

Why Describe African Languages?

Larry M. Hyman

University of California, Berkeley
hyman@socrates.berkeley.edu

Whether formal modeling is treated simply as programming for some practical purpose, or as a method of investigating the properties of the cognitive systems involved, it can and should be separated in most cases from the problem of determining the facts and the descriptive generalizations. (Akinlabi & Liberman 2000:29)

I have been asked to speak on the theme this conference, which I repeat in (1).¹

(1) *The description of African languages in a rapidly changing field.*

I think that it is particularly timely for us to focus on this theme, not only in terms of how things look today, but also concerning our future. It is particularly fitting that this discussion take place at Rutgers, whose dramatic growth and excellence in both linguistics and African studies are well known to us all.

As you will hear, I have some things to say on the subject, but unfortunately, I will have as many questions for us as I do answers. I apologize in advance for my perspective as someone who works in the United States, but I think that much of what we have to discuss will have parallels wherever the description of African languages takes place. I also apologize for any Benue-Congocentrism that may be apparent.

So, let us begin by unpacking the content words in the conference theme and ask the questions in (2).

(2) Interpreting the theme

- a. what “field”?
- b. how is it “rapidly changing”?
- c. what is meant by “description”?
- d. what counts as an “African language”?²

The answer to the first question, “what field are we talking about?”, may seem obvious to North Americans, who will immediately answer, “linguistics”. This is because, unlike Europe and Africa itself, the study of African languages did not begin until the modern linguistic era. The first dissertation written on an African language in the United States was entitled “A Grammar of Chichewa” written by Mark Hanna Watkins in 1937 and directed by Edward Sapir at the University of Chicago. Shortly after that, we begin to see the descriptive work of my teacher, William Welmers, who in turn had studied with Zellig Harris at the University of Pennsylvania. Setting aside

¹ I would like to express my thanks to Akinbiyi Akinlabi and the other organizers for the invitation and the opportunity to speak at the combined WOCAL4/ACAL34 conference, as well as the many thoughtful colleagues with whom I was able to speak about the issues of concern to me in this talk. Thanks also to my Berkeley colleague, Lynn Nichols, for her reactions to a draft of this paper and for many stimulation discussions of the issues treated herein—with which we are both concerned.

² I will have nothing to say about this last question.

the unique case of Joseph Greenberg, who was a linguist unto himself, African linguistics easily makes the transition into generative grammar in the mid 1960s and up to this day. As a consequence of America's late entry, and with very few exceptions, the study of African languages takes place within linguistics departments in both the United States and Canada.

In contrast, unlike North America, the study of African languages easily predates 20th century synchronic linguistic frameworks. In Bantu studies, for example, research by Bleek, Meinhof, Guthrie, Meeussen and their students has a distinctly comparative and diachronic character that begins more neogrammarian than structuralist. Whereas many African universities have departments that join the study of African languages with linguistics, in Europe, African linguistics may either be a separate department or part of African studies, or even a museum, e.g. le Musée Royal de l'Afrique Centrale in Tervuren.

Be this as it may, it seems that the trend is for Africanist linguists to identify with the field of general linguistics, and so we must ask the questions in (3) and (4).

- (3) How is the field of linguistics rapidly changing?
- (4) How are these rapid changes likely to affect the description of African languages?

In response to these questions, there's good news and bad news. The good news is that all of linguistics seems now to accept, if not enthusiastically encourage, the study of "endangered languages" as well as "minority languages", or what I generally refer to as "unempowered languages". We can debate how deep this commitment is, but we have come a long way since my early generative student days, when one professor said to me that we should all work on English, and that we weren't ready yet to work on any other language, except maybe French.

Times have certainly changed in this respect. There is now an Association for Linguistic Typology and both its journal and many others welcome good descriptive and cross-linguistic work. I think most linguists accept that we cannot appreciate the rich possibilities inherent in language without considering a representative sample of the world's 6,000 languages, about 2,000 of which are to be found in Africa. Everyone knows that African languages have made special contributions to general linguistics, e.g. in the areas of tone, ATR vowel harmony, serial verbs, grammaticalization, focus constructions, and so forth. However, if I were to raise the question, "Why describe African languages?", as I do in my title, I think it would be incomplete to respond, "because of their importance to linguistic theory." Africanist linguists are accustomed to making this case, and the ACAL conferences have had guest lectures and symposia on subjects like "contributions of African languages to syntactic theory" or "to phonological theory" and so forth.

Why describe African languages? Do linguists defend the study of African and other "exotic" languages out of pure self-interest? Such a charge has been made in a recent *New York Times* essay by science writer David Berreby, from which I quote in (5).

- (5) For the past decade, scholars and political activists have been working to get the rest of us worried about the future of the world's 6,000 or so spoken languages.... It is no surprise that linguists and activists promote maintaining spoken languages. Just as the Poultry and Egg Council wants us to eat eggs, linguists want languages to study.... It would be a terrible thing to run out of

languages. But there is no danger of that, because the reserve of language, unlike the gas tank, is refueled every day, as ordinary people engage in the creative and ingenious act of talking. Old words, constructions and pronunciations drop away, new ones are taken up, and, relentlessly, the language changes. (David Berreby, *New York Times*, May 27, 2003)

I'm not sure if it was published, but Jerrold Sadock wrote an effective response from which I excerpt in (6).

- (6) David Berreby's recent article denying the significance of the world-wide extinction of languages is insensitive and uninformed. What he fails to appreciate is that each living language is a special record of the history of a culture dating right back to the origins of modern man. This record is engraved in the collective subconscious of living speakers, as a special mode of thought, a special set of 'thought grooves', as Edward Sapir put it eighty years ago. When the chain of transmission of a language in a society is broken, the unique pattern of thinking that it bore ends and, contrary to what Berreby says, cannot be resurrected. Even a large dictionary and grammar do not contain anything like the deep, unfelt intuitions that a native speaker automatically develops. Written records are to a living language what a movie of a panda is to a real panda. (Jerrold M. Saddock)

Within this context, Africa has a special status as the place of origin of the human species, and consequently of language itself. As Tucker Childs (2003) puts it in his recent *Introduction to African Languages*, in (7),

- (7) In addition to their purely linguistic fascination and the urgency to their study, African languages are appealing from a socio-historical perspective, not the least for their widespread legacy. If Africa is indeed the continent where the human species first appeared, then perhaps studying the continent's languages can provide some insights into how language arose and spread, and how it has changed over space and time.... In addition, the study of languages can contribute to understanding Africa's history... and the reverse is true as well. (Childs 2003).

Why describe African languages? The loss of a language, African or otherwise, is a serious event which is compounded by everything that we failed to do in terms of documentation. Least of our failures may be our inability to know what it might have told us about Language in general. Once in 1970 when I visited Robert Armstrong, then director of the Institute of African Studies in Ibadan, I walked in on a session where he and his assistants were transcribing some mega-recitations of Idoma chants, which had been laboriously recorded over a period of weeks with a single reciter. When I expressed my fascination, Armstrong replied, "Imagine all of the Iliads and Odyssey's that are walking around out there", succinctly characterizing the importance and vastness of the situation. Even if a language is not itself endangered, certain aspects of the culture may be. So, we are sensitized to the urgent need to get out there, even if many of us don't drop everything and go and do it.

As general linguists, we want to describe languages, by which we usually mean, grammar. But as such, we are subject to the rapidly changing field of linguistics whose theories are not necessarily description-friendly. What would a minimalist grammar of Luganda look like? Or a reasonably complete OT phonology of that very complex language? These and other frameworks work well for certain things, not so well for others, and, in any case, their goal is not to provide a framework in which a descriptive grammar might be written.

Let's consider linguistic theory for a moment. I myself tend to view theory as an aid in discovery, or in expressing insights. Formal models help us discover new things,

patterns or relationships within a grammar which might not otherwise have been evident. The dominant assumption, however, is that there is a relation between doing formal theory and studying the human mind. We often hear that language is a mirror into the mind, but it is also a mirror into culture and into history. Just as we carry human history in our genes, so do we carry human history in our language. We must describe the 2,000 African languages not only because they tell us something about Language (with a capital L), but also because they tell us about human history: Where do these people and their language come from? Who were they related to? With whom did they have contact? What is or was their “special mode of thought”, as Saddock put it, their wisdom and unique contribution to the human experience? Descriptive grammars and lexicons contain hidden markers that bear on these questions which obviously transcend Language (even with its capital L).

I ask myself therefore how linguistics is rapidly changing, and whether it will be good for those of us who want to describe African languages. Consider the statement in (8) which occurs in a report prepared for the Ford Foundation:

(8) The Special Case of Linguistics

“It developed originally as a branch of cultural anthropology and philology but has developed in past decades as a branch of cognitive science.” (Professional Master’s Programs in the Social Sciences: Current Status and Future Possibilities, a report to the Ford Foundation, Council of Graduate Schools, p.18)

First of all, I don’t know if I want to be a branch of anything, but we have to be a little concerned about any subordination to cognitive science. There has been an increasing trend for linguists to be inspired by computational or neurocognitive research which seems quite distant from the goals of describing African languages. Should such research become the main activities of linguistics departments, even so-called theoretical linguists who work on African languages may be in trouble.

Which brings me to still another thing that is rapidly changing in all fields: technology. The development of computers has greatly facilitated our work in more ways that I can enumerate. The laptop is but one piece of equipment that can be taken into the field to help in the recording, storage, and analysis of African language data. I myself have found it greatly beneficial to produce lexicons directly in a data base program, which can be immediately searched to discover distributional generalizations and test hypotheses. With websites and email, the possibilities for sharing information have increased dramatically, and intercontinental exchanges have been greatly facilitated. See, for example, the following website address which can be consulted to find appropriate resources for the study of African languages:

<http://www.african.gu.se/webresources/index.html>

So, if I were to summarize thus far, I’d say that it’s been a good relationship with linguistics overall, and that we have been able to make extraordinary progress in describing African languages because we have been trained as linguists—but that our future relationship is uncertain as long as we don’t know where linguistics is going. I might add parenthetically, that some major linguistics departments in the US don’t have a field linguist or a historical linguist—and there appears to be a trend for fewer such specialists to be hired in foreign language departments as well. We must be concerned that Africanist and other descriptive and historical linguists, however theoretically minded, do not go the way of philology and Indo-European.

This brings me to the key word in the theme of the conference, which is “description”. What is meant by description? Description of what? And as opposed to what?

In general linguistics the typical discussion concerns theory vs. description. The mantra that was repeated to me as a graduate student is that you can’t describe without having a theory. When I protested, naively suggesting that certain of my work was “just” description and that I had no theory, my professor responded, “No, you have an implicit theory.” After all, I used such concepts as abstract underlying vowels and floating tones, which clearly could not be observed, let alone measured, so it must be theory. And since you can’t do description without theory, please everyone do theory, preferably their theory.

Well, there’s been a lot of discussion in linguistics about what science and theories are about. In a recent discussion among phonologists concerning what constitutes a “theoretical contribution”, a distinguished colleague wrote in (9),

(9) What constitutes a “theoretical contribution”?

The shared belief of many in the field appears to be that a paper making a theoretical contribution must (a) propose some new mechanism, which adds to or replaces part of some current theory, or (b) contradict some current theory. Papers that do neither, or those that do either but in a relatively minor way, are not looked at as making a theoretical contribution.

By this view, papers that provide a compelling analysis of an interesting, perhaps previously unattested linguistic phenomenon would presumably not be seen as theoretical. Hmm.

What I would like to do here is make the case that description is essentially indistinguishable from theory. Already, by using the phrase “in a relatively minor way”, the colleague in (9) seems to be saying that the notion of being “theoretical” is somehow scalar. I’m a little slow sometimes, but what I’ve increasingly come to realize is that description and theory are very hard to disentangle—and, when each is done right, they have the same concerns. But what do we mean by “description”?

First, consider in (10) what description is not.

(10) Description is not...

- a. data
- b. a list
- c. an exercise

You hear people contrast “theory” and “data”, as in “some people are more interested in the data than in theory”, or “we have enough data, what we need is more theory”. There’s a sense in which Robert Armstrong’s actual tape recordings of the Idoma chants are data, but both the transcription and the description require an analysis of considerable sophistication. That’s why description cannot also be a list, whether an enumeration of minimal pairs, a word list, or a list of sentences in a language. A description must be self-consciously analytical.

I have indicated in (10c) that a description should also not be an exercise, but the same holds of theory—and specifically about studies that try to combine the two. I

think of certain articles, mémoires and theses that attempt to provide a description in some formal framework: tagmemics, systemic linguistics, transformational grammar, and so forth. You go to these works to find out how a language works, what's going on, what insights the writer has—and instead, you find yourself in a morass of some outdated (or current) framework that you have to unpack, sometimes determining that it wasn't worth the effort (or the money).

Description is not an exercise in how to apply a model. Whether you call it descriptive or theoretical, a PhD dissertation should not consist of an exercise in plugging African data into a pre-existing framework, whether Pike's, Halliday's, Chomsky's or whoever's. In fact, I'm not even sure that most formal linguists would recommend that descriptivists work this way. Consider what Akinlabi and Liberman (2000) have to say in (11).

- (11) The documentation of... descriptive generalizations is sometimes clearer and more accessible when expressed in terms of a detailed formal reconstruction, but only in the rare and happy case that the formalism fits the data so well that the resulting account is clearer and easier to understand than the list of categories of facts that it encodes.... [If not], subsequent scholars must often struggle to decode a description in an out-of-date formal framework so as to work back to... the facts.... which they can re-formalize in a new way. Having experienced this struggle often ourselves, we have decided to accommodate our successors by providing them directly with a plainer account. (Akinlabi & Liberman 2000:24)

In my descriptive work, I like to say that the approach I take to formalisms is strictly opportunistic: If a tableau is worth a thousand words, use it! In other words, having a proper command of theory means not only knowing when to use it, but when not to.

So, what is description? As seen in (12), I consider description to be another word for analysis:

- (12) Description is *analysis* and should ideally be
- | | |
|------------------|--|
| a. rigorous | (it should “compute”, i.e. be accurate) |
| b. comprehensive | (it should cover the required ground, richly exemplified) |
| c. rich | (it should investigate as many aspects of a problem as possible) |
| d. insightful | (it should “explain” or “account for” phenomena) |
| e. interesting | (it should tell you something worth knowing) |

In (12) I've listed five properties of an ideal description. I'll be giving examples shortly, but let me mention each briefly. First, a description should be rigorous and comprehensive such that each of the potential factors is considered and tested as exhaustively as possible against a wide range of contexts. The description should be rich and insightful, drawing on whatever tools or perspectives are needed to tell the full story. One might draw simultaneously from morphology, syntax, and semantics, and there may also be a need to contrast synchrony and diachrony. An ideal description should be richly exemplified, whether from elicitations, texts, or naturally occurring discourse. One should never assume that a description is complete, that the interpretation is definitive, or that we have foreseen all possible purposes to which the description may be applied. But it seems to me that descriptions should be interesting, i.e. have some potential for impact beyond the immediate act of describing.

A moment's reflection will reveal that the notion of analysis and the five qualities in (12) also apply to theory. Whether you do formal modeling or not, both theory and

description are at their exciting best when they take distance from the data and show you that something isn't what it appears to be on the surface. To some extent theory is to description as general is to specific, and that's one reason I always talk about general linguistics, rather than theoretical linguistics.

Well, so much for the general part of this talk. Let's now do some description.

I have chosen three vignettes to illustrate some of the points I've made and concretize some of the reasons for describing African languages. I'll present them, after which we'll try to tie them together to reach some kind of conclusion.

The first illustration comes from the tonal work of Jan Voorhoeve on the Bangangte dialect of Bamileke, spoken in Cameroon, known as Medumba, est. 210,000 speakers by the Ethnologue (1991 UBS). The nouns in (13) all are pronounced with a H tone in isolation:

- (13) High (H) tone nouns in Bamileke-Bangangte (Medumba) (Voorhoeve 1971)
- | | | | | | |
|----|----|---------|----|-----|---------|
| a. | tí | 'tree' | c. | sáj | 'bird' |
| b. | yú | 'thing' | d. | mén | 'child' |

When we put these nouns into the genitive or associative construction, however, different tonal patterns emerge. In (14a), 'tree of bird' is pronounced H-H, while in (14b), 'tree of child' is pronounced H followed by a downstepped H:

- (14) a. tí sáj [- -] 'tree of bird' /sáj/ (noun class 1a, no prefix)
 b. tí [!]mén [- -] 'tree of child' /[^]mén/ (noun class 1, L prefix)

The difference must be due to the second noun, and as seen to the right, the word 'bird' belongs to noun class 1a, which lacks a prefix, whereas all other noun classes have a floating L tone prefix, as seen in the word 'child' to the right in (14b). As shown in (15), a floating L tone that is wedged between two H tone syllables will cause the second one to be downstepped:

- (15) (= syllable; [!] = downstep)
- | | | | |
|---|---|---|--------------------|
| | | | |
| H | L | H | H - [!] H |

Having established that /sáj/ 'bird' has no prefix, now consider the forms in (16).

- (16) a. tí sáj [- -] 'tree of bird' /sáj/ (noun class 1a, no prefix)
 b. yú [!]sáj [- -] 'thing of bird' /[^]yú[`]/ (noun class 7, L prefix)

As before, 'tree of bird' is pronounced H-H in (16a), but 'thing of bird' is pronounced H-[!]H in (16b). In this case, it has to be because of the difference between 'tree' and 'thing'. So, as shown to the right in (16b), 'thing' not only has a L prefix (because it belongs to class 7), but also a lexical floating L tone which is part of its stem and causes downstep.

The final and most intriguing example is in (17).

(17) yú [!]mén [- -] ‘thing of child’ i.e. H + double-downstepped H

In this case, we have a H tone followed by a double-downstepped H. In other words, as summarized in (18), when we put together two H tone nouns in an associative construction in this language, we get one of three outputs:

- (18) a. [H] + [H] H-H e.g. tí sánj ‘tree of bird’ (16a)
 b. [H] + [H] H-[!]H e.g. yú [!]sánj ‘thing of bird’ (16b)
 c. [H] + [H] H-[!]!H e.g. yú [!]mén ‘thing of child’ (16a)

How can we account for the double-downstep in (18c)?

Voorhoeve’s answer was to recognize the three sources of floating tones in (19).

- (19) Three sources of floating tones in the above examples
- a floating L tone prefix in all but class 1a
 - a floating L stem tone following some H tones, e.g. /yú` / ‘thing’)
 - a floating H associative (genitive) tonal morpheme

As seen in (20), the representation for ‘thing of child’ has two syllables, but seven tones, five of which float:

- (20) The double-downstep in (17) results from a floating L H L sequence between two H tone syllables

‘thing’	‘of’	‘child’
yu		mén
L H <u>L</u>	<u>H</u>	<u>L</u> H L
H	! <u>H</u>	! H
	∅	

Both nouns have a floating L prefix, and both have a final floating L stem tone. As seen, the associative marker is a floating H, which is downstepped by the preceding floating L. The H of ‘child’ is in turn downstepped by its L prefix. When the underlined !H associative tone is deleted, as shown, the result is two downsteps in a row, i.e. a double-downstep.

I cite this example for at least three reasons. First, it shows that description is not easy. It took Jan Voorhoeve some time to figure this out, and the rest of us are the fortunate beneficiaries of his toil. Second, it shows the need for description to rigorous. As schematized in (21), in a Bamileke, i.e. Eastern Grassfields Bantu language, in order to study noun + noun associative tones, you need to elicit at least 32 different sequences:

(21) The 32 sequences of Noun + Noun in an Eastern Grassfields Bantu language

4	x	2	x	4	= 32 combinations
prefix-stem	+	associative	+	prefix-stem	
<u>L</u> H		<u>H</u>		<u>L</u> H	
H <u>L</u>		<u>L</u>		H <u>L</u>	(underlined = floating)
L				L	
L <u>H</u>				L <u>H</u>	

As seen, L tone stems can also be followed by a floating H tone, and an associative tone can be H or L. Thus, we must consider 4 x 2 x 4 tone combinations, or a total of 32 ‘noun of noun’ sequences. As shown in (22), if we put prefixless class 1a nouns in each of the two slots, we then multiply the possibilities and have to look at as many as 96 combinations:

(22) 128 combinations if we then look at [+class 1a] vs. [-class 1a] nouns³

a.	[-class 1a]	+	[-class 1a]	=	32	(as in (21))
b.	[-class 1a]	+	[+class1a]	=	32	
c.	[+class 1a]	+	[-class 1a]	=	16	
d.	[+class 1a]	+	[+class1a]	=	<u>16</u>	
					96	

A third, and perhaps most important lesson to take from Medumba is that a careful and detailed description allows us to draw important historical inferences, namely, that despite earlier claims to the contrary, the highly evolved Eastern Grassfields languages are Bantu. Voorhoeve’s floating tones can easily be shown to be the historical result of losing the vowels on which they were once pronounced. Thus, the corresponding seven-syllable Proto-Bantu reconstruction of ‘thing of child’ is shown in (23).

(23)	<i>Proto-Bantu</i>		<i>Medumba</i>	
	*kì-jùmà + kí-á + mù-jánà	>	yú !!mén	‘thing of child’
	L H L H L H L		L H L H L H L	

Well aware of this, Voorhoeve writes in (24):

(24) From a diachronic viewpoint these results are rather satisfactory. The first non-segmental L can be related to the proto-Bantu noun prefix, which has L tone. Both tones of the normal Bantu disyllabic noun stems are preserved. The grammatical morpheme seems to correspond to the proto-Bantu connective [associative], which has L tone in singular Classes 1 and 9 (containing most of the animates...) and H tone in all other classes. (Voorhoeve 1971:52)

The diachronic account in (24) gives a good explanation as to why Medumba has all these floating tones. We can, of course, be concerned with the issue of how to represent the effects of floating tones synchronically. I give a few possibilities in (25).

³(22c,d) show only 16 combinations, because initial class 1a takes only a L associative tone.

(25) Some of the possible ways of representing the tone of /yu/ ‘thing’

a. diacritic	b. floating L	c. feature geometry	(etc.)
yu H [+d]	yu H L	yu H L	

These issues were very much a preoccupation of mine and of the members of the Grassfields Bantu Working Group in the 1970s, when Maurice Tadadjeu and I surveyed the tonal developments in 30 different dialects or languages (Hyman & Tadadjeu 1976).

Since one might argue that the true explanation for the current properties of Medumba tone is a diachronic one, I want now to cross the Nigerian border and consider a second example where the synchronic issues are perhaps more pressing. In his thesis, Alexander Iwara provides a comprehensive description of a rather unusual noun compounding process which characterizes Lokaa (est. 120,000 speakers) and a number of closely related Upper Cross languages.

As seen in (26), most nouns in Lokaa begin with a noun class prefix:

(26) Some of the noun class prefixes in Lokaa [lòkǎǎ] (Iwara 1982)

a. o-/ɔ-	ò.dóm	‘man’	ò.bàsè	‘rain’
b. e-/ɛ-	è.tì	‘road’	è.tó	‘house’
c. ke-/kɛ-	kè.tí	‘tree’	kè.kpáŋ	‘cover’
d. le-/lɛ-	lè.tú	‘head’	lè.mà	‘hole’
e. lo-/lɔ-	lò.wí	‘water’	lò.sè	‘earth’
	etc.			

There is a full noun class agreement system, as well as vowel harmony (cf. Akinlabi & Iwara 2003).

The amazing fact about Lokaa noun compounding is seen in (27).

(27) Pref ₁ .Stem ₁ + Pref ₂ .Stem ₁		Pref ₂ .Stem ₁ -Stem ₂	
a. ò.bàsè	+	kè.kpáŋ	kè.bàsè-kpáŋ ‘sky’ (rain-cover)
rain		cover	
b. è.tì	+	lè.mà	lè.tì-mà ‘door, doorway’ (road-hole)
road		hole	
c. ò.ká:lán	+	è.tó	è.ká:lán-tó ‘house like a white man’s’
white man		house	

In (27a), the input noun ‘rain’ has the prefix ò-, while the input noun ‘cover’ has the prefix kè-: ò.bàsè, kè.kpáŋ. However, the compound in the output, kè.bàsè-kpáŋ has has

the prefix *kè-* followed directly by the two noun stems. The examples in (27b and c) are similar: the prefix of the second noun is realized at the beginning of the noun compound.

As summarized in (28), it appears superficially that Lokaa compounding is accomplished by moving the prefix from the second (or head) noun onto the first noun, whose prefix is thereby replaced.

(28) Lokaa compounding (first approximation)

Move the prefix from the second (or head) noun onto the first (non-head) noun, whose prefix is thereby replaced

Result: prefix₂ + stem₁ + stem₂

The resulting string therefore consists of the second prefix followed by the first stem, followed by the second stem.

Perhaps we should refer to these as “floating prefixes”! As in Voorhoeve’s tonal analysis, Iwara studies all of the logical combinations and potentially interacting factors to show that the analysis in (28) cannot be right—that is, it is not exactly the case that prefixes are moving around. The evidence comes from prefixless nouns such as those in (29).

(29) Prefixless nouns in Lokaa

- a. *jèn* ‘eye’ b. *ỳ̀l* ‘snake’ c. *ẁ̀l* ‘body’

As Bantuists and Benue-Congoists will recognize, these are nouns whose prefix is likely to have fused with the stem, which may have been vowel-initial. The following question naturally arises: What will happen if the second noun of a compound is prefixless? One possibility is that nothing will happen, because the nouns in (29) don’t have a prefix to send over to the first noun. As Iwara shows in the examples in (30), this is not the case.

(30) A prefixless second noun assigns an unpredictable prefix to the first noun

- | | | | | | |
|----|--------------|---|-------------|--------------------|---------------------------------------|
| a. | <i>è.wì</i> | + | <i>jèn</i> | <i>lè.wì-jèn</i> | ‘sun’ (sunshine-eye) |
| | sunshine | | eye | | |
| b. | <i>lò.wí</i> | + | <i>ỳ̀l</i> | <i>ó.wí-ỳ̀l</i> | ‘river-snake’ |
| | water | | snake | | |
| c. | <i>ò.dám</i> | + | <i>ẁ̀l</i> | <i>lì.dám.ẁ̀l</i> | ‘physique like a man’s’
(man-body) |
| | man | | body | | |

As seen, a prefixless second noun may assign one of several prefixes to the first noun. In (30a), for instance, the noun *jèn* ‘eye’ seems to hold back a prefix *le-* which replaces the prefix *è-* on *è.wì* ‘sunshine’. In (30b) *ỳ̀l* ‘snake’ assigns an *o-* prefix, while in (30c), *ẁ̀l* ‘body’ assigns a *li-* prefix. So, we have to modify the analysis as in (31).

(31) Lokaa compounding (second approximation)

The second (head) noun assigns a noun class feature to the first noun, which is then spelled out as its prefix.

Result: prefix₂ + stem₁ + stem₂

Given this discovery, we realize that we have considered only two of the four logical possibilities, those in (32a,b).

(32) Logical possibilities in noun+noun compounding

- a. pref-stem + pref-stem
- b. pref-stem + Ø-stem
- c. Ø-stem + pref-stem
- d. Ø-stem + Ø-stem

Iwara thus also considers the combinations in (32c,d). As seen in (33), whenever the first noun is prefixless, it remains so. In (33a) there is thus no change, while in (33b), the second noun loses its prefix.

(33) When the first noun is prefixless, it remains so

- | | | | | | |
|----|-------|---|---------|----------|----------------|
| a. | yəl | + | wəl | yəl-wəl | 'snake's body' |
| | snake | | body | | |
| b. | jèn | + | kè.kpán | jèn-kpán | 'eye-lid' |
| | eye | | cover | | |

So what we need to add is that prefixless nouns are morphologically saturated so that they cannot accept a prefix under any circumstance. They are, in effect, words, not just stems. This leads then to the possible interpretation of Lokaa as having noun-stem compounding, as in (34).

(34) Lokaa noun-stem compounding (almost-final version)

In noun-stem compounds, the noun class feature carried by the first (non-head) noun stem is not realized; the noun class feature carried by the second (head) noun stem is spelled out as a prefix on the compounded stem.

i.e. stem₁ + stem₂ stem₁ + stem₂ prefix₂ + stem₁ + stem₂
 [1] [2] Ø [2]

Other alternatives involving feature percolation also come to mind.

But we are still not finished. Not content with having exhausted the logical combinations in (32), Iwara then gives the generalizations concerning the tone that is found on the one prefix in compounds. First, in (35), he states that:

(35) "The tone of a compound noun prefix is low, when all the component nouns have prefixes marked by low tone." (Iwara 1982:118)

- | | | | | | |
|----|--------|---|-------|----------|-----------------|
| a. | è.fəl | + | lè.mà | è.fəl-mà | 'rabbit's hole' |
| | rabbit | | hole | | |

- b. yà.tù + lè.kòlò lè.tù-kòlò ‘door, doorway’ (road-hole)
road hole

More interestingly, he goes on to state in (36):

(36) “But the tone of a compound noun prefix is high, when at least one of the component nouns has a prefix marked by high tone.” (Iwara 1982:119)

- a. é.fém + lè.gà lé.fém-gà ‘crocodile’s tooth’
crocodile tooth
- b. ò.nò:tàm + é.kò é.nò:tàm-kò ‘old cloth’
old person cloth

In (36a), the H tone is contributed by the prefix *é-* of the first noun, while in (36b) the H tone is contributed by the prefix *é-* of the second noun. This is somewhat paradoxical, since the shape of the surfacing prefix is supposed to be entirely predictable from the second (head) noun. This raises the question of whether the H tone is an underlying property of a prefix or is an initial floating H stem tone, something which has been a gap in the tonal literature.⁴

Why describe African languages? In case we need a reminder of how interesting African languages are, this is fascinating material. Lokaa compounding makes synchronic sense, but I think it is quite unusual. Although this merits further exploration, it seems to me that the preferred compound structures are those given in (37a,b), where the underlining indicates the head of the compound, which both governs and is adjacent to its prefix or suffix:

- (37) Stem-compounding structures (underlined = head)
- | | | | |
|---|-----|--|----------------|
| preferred | | dispreferred | |
| a. pref ₁ - <u>stem₁</u> -stem ₂ | vs. | c. pref ₂ -stem ₁ - <u>stem₂</u> | = Upper Cross! |
| b. stem ₁ - <u>stem₂</u> -suffix ₂ | | d. <u>stem₁</u> -stem ₂ -suffix ₁ | |

The structures in (37c,d) are less fortunate in that the affix is not adjacent to its governing head. As seen, Upper Cross decides to adopt (37c) anyway. We could model this as a very nice OT conflict between head-affix adjacency vs. the need to express one or both prefixes. My impression is that head-final compounding is rare in prefixing Benue-Congo languages—perhaps a strategy adopted to avoid the conflict altogether. Be this as it may, unlike the case of floating tones in Medumba, it’s not clear that the diachronic dimension has much more to add to what may be going on here, unless compounding bears on the syntactic changes that characterize Upper Cross in general.

⁴Further examples in Akinlabi & Iwara (2003) indicate that the tonal facts in (36) are quite robust in the language. I have suggested that the segmental shape of the compound prefix can be determined by feature percolation from the second (or last) head noun. Since either noun in a compound can contribute a H tone to the surfacing prefix, it would seem that the H of the head noun must also percolate, as if it were the realization of a root-level feature [+H]. However, since the first noun can also contribute a H, we must assume that its [+H] may also percolate. If L tone prefixes lack a feature specification (they are presumably more frequent), there will be no conflict: the [+H] specification will always be realized as a H prefix, no matter what its source. It would be interesting in this context to systematically investigate the realization of tone in tripartite and other “layered” compounds.

So, for my third illustration, I will remain in Upper Cross and discuss some quite unexpected syntactic facts from Leggbó, a language with ca. 60,000 speakers, which we studied in my field methods course in Berkeley two years ago with the able assistance of Imelda Udoh.

In Leggbó, Lokaa, and closely related languages there is a word order difference between affirmative and negative clauses. As seen in (38), the word order is SVO in the affirmative, but as seen in (39), the word order is SOV in the negative:

(38) Affirmative word order in Leggbó: SVO

- a. bàdum sé ba dzi lídzil 'the men ate food'
 men the 3pl ate food
- b. *bàdum sé lídzil ba dzi
 men the food they ate

(39) Negative word order in Leggbó: SOV

- a. bàdum sé lídzil bè àà dzi 'the men didn't eat food'
 men the food 3pl neg eat
- b. *bàdum sé bè àà dzi lídzil
 men the 3pl neg eat food

The word order difference is only the beginning, however. As also seen, I've underlined the subject-verb agreement, which is obligatory in both the affirmative and the negative, although the markers are different (*ba* vs. *bè*). While the affirmative agreement marker *ba* must occur immediately before the verb in (38a), (40) shows that its position is variable in negative clauses:

(40) Variable placement of subject agreement (*bè*) in negative clauses

- a. bàdum sé lídzil bè àà dzi 'the men didn't eat food'
 men the food 3pl neg eat
- b. bàdum sé bè lídzil àà dzi
 men the 3pl food neg eat
- c. bàdum sé bè lídzil bè àà dzi
 men the 3pl food they neg eat
- d. *bàdum sé lídzil àà dzi
 men the food neg eat

The agreement marker *bè* occurs before the verb in (40a), but before the object in (40b). In (40c) it occurs in both places, while (40d) shows that the utterance is ungrammatical without an agreement marker. As seen in (41), the same facts hold whether the subject noun phrase *bàdum sé* 'the men' is present or not:

(41) Same variable placement of *bè* when subject NP is not present

- a. lídzil bè àà dzi 'they didn't eat food'
- b. bè lídzil àà dzi
- c. bè lídzil bè àà dzi

When we add more elements, the plot thickens. (42) shows the one possible outcome in the affirmative when there are two objects:

(42) Affirmative double object construction: SVOO

bàdum sé ba ni wàé sé lídzil 'the men gave the child food'
 men the 3pl gave child the food

As in the case of Medumba tone and Lokaa noun compounding, in (43) we consider all of the logical possibilities for placing the subject agreement marker *bè* in the corresponding negative:

(43) Eight logical possibilities for placing *bè* in the negative double object construction

	men	the	3pl	child	the	3pl	food	3pl	neg	eat		
a.	bàdum	sé	—	wàé	sé	—	lídzil	<u>bè</u>	aà	dzi	(one <i>bè</i>)	'the men didn't give the child food'
	bàdum	sé	—	wàé	sé	<u>bè</u>	lídzil	—	aà	dzi		
	bàdum	sé	<u>bè</u>	wàé	sé	—	lídzil	—	aà	dzi		
b.	bàdum	sé	—	wàé	sé	<u>bè</u>	lídzil	<u>bè</u>	aà	dzi	(two <i>bè</i> 's)	
	bàdum	sé	<u>bè</u>	wàé	sé	<u>bè</u>	lídzil	—	aà	dzi		
	bàdum	sé	<u>bè</u>	wàé	sé	—	lídzil	<u>bè</u>	aà	dzi		
c.	?bàdum	sé	<u>bè</u>	wàé	sé	<u>bè</u>	lídzil	<u>bè</u>	aà	dzi	(three <i>bè</i> 's)	
d.	*bàdum	sé	—	wàé	sé	—	lídzil	—	aà	dzi	(no <i>bè</i>)	

(43a) shows that a single *bè* suffices in any of three positions: before the verb, between the two objects, or between the subject and the first object. (43b) shows that any combination of two *bè*'s is acceptable. The sentence in (43c), with three occurrences of the agreement marker, is probably OK, although it feels like there's a little too much *bè* in it. Finally, (43d) agains shows that a sentence with no subject agreement is not grammatical.

The word order in the affirmative is quite fixed. Although I don't show it here, there is much freer word order among the preposed elements in the negative, such that Good (2003) proposes that that these elements are adjuncts not arguments. Be that as it may, the potential relevance of these facts to Niger-Congo studies is not to be missed. Leggbó and related languages are surrounded by SVO languages: Igbo to the West, Ibibio to the South, Bantoid to the North and East. The SOV word order is definitely an innovation. I will show this in two ways.

First, as seen in (44), the verb *dzè* 'finish' can be serialized with the two indicated meanings in the affirmative:

(44) *dzè* 'finish' in the affirmative = 'already', 'finish'

ba dzi lídzil (b)à dzè 'they have already eaten food, 'they have finished eating food'
 3pl ate food 3pl finish

In the corresponding negative in (45), *dzè* has clearly been grammaticalized and does not take subject agreement:

(45) dzè ‘finish’ in the negative = ‘not yet’, ‘never’, ‘negative habitual’

- a. — lídzil bè aà dzi dzè ‘they haven’t eaten food yet, they never eat/ate food’
 b. bè lídzil — aà dzi dzè
 c. bè lídzil bè aà dzi dzè
 3pl food 3pl neg eat finish = ‘yet’ (*...dzi bè dzè)

As shown, bè can appear once or twice, but not between ‘eat’ and dzè. What’s interesting is that dzè can also appear before the verb, i.e. it can be preposed as in (46).

(46) Preposed dzè ‘finish’ in negative

- a. NEG: with subject marker in one position
 — lídzil — dzè bè aà dzi ‘they haven’t eaten food yet’
 — lídzil bè dzè — aà dzi
bè lídzil — dzè — aà dzi
- b. NEG: with subject marker in two positions
 — lídzil bè dzè bè aà dzi
bè lídzil bè dzè — aà dzi
bè lídzil — dzè bè aà dzi
- c. NEG: with subject marker in three positions
bè lídzil bè dzè bè aà dzi
 3pl food 3pl finish 3pl neg eat

In fact, as seen in (47), there can even be more than one dzè, just like there can be more than one bè:

(47) dzè can appear in two or three positions!

- a. NEG: with dzè in all three positions (V2, middle, post-verbal)
 (bè) dzè lídzil (bè) dzè (bè) aà dzi dzè
 3pl yet food 3pl yet 3pl neg eat yet
- b. NEG: with dzè in two positions
 (bè) — lídzil (bè) dzè (bè) aà dzi dzè
 (bè) dzè lídzil (bè) dzè (bè) aà dzi —
 (bè) dzè lídzil (bè) — (bè) aà dzi dzè

Clearly Leggbó has gone off in a certain direction where each preposed constituent is a potential agreement domain, which must necessarily have been an innovation.

Finally, let’s see what happens to objects with serial verbs. (48) shows an affirmative serial verb construction involving an instrument:

(48) Serial verb in affirmative, using kaa ‘carry’ to express an instrument

- ba kaa izbòòm (b)à vili ètèèn ‘they cut meat with a knife’
 3pl carry knife 3pl cut meat

governments or the public. But if we hope to have a healthy field, we must also ask the question ourselves, as Africanist linguists.

I started this talk by discussing some of the concerns I have about the description of African languages in a rapidly changing field. It is encouraging to see how good Voorhoeve's and Iwara's analyses look after 20 to 30 years. The question is whether it is enough to continue doing as we have been, or whether we need to bring more attention to the important and exciting phenomena that we discover in describing African languages.

I want to suggest a few ways to do just this:

First, let's all go out and describe. Forget about theory vs. syntax: Whether in the field or with linguistic consultants at home, we should just do good work!

Second, let's try to get every linguistics program to offer a course in field methods, in which unempowered languages can be investigated and described. This would be important not only for the information one could contribute concerning these languages, but a field methods course is in fact an essential part of the training of new linguists—who, in turn, may actually go out into the field.

Third, let's try to get away from those kinds of studies that I have called exercises, i.e. those which are likely to become quickly dated, and often contain relatively little descriptive information about a language. I know that we can't all write Paul Newman's Hausa reference grammar or Hulstaert's *Grammaire du Lómóngò*.⁶ But, I personally think it should be possible in any linguistics department in the country to write a properly rich descriptive grammar of a language as a PhD dissertation, for example, like Jay Nash's (1992) 1400-page grammar of Ruund, which he, I suspect proudly, submitted at the University of Illinois.⁷

Finally, I hope there will continue to be departments of linguistics where one can be both educated and trained to do African linguistics. When I went to graduate school, in addition to two years of language classes in Igbo and Hausa, and my informant work on Bamileke, I took a course on the Structure of African Languages with William Welmers, on African language classification with Charles Kraft, on the Structure of Bantu with Talmy Givón, and field methods courses on Nupe and Luganda. It was through these experiences (and much prodding from Erhard Voeltz) that I came to be familiar with Westermann and Meinhof, the Guthrie-Greenberg controversy, Bernd Heine's *Togorestsprachen*, and reconstructing proto African languages. I think that the most crucial part of training as an Africanist linguist is in historical linguistics.

⁶I was reminded at the conference that there are three recent reference grammars of Hausa. See also Wolff (1993) and Jaggar (2002).

⁷I use the word "possible" in two senses. First, one should be allowed to write a descriptive grammar as a descriptive grammar as a PhD dissertation. Second, one's career should not be negatively affected by so doing. Although I tried to avoid the issue in the talk, I was asked in the question period about the perception many have that if you do a descriptive dissertation, you won't get a job—at least not within a top linguistics department in the United States. It is true that one won't get a theoretical job unless one does theoretical work. But there are a lot of positions in departments whose wide-ranging needs would best be met by an enterprising descriptive linguist who has a lot to offer in different areas. Again, my recommendation would be "do good work!", which hopefully will be recognized. Needless to say, these are issues worthy of a separate discussion

Unfortunately, I don't think many departments have been able to maintain this much of an Africanist presence, either in personnel or student interest.⁸

Of course we would all love it if every department of linguistics had at least one Africanist. That may seem like too much to ask for—but isn't 2,000 out of 6,000 languages a big enough constituency for one position? So, to conclude, here's my political campaign slogan: One language, one vote!

References

- Akinlabi, Akinbiyi & Alexander Iwara. 2003. Transparency and opacity in Lokaa vowel harmony. Paper presented at WOCAL4/ACAL34, Rutgers, June 19, 2003.
- Akinlabi, Akinbiyi & Mark Liberman. 2000. The tonal phonology of Yoruba clitics. In B. Gerlach and J. Grizjenhout (eds). *Clitics in phonology, morphology and syntax*, 31-62. Amsterdam: Benjamins.
- Childs, Tucker. 2003. *An introduction to African languages*. Amsterdam: Benjamins (in press).
- Claudi, Ulrike. 1993. *Die Stellung von Verb und Objekt in Niger-Kongo-Sprachen: Ein Beitrag zur Rekonstruktion historischer Syntax*. Köln: Institut für Afrikanistik, Universität zu Köln.
- Gensler, Orin. 1997. Grammaticalization, typology, and Niger-Congo word order: progress on a still-unsolved problem. *Journal of African Languages & Linguistics* 18.57-93.
- Givón, Talmy. 1975. Serial verbs and syntactic change: Niger-Congo. In Charles N. Li (ed.), *Word order and word order change*, 47-112. Austin: University of Texas Press.
- Good, Jeff. 2003. When arguments become adjuncts: Split configurationality in Leggbó. To appear in *Proceedings of Chicago Linguistic Society* 39.
- Grimes, Barbara F. 1996. *Ethnologue: languages of the world*. 13th edition (web version) Dallas, Texas: Summer Institute of Linguistics.
- Heine, Bernd. 1968. *Die Verbreitung und Gliederung der Togoestsprachen*. Berlin: Dietrich Reimer Verlag.
- Heine, Bernd. 1976. *A typology of African languages*. Berlin: Dietrich Reimer Verlag.
- Heine, Bernd. 1980. Language typology and linguistic reconstruction: the Niger-Congo case. *JALL* 2.95-112.
- Hulstaert, Gustaaf. 1961/1965/1966. *Grammaire du Lómóngó*. Part I: Phonologie, Part II: Morphologie, Part III: Syntaxe. Tervuren: Musée Royal de l'Afrique Centrale.
- Hyman, Larry M. 2001. Fieldwork as a state of mind. In Paul Newman & Martha Ratliff (eds), *Linguistic Fieldwork*, 15-33. Cambridge University Press.
- Hyman, Larry M. & Maurice Tadadjeu. Floating tones in Mbam-Nkam. In Larry M. Hyman (ed.), *Studies in Bantu tonology*, 57-111. Southern California Occasional Papers in Linguistics 3, University of Southern California.
- Iwara, Alexander Ubi. 1982. *Phonology and grammar of Lókóó: A preliminary study*. Thesis submitted for the degree of Master of Philosophy of the University of London. Department of Africa, School of Oriental and African Studies.
- Jaggar, Philip J. 2001. *Hausa*. Amsterdam: Benjamins.
- Nash, Jay. 1992. *Aspects of Ruwund grammar*. PhD dissertation, University of Illinois at Urbana-Champaign.
- Newman, Paul. 2000. *The Hausa language: an encyclopedic reference grammar*. New Haven: Yale University Press.
- Noonan, Michael. 2002. Grammar writing for a grammar-reading audience. Paper presented at the Symposium on Grammar Writing, SIL, Dallas, October 19, 2002.
- Voorhoeve, Jan. 1971. Tonology of the Bamileke noun. *Journal of African Languages* 10.44-53.
- Watkins, Mark Hanna. 1937. *A grammar of Chichewa, a Bantu language of British Central Africa*. Supplement 24 to *Language*.
- Williamson, Kay. 1986. Niger-Congo: SVO or SOV. *Journal of West African Languages* 16.5-14.

⁸Unlike graduate student bodies today, many of my classmates in the late 1960s were Peace Corps returnees—such that I was one of the few taking African linguistics courses who hadn't been to Africa yet.

Wolff, Ekkehard. 1993. *Referenzgrammatik des Hausa: zur Begleitung des Fremdsprachenunterrichts und zur Einführung in das Selbststudium*. Munster: Lit.