although both are pronounced H-H in isolation, the two Aghem nouns kí--fú ‘rat’ and kí--wó ‘hand’ have very different tonal behavior: The H of kí-fú spreads onto a following L, while the H of kí-wó doesn’t. The noun kí-wó causes a following H to become downstepped, but the noun kí-fú does not. The proposal was to recognize an underlying floating L tone in the stem /-wó/ which both prevents the preceding H from spreading, as well as conditioning a downstep.

I first presented this analysis in a class on phonological theory and African tone which I co-taught with Will Leben at the Stanford LSA institute in 1987. After presenting the above analysis, one student raised her hand and objected: “But you are putting in the floating L tone only because it works.” I responded emphatically, “Yes, right! That’s why I’m putting it in.” I don’t recall what else I said, or if I satisfied the student’s objection or not. However, I have thought about what I do remember about this exchange many times since. If I had the opportunity to respond again, here is what I would now say:

First, it is clear that there are two kinds of H tone stems in Aghem. One is like -fú ‘rat’, the other is like -wó ‘hand’. Anyone working on Aghem has to account for this fact. The question is how? It seems to me that we have two choices. On the one hand, we can use a diacritic, which can be placed either on the TBU, e.g. -fú₁ vs. -wó₂, or on the tone, e.g. H₁ vs. H₂. On the other hand, one can use floating tones, as proposed above. Having established this much, we can now ask the further question: What’s the difference?

Phonologists will immediately argue that the floating L is “explanatory”, whereas the diacritic is not. The argument is based both on elegance as well as naturalness: A floating L tone is expected to block a preceding H from spreading onto a following L, just as it is expected to cause a downstep on a following H. That is, the effects of the floating L are “natural” in the sense of phonetically plausible. Diacritics, on the other hand, make no such claims: they could have produced any tonal effect and in any combination. For example, rather than the way the two tones work in (6), a diacritic analysis could as easily distinguish a H₁ which spreads and produces downstep vs. a H₂ which neither spreads nor produces downstep. It would be hard to express such a distinction in a floating tone analysis. Since such an alignment of properties is unknown, this is a good result, and hence a strong argument for floating tones, if not ultimately for autosegmental representations.

It should be said that not all linguists would find this argument persuasive. There are those who might argue that there is no synchronic “explanation”, but rather the properties of Aghem H tones are the way they are because of historical events. We know that the floating L tone is due to the loss of a syllable (cf. Proto-Bantu *-bókò ‘hand’). So, this argument goes, from the
synchronic point of view we should look at the floating L tone as another kind of diacritic. In other words, there’s really no difference between the two approaches.

I would like to conclude by giving another reason to adopt the floating tone analysis of Aghem, which is really an argument in disguise for doing phonological theory:

Let us assume that the hypothetical nay-sayers are correct. Let’s even assume that there is no empirical difference between using diacritics or floating tones to account for Aghem or other comparable facts. If we use diacritics, we have a one-to-one representation of “the facts”. But does the diacritic analysis reveal anything interesting about Aghem? Does it naturally allow us to generalize to any other property of the language or to anything that occurs in any other language? Does it raise new questions or lead one to establish a research agenda? I think the answer to all of these questions has to be “no”.

Now compare the floating tone analysis. Here we have more than a one-to-one representation of “the facts”. We have a proposed representation which can be evaluated and applied to any number of situations. For example, since floating tones are sequenced, we can ask where they can vs. cannot occur. Can two or more floating tones occur in sequence (the answer is yes!). Since they have a featural composition, natural questions concern whether all tones (H, M, L etc.) can float? Do floating tones occur in the output as well as the input, or must they link in some feature-geometric structure in the postlexical phonology? We also can ask which other features may float, and whether these floating features have exactly the same properties as floating tones? The final question I want to raise concerns what a floating tone can do vs. not do? Whereas diacritics are presumably free to do anything that we might stumble on, there is reason to believe that floating tones are constrained in an identical way to linked tones.

To see the value of this last point, consider the Kpelle data in (22):

(22) 3sg prefix in Kpelle

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welmers</td>
<td>Bird</td>
</tr>
<tr>
<td>a.</td>
<td></td>
</tr>
<tr>
<td>` -pòlu</td>
<td>m-pòlu</td>
</tr>
<tr>
<td>` -túe</td>
<td>h-túe</td>
</tr>
<tr>
<td>` -kóo</td>
<td>ñ-kóo</td>
</tr>
<tr>
<td>` -fíi</td>
<td>m-fíi</td>
</tr>
<tr>
<td>` -sèñ</td>
<td>h-sèñ</td>
</tr>
<tr>
<td>b.</td>
<td></td>
</tr>
<tr>
<td>` -barâñ</td>
<td>m-barâñ</td>
</tr>
<tr>
<td>` -lee</td>
<td>h-lee</td>
</tr>
<tr>
<td>` -yée</td>
<td>ñ-yée</td>
</tr>
<tr>
<td>` -wóli</td>
<td>ñ-wóli</td>
</tr>
</tbody>
</table>
As seen, Welmers (1962) posits a floating L tone 3sg prefix in the inputs, whereas Bird (1971) (and subsequent scholars) have posited a L tone nasal prefix. As seen in the outputs in (22a), stem-initial voiceless obstruents acquire L tone full voicing after the prefix and are apparently longer in duration (Dwyer 1974). In (22b) we see that stem-initial sonorants (including the bilabial implosive) become a L tone nasal which both scholars transcribe as short.

There are two questions of interest concerning Welmers’ analysis. First, can a floating L tone cause voicing (and potentially gemination) of a obstruent? Second, can it cause nasalization of a sonorant consonant? Hyman & Schuh (1974) claim that tones do not affect consonant voicing (rather, it is the reverse), and although a few counterexamples have been cited, one must acknowledge that such examples are rare at best. Concerning nasality, there is no reason for a L (or H) tone to nasalize a consonant. Instead, both voicing and nasalization are expected to occur after a nasal prefix, e.g. after the H tone 1sg nasal prefix, e.g. m-bólu ‘my back’, n-dúé ‘in front of me’, n-máran ‘my companion’, n-ée ‘my mother’ etc. We thus concur with Bird (1971) that the 3sg prefix is a L tone nasal which has the same effects as the 1sg H tone prefix, except that it drops out after voicing an initial obstruent.

The question “can a floating L tone do this?” is of course one that makes little, if any sense, if asked instead of a diacritic. In the above case the diacritic alternative would presumably refer directly to [+3sg]. Diacritics come in all shapes and sizes and are at times required to account for unpredictable properties within phonology and morphology. Should we then view the tonal diacritics considered in the Aghem case as morpheme features, e.g. like declension class features? As this question indicates, the diacritic analysis sends us off in a different direction from the floating tone analysis: It encourages us to relate the tonal diacritics with apparently unrelated phenomena. It is of course logically possible to ask the same question about the floating L in /-wó/ ‘hand’, but we normally wouldn’t. That’s because the decision to invoke floating tones sends us off in a clearly phonological direction—which is in this case correct.

To summarize this last argument, the reason to use floating tones is the same reason why we should adopt the representations made available in autosegmental phonology (cf. also Clements & Ford 1979 proposal to represent downstep as an unlinked L tone). The advantage of a theory is that it can simultaneously serve as a guide to discovery and a tool for expressing insights. When successful, a theory or framework opens doors. That’s what autosegmental phonology has over the kind of phonology that preceded it and over analogous, intuitive approaches which are not formalized. That’s also what floating tones have over diacritics. We have gone in one vs. another theoretical direction arguing on the basis of truth, elegance, naturalness, and interest. People choose what to work on based on a wide range of factors, high among which is their own personal interest. For a general linguistic theory to have interest for those of us involved in the analysis of languages, it should send us off into new areas where discovery and insight converge. Not discussed in this paper are cases where the theory-description connection is not optimal, e.g. when “applying” a theory to a set of data becomes an exercise, or when a theory becomes descriptively non-friendly. I discuss these and other issues with respect to African languages in Hyman (2003). For my present purpose it suffices to say that phonological theory has opened a lot of doors for Africanists, and we have entered to great advantage.
Acknowledgements

I would like to thank my colleagues Sharon Inkelas and Lynn Nichols for very stimulating discussions of some of the issues I will discuss, as well as Lisa Cheng for her extraordinary patience in waiting for the completed manuscript.

References


Rose, Sharon & Rachel Walker. 2003. A typology of consonant agreement as correspondence. [Rutgers Optimality Archive; ROA-458-0801].