bls Berkeley Linguistics Society

Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: Special Session on Historical Issues in African Linguistics (1994)

PROCEEDINGS OF THE TWENTIETH ANNUAL MEETING

OF THE

BERKELEY LINGUISTICS SOCIETY

February 18-21, 1994

SPECIAL SESSION

ON

HISTORICAL ISSUES

IN

AFRICAN LINGUISTICS

PROCEEDINGS OF THE TWENTIETH ANNUAL MEETING

OF THE

BERKELEY LINGUISTICS SOCIETY

February 18-21, 1994

SPECIAL SESSION

ON

HISTORICAL ISSUES

IN

AFRICAN LINGUISTICS

edited by

Kevin E. Moore

David A. Peterson

Comfort Wentum

TABLE OF CONTENTS

SPECIAL SESSION

TABLE OF CONTENTS
FOR GENERAL SESSION
TO BE FOUND IN BLS 20

GENERAL SESSION

PREFACE

We herewith present the second Africanist volume in the series of Special Session proceedings devoted to Berkeley's areal specializations.

There are a number of people to be thanked for their assistance at various stages. Above all, many thanks are due to Jeri Moxley for her participation in running the Special Session and getting it set up in the months beforehand. Larry Hyman was also of considerable help in numerous respects at all stages. Others to whom we express appreciation include Joyce Mathangwane, Josephat Rugemalira, Cheryl Zoll. Thanks also to the organizers of the General Session, Andy Dolbey, Susanne Gahl, and Chris Johnson, and to the participants at the Special Session.

Kevin E. Moore
David A. Peterson
Comfort Wentum

On Reconstructing the Syntagm S-Aux-O-V-Other to Proto-Niger-Congo

Orin D. Gensler
University of California, Berkeley

[This paper is dedicated, with much respect and much affection, to Charles Fillmore, part-time Africanist — in whose field methods class on Bakweri I had my first linguistic exposure to a Niger-Congo language.]

1. Background and Overview

'Niger-Congo: SVO or SOV?' This question, the title of a 1986 article by Kay Williamson, epitomizes a twenty-year debate over the clause-level word order of Proto-Niger-Congo (PNC). Niger-Congo languages by the hundreds are overwhelmingly Subject-Verb-Object (SVO) in dominant word order; only Ijo shows clearcut OV clause-level syntax, while the Mande languages are technically 'SOV' but in fact place everything except the object after the verb. Hence it might seem that Proto-Niger-Congo should be reconstructed, unproblematically, as SVO. Contrary to this position, Talmy Givón and others (Givón 1975, 1979, Hyman 1975, and the Williamson article just cited) have argued vigorously for an original SOV order as in Ijo, drawing on 'minor' word order patterns throughout the family and making heavy use of the 'morphology as frozen syntax' (MAFS) principle. On the other hand, Bernd Heine and others associated with him (Heine 1976, 1980; Heine and Reh 1984; Claudi 1988, 1993), appealing to standard comparativist methodology and invoking the overall word order profile of Niger-Congo (including NP-level syntagms), have argued for SVO as the original word order in the family. The debate becomes an explicit head-to-head confrontation in Givón 1979 and Heine's 1980 rejoinder in successive numbers of the same journal.

The question/title quoted above sums up not only the debate but the presuppositions shared by both sides in the debate — presuppositions highly characteristic of the era when the question was first formulated. Word order studies in the '70s were dominated by the VO/OV 'seesaw theory' associated with Lehmann and Vennemann, whereby long-term change in word order was seen as a slow pendulum swing toward or away from ideal VO (Head-Modifier) or OV (Modifier-Head) macro-types. In this tradition clause-level elements other than S, V, and O — that is, adverbials of any sort — are given at best cursory attention. Specifying the position of the object (more specifically, the direct object) vis-a-vis the verb is implicitly taken as a shorthand way of indicating the position of any verbal argument or adjunct vis-a-vis the verb. And indeed, in the great majority of the world's languages this is the case.

Three assumptions appear to be taken for granted in both Givón's and Heine's approach to Proto-Niger-Congo word order:

a) The 'interesting' question to ask is whether Proto-Niger-Congo was SVO or SOV;

b) This question ought properly to have a clear-cut, yes/no answer;

c) Given (a) and (b), the right approach to the answer must necessarily focus on articulating paths and mechanisms of *change*, whereby modern-day SVO features can be shown to derive from earlier proto-SOV, or vice versa.

This paper comes at the problem of Proto-Niger-Congo word order in a very different way, one which ignores or takes issue with all of these assumptions. The question it asks — the question I consider to be the really interesting one — bypasses the usual SVO/SOV dichotomy, and indeed ignores the matter of *overall* clause-level word order type almost entirely; rather, it focuses on one particular and very distinctive word order subconfiguration of Niger-Congo: S-Aux-O-V-Other. Far from seeing the abstract choice between SVO and SOV as a clearcut yes/no dichotomy, it observes that S-Aux-O-V-Other (and more generally S-O-V-Other) represents an intermediate type that is neither SVO nor SOV in any straightforward sense (though it can be and has been 'argued into' the status of SVO or of SOV). And its focus is not on change (regrammaticalization) but on continuity. It will be argued that the given configuration *as a configuration* — a template — should be reconstructed all the way back to Proto-Niger-Congo as a pattern that was always there, constantly emerging and re-emerging with new morphemic material as far back as we can envision in the linguistic prehistory of the family. This in turn has significant implications for how we see the overall clause-level word order type of PNC.

2. Terminology

In this paper I will be consistently using the terms Aux (= Auxiliary) and MainVerb in a particular way. I take the MainVerb in a clause to be that element which belongs to an open paradigmatic class and which conveys the lexical-semantic content of the verbal action or state or process in the given clause. An Aux, by contrast, is taken as an element which is syntagmatically separate from the MainVerb (thus not an affix), which belongs to a small, closed paradigmatic class, and which conveys grammatical notions such as tense, aspect, mood, and/or negation. In the English clause 'I have been thinking', the MainVerb is 'think' and the Auxes are 'have', 'be'. These definitions are deliberately neutral as to the categorial assignment of Aux and MainVerb, in order to foster comparability across the whole family. Thus the Aux may be a particle, a special subclass of verb (including a serial verb), or its own distinct part of speech; the MainVerb can perfectly well pattern morphosyntactically as a (verbal) noun or a participle, and often does. Normally the MainVerb is obligatory; in some languages the Aux is as well.

The term 'SOV' will be understood in this paper as representing 'canonical SOV', that is, a word order where the verb (MainVerb) basically appears in *final* position in the clause. This terminological stipulation reflects the fact that, crosslinguistically, languages which position O before V overwhelmingly position the verb clause-finally; even in languages described as 'flexible SOV', the unmarked position of the verb is still clause-final. The corollary is that, understood in these terms, the structure S-Aux-O-V-Other (taken strictly as a surface string and ignoring internal constituency) is neither SOV nor SVO. Unlike canonical SOV but like SVO, the MainVerb is not clause-final but clause-medial; unlike SVO but like SOV, the Object precedes the Verb.

'Other' covers everything except Subj and (Direct) Obj, i.e. Adverbials of all sorts. In languages where indirect objects are expressed as adpositional phrases ('to John'), it may cover IndObj as well. On the other hand, 'O' (= Object) may include IndObj and other secondary objects in addition to DObj, depending on the language.

3. The Phenomenon

In many branches of Niger-Congo, some languages — and, in at least one branch (Kru), apparently all languages (Marchese 1986:218) — show a characteristic constructional pattern whereby sentences having an Aux appear in a special word order: Subject - Aux - Object(s) - MainVerb - Other.   (In many of the languages this contrasts sharply with Aux-less clauses, where the order is straightforwardly SVO.) In some languages, only one Obj may precede the MainVerb; other languages allow multiple preverbal Objects. All previous works on Niger-Congo word order diachrony have of course acknowledged this pervasive construction, but only Lyn Marchese's excellent treatment of the syntagm in Kru has made it the focus of attention. In other studies the construction appears in the penumbra of something else — not as a phenomenon of interest in its own right, but in a supporting role, as part of the arsenal mustered for or against original SVO or SOV order.

Below I present, in condensed form, representative data displaying the construction as it appears in six branches of Niger-Congo. The construction is a dominant feature of Kru, Senufo, and Mande. It appears elsewhere in other branches, but is conspicuously rare in Bantu — a point we will return to in sec. 8.

A. Mande

The basic order is   Subj (Aux) DObj V Other; this order appears to hold rigidly throughout the family, regardless of whether or not an Aux is present (cf. Heine and Reh 1984:198–202).

The situation is clearest in Mandinka (Creissels 1983). Here the Aux element (Creissels' 'prédicatif') is an obligatory component of verbal clauses (1983:20), and the standard verbal syntagm (26–27) is:

Subj Aux DObj V Other

The postverbal 'Other' includes IObj, which is realized as an adpositional phrase and thus counts as an oblique (Creissels' 'circonstant'); only the direct object is preverbal. Examples:

1) [mòolu   ye   kinoo dii]   n na
   people PAST food give   me to
   'The people gave me food'   (134)

2) [i   kana   wò fo]   à ye
   you PROHIB that say   him to
   'You mustn't say that to him'   (125)

In Mende, the basic word order is S (Aux) O V-Tns/Asp Other (Aginsky 1935:99, Innes 1962:127ff.); the sequence [O-V] forms a tight constituent that is identical structurally to one subtype of [Gen-N] construction (Innes 1963:54ff.). The indirect object counts as 'Other' and is postverbal. Tense/aspect distinctions (Aginsky 1935:31ff., Innes 1962:122ff.) are basically coded by suffixes on the verb, and most tense/aspect forms do not involve an Aux; only a handful of elements may occur between S and O, including Neg and a few aspectual/adverbial particles (Innes 1963:116, 122, 134–35), so that 'Aux' appears to be a relatively impoverished category.

Significantly, however, subject pronouns occur in a number of distinct series, depending on the choice of verbal tense/aspect (Migeod 1908:69–70, Aginsky 1935:20–22, Innes 1962:31ff.). Some of these series are analyzable as the phonological fusion of the 'basic' pronominal series (the *ngi*-series) with a following vocalic element (Innes 1963:126–27, 132ff.); and this vocalic element, which conveys tense/aspect information, plausibly represents the residue of earlier Aux elements. Moreover, subject pronouns are of extremely frequent occurrence because full-NP subjects (especially plural NPs) are commonly doubled by a following subject pronoun (Migeod 1908:70, Aginsky 1935:23–24, Innes 1963:135–36 and passim). Hence the Aux residue, as an integral part of the Subj pronoun, is likewise very common. Examples (Innes 1963):

3) Kpana ii   nikɛ-i   waa-ma ha   'Kpana is not killing the cow today'
         Neg cow-Def kill-Sfx today   (explicit Aux *ii*) (118)

4) ngi mbɛ-i   lɔ   mɛ-ni   'I ate the rice' (121)
   I   rice-Def Emph eat-Sfx   (no Aux)

5) nga mba mɛ lɔ   'I eat rice (habitually)' (131)
   I   rice eat Emph   (frozen Aux: *nga* = *ngi* + *a*)

Essentially the same overall pattern appears to hold for Kpelle (Welmers 1973:395ff.).

B. Kru

Marchese devotes a major portion of her excellent study of tense/aspect in Kru to the construction S Aux O V (1986, especially chapter 5). Throughout the Kru family, there is a sharp distinction between verb constructions with and without Aux (1986:21, 24, 225):

Subj Aux IObj DObj V Other
Subj V   IObj DObj   Other

(Not all Aux elements show this patterning; Marchese separates out a subclass she terms 'sentence-second particles', with word order 'S Aux V O' (24).) Examples from Wobé (Western Kru):

6) [ɔ   se   kei ko kpa]   dè   dɔɔ
   he NEG Kei rice bring LOC market
   'He didn't bring rice to Kei at the market' (243)

7) [ɔ   se   (e) dɔɔ   mu]
   he NEG to market go
   'He didn't go to the market' (91)

Kru languages show variability regarding the positioning of adverbials when an Aux is present: in some languages the adverbial is preverbal, in some postverbal (219, 225), and different classes of adverbials can pattern in different ways. However, one adverbial subtype shows consistent behavior throughout the family. Locative complements of verbs of motion (see ex. [7]), which are subcategorized for by their governing verb, always come between Aux and V — just like objects. This is no accident, Marchese argues, for such a verb is properly to be taken as *transitive*; the locative complement is its Object (89–92). With other types of adverbials — non-complement locative, temporal, manner, etc. (219ff.) — there is a rough dividing line between Eastern and Western Kru: often Eastern Kru allows preverbal position, while Western Kru does not (219).

C. Kisi (Atlantic)

The same structural opposition evident in Kru exists in the Atlantic language Kisi (Childs 1988:29, 32):

Subj Aux IObj DObj V Other
Subj V   IObj DObj   Other

No explicit statement is made in the grammar about the ordering of IObj and DObj, but the above formulas fit the examples. Thus:

8) í bɛ́í ndú
   I hurt him
   'I hurt him' (29)

9) í có ndú bɛ̀ì
   I FUT him hurt
   'I will hurt him' (29)

10) ò có   ndú kóná   dóónɔ̆ng   (preverbal IObj, DObj)
    he FUT him message pour.forth
    'He will relate the message to him' (139)

11) ò cìí   yɔ̀mndé lòmɔ̀ɔ̀   múɛ́ɛ́ng   (postverbal non-Arg)
    he finish wood   burn   IDPH
    'He finished burning the wood completely' (143)

D. Tunen, Ewondo (Bantu)

As indicated above, S-Aux-O-V-Other is highly unusual in the syntax of Bantu languages. Interestingly, the one or two (narrow) Bantu languages where it does occur are spoken in Cameroon, the area of the Proto-Bantu homeland. These basically isolating languages lack the characteristic agglutinative structure typical of

most Bantu languages. Clearest of all — though apparently unique in Bantu (Dugast 1971:6) — is Tunen, where all objects (noun or pronoun, direct or indirect) occur between Aux and Verb while adverbials are postverbal (6, 171, 309):

> 12) á ndò miaŋó menyàma húlən   'he brought me the meat' (171)
>   he AUX me   meat   bring

In Ewondo, pronoun objects (but not noun objects) come between Aux and main verb (Redden 1980:126, 166–67):

> 13) a- kad   mə dzɔ vɔ́   'he usually gives it to me' (167)
>   he-do.usually me it give

In both languages, indirect objects precede direct (1971:309, 1980:166).

E. Ewe (Kwa)

Heine (1980:104) provides the following examples from Ewe. The language has basic SVO order, but in progressive and ingressive aspects the object is preverbal (cf. also Heine and Reh 1984:188–90):

> 14) éle   nú-ɖu ḿ   15) mele   é-dí   ḿ
>   he-is thing-eat PROG   I-am   him-want PROG
>   'He is eating'   'I am looking for him'

Note the fusion of Aux onto the Subj (as in Mende).

F. Senufo (Gur)

The syntagm S-Aux-O-V-Other is standard in the Senufo subgroup of the Gur family; the indirect object is post-verbal (Carlson 1991:201, Claudi 1993:112; also Givón 1975:60). Thus:

> 16) u   a   kù kàn mìì-á   's/he gave it to me' (Supyire)
>   s/he PERF it give me-to   (Carlson 1991:217)

4. S-Aux-O-V-Other as a Rare Type Globally

As the preceding section demonstrates amply, the syntagm S-Aux-O-V-Other as a clearly profiled, salient word-order type is a fixture of many (though not all) branches of Niger-Congo. In striking contrast to this is the rarity of the type worldwide. In a reasonably well-balanced global sample of over 75 languages (created for other purposes), I found no cases of the syntagm outside Africa; elsewhere in Africa, it exists only (?) in some dialects of the non-Niger-Congo language Songhai (Carlson 1991:201), a neighbor of Mande. At first glance the Mandarin Chinese *bǎ* construction,   Subj *bǎ* DObj Verb   (Li & Thompson 1974; 1981:463ff.), might appear to be an example of the configuration; but whatever the precise function of *bǎ* (apparently a combination of case-marker and indica-

tor of certain kinds of discourse salience, 1981:482ff.) it is certainly not a tense/aspect/mood 'Aux', nor is the element 'Other' constrained to postverbal position (1981:19ff.).

The ordering S-O-V-Other (without Aux) would appear to be almost as rare. Manner adverbs and especially adpositional phrases — two of the major constituent types subsumed under the rubric 'Other' — tend very strongly to appear on the same side of the verb as the Object (Dryer 1992:92–93), in line with the remarks made above (sec. 1) to the effect that the vast majority of OV languages are actually verb-final. Two instances where S-O-V-Other is presented as the favored word order are the Muskogean language Koasati (Kimball 1991:513–14) and (in the genre of legal texts) the atypical, peripheral dialect of Akkadian spoken at Ugarit (Huehnergard 1989:220ff.). However, as the authors make clear, these are flexible-word-order languages, and S-O-V-Other is merely a preferred option — quite unlike its canonical status in families like Mande. (The New Guinea language Kobon (basically SOV) may represent another example; but the grammar is self-contradictory in this regard (Davies 1989: 49, 51).)

This global rarity establishes the syntagm S-Aux-O-V-Other as a widespread (though not universal) family-wide *quirk* of Niger-Congo — a fact which will have considerable importance in arguing for the existence of the syntagm in the protolanguage.

5. The Methodology of Syntactic Reconstruction

The concerns and the approach of the present paper are very different from those of earlier diachronic work on Niger-Congo word order, and the differences bear some discussion. A great deal of attention has been devoted in previous literature to the mechanisms and scenarios whereby the verbal structures of present-day Niger-Congo languages could have emerged by grammatical reanalysis from earlier stages of the language — grammaticization processes variously involving serial verbs, verb-noun periphrastic tenses, or adverbs (see Heine and Reh 1984:114ff., and most recently and comprehensively Claudi 1993). As applied to the genesis of S-Aux-O-V-Other, such processes show how the construction can be seen as emerging from an earlier Aux-less construction (SVO or SOV). But it is one thing to explore the genesis of a particular individual instance of the construction (involving some particular concrete Aux) from an earlier Aux-less construction, and something very different to ask how and when the construction as a whole first emerged in the given language or branch of Niger-Congo. In particular, the just-mentioned scenarios, in reconstructing Aux-less SVO or SOV syntagms at an earlier stage, say nothing about whether at that earlier stage the language might not *already* have had the syntagm S-Aux-O-V-Other alongside Aux-less SVO or SOV — with different morphs in the Aux slot, of course. The present paper focuses exclusively on this latter question, and its concerns are thus largely complementary to those of earlier studies.

This paper also differs from previous studies of Niger-Congo word order in its methodology: the orthodox comparative method, as applied to syntactic material. As in traditional phonological and morphological reconstruction, a word order reconstruction for a protolanguage should above all be comparative: it should be responsive first and foremost to the observed syntactic distribution throughout the family, and should yield the most plausible path of evolution from the proto-phenomenon to the observed facts. Such considerations as synchronic 'naturalness'

or typological consistency, though important, are distinctly secondary. A proto-language, on this view, is not an idealized schematic construct but a language like any language; particular languages routinely embody 'inconsistent' and quirky phenomena, and so may the reconstructed protolanguage — if the daughter languages warrant it. Earlier work has been based heavily on internal reconstruction and on plausible assumptions about how change 'ought' to proceed, notably the 'morphology as frozen syntax' principle (associated with Talmy Givón) and the notion that word order change is dominated by vast tidal shifts between ideal VO and OV macrotypes (associated with Lehmann and Vennemann). Both of these are valuable principles, and reflect phenomena which are indeed common and natural. But they are not laws, and should never be appealed to blindly. Crucially, pure internal reconstruction by its very nature produces neat origins for un-neat data. But origins may sometimes be just as messy as the descendant reflexes.

Earlier work has tended to view syntagms that do not conform to straight-forward VO or OV patterns as deviations from an ideal, both synchronically and diachronically, and thus to view the 'quirky' syntagm S-Aux-O-V-Other as a problem in need of explanation; the explanation would involve an appeal to some more canonical word order (SVO or SOV), either recasting S-Aux-O-V-Other as actually being an instance of one or the other 'basic' configuration, or else deriving it historically from such a source. This procedure, however, is exactly backwards from the comparative method as regularly applied in standard phonological and morphological reconstruction. When a distinctive and quirky phenomenon recurs all over a family F, the comparative method may indeed seek to explain where the phenomenon comes from, especially if suitable data is available from the superfamily encompassing F; but within the family F itself, the method's primary stance is simply to accept the phenomenon as such — and, most typically, to reconstruct it in some form to the protolanguage. A quirky phenomenon thus represents not so much a problem to be solved as rather a key to solving other problems. In particular, family-wide quirks are prized as one of the very best proofs of genetic related-ness (see e.g. Meillet 1967:41). An example is the quite peculiar cooccurrence of *t*- as marker of both 2-sg and 3-f-sg (but not 3-m-sg) in Cushitic (e.g. Afar), Berber, and all of Semitic. This quirk is not normally seen as a datum in need of explanation, but rather is accepted as one of the structural cornerstones of the Afroasiatic superfamily.

In the domain of syntax, the only secure way to ground the notion 'quirk' is by an observational appeal to typology: a quirk is a pattern that is observed to be common in the language grouping in question but rare worldwide. It is in this light that I will be approaching the syntagm S-Aux-O-V-Other, rare globally yet commonplace in Niger-Congo: as a family-wide and family-specific quirk which, precisely because of its quirkiness, merits reconstruction to the protolanguage.

Of course, there is a difference between a morphological and a purely syntactic reconstruction. Morphological reconstruction posits a particular concrete morph or set of morphs which, subject to possible phonetic and semantic change, can be traced from the protolanguage down to the attested daughter languages. 'Mixed' morphosyntactic reconstruction involves the reconstruction of both concrete morphs and a syntactic pattern. Purely syntactic reconstruction, however, posits only a pattern (or diachronic succession of patterns), which may be filled with different concrete material in different languages at different periods. It is extremely important to elucidate the mechanisms (metanalysis, grammaticization) whereby this kind of serial replacement of morphemic material occurs. Equally important, however, is the proto-pattern itself — the focus of the present study.

6. The Reconstruction

On analyses of the sort presented in Claudi 1988 and 1993, the syntagm S-Aux-O-V-Other is present and constantly reemerging (by various mechanisms) over much of Niger-Congo. (For Bantu see sec. 8 below.) This is observationally true synchronically, and the process can be observed today at all stages (Claudi 1988:53); to posit it as also occurring in the past is totally natural. In light of these considerations and those presented in the previous section, it is the thesis of this paper that the construction S-Aux-O-V-Other, in some form or other, has been part of Niger-Congo as far back in time as one can envision, and hence should be reconstructed in some form or other to the protolanguage. And it reconstructs as a pattern, a template conducing speakers (and the language itself) to regenerate and re-regenerate new elements in the Aux slot as old Aux elements underwent phonetic erosion and absorption by the Subject (as e.g. in Mende).

A central feature of this reconstruction is that the regeneration process, both of Aux morphemes and thus of the syntagm as a whole, is assumed to have gone through (unknowably) many cycles between Proto-Niger-Congo and the present. This dynamic cyclicity seems well-nigh unavoidable if we are aiming at a realistic reconstruction reaching all the way back to the actual protolanguage of the Niger-Congo family. Thousands of years have gone by since the time of Proto-Niger-Congo, and it seems improbable that any grammaticization scenario for Aux-formation would have taken that entire time span to slowly complete just one cycle. One need only think of Romance, where the evolution from full verb 'have' via Aux to a future suffix (as in French *parler-ai*) ran its course in less than a millennium. The time depth of Niger-Congo is much greater than that. There is of course another conceivable single-cycle option, which is to posit that the syntagm S-Aux-O-V-Other is recent in Niger-Congo; on this view we would envision a single cycle which ran its course relatively quickly, taking us back to some fairly recent intermediate boundary point before which the syntagm did not exist at all in Niger-Congo. But this would not yield a reconstruction to Proto-Niger-Congo — unless perhaps we imagined, with no evidence at all, that the word order profile of the posited intermediate stage had been preserved over millennia since the time of PNC. Any such intermediate boundary point, moreover, would obviously be an arbitrary guess. And there are in fact several positive arguments that the syntagm is actually quite old in the family; see sec. 9.

None of these difficulties arise on a multicyclical view. Natural cyclical processes, observable at many stages all over contemporary Niger-Congo, are simply assumed to have been at work indefinitely far back in time, producing variations on the same pattern again and again.

It should be stressed that the proposed reconstruction of the syntagm S-Aux-O-V-Other to Proto-Niger-Congo is not intended as a competitor to (or mutually exclusive with) the various proposed 'mechanism'-oriented accounts deriving the syntagm from other syntagms, but rather as something compatible with and complementary to them. For, whatever the mechanism which is posited for deriving any particular case of S-Aux-O-V-Other from an Aux-less syntagm, the process could not help but be facilitated by the prior existence of S-Aux-O-V-Other as a preexisting template. Any diachronic process will surely be favored if its output pattern is something already well-installed in the grammar and in speakers' competence, rather than something completely alien to the language. Speakers will

not then have to augment their grammar with a completely novel syntagm, but merely add a new paradigmatic option to a preexisting syntagmatic pattern.

A concrete demonstration of just such a template-sensitive process in action can be seen — at the morphological level — in Bantu, the one Niger-Congo family containing a language for which direct historical data is available. In Swahili, the basic verbal slot-sequence is the ordinary Bantu one,

Subj-Tns/Asp-Obj-Stem .

The Tns/Asp slot, however, represents 'an open set' of morphemes (Nurse and Hinnebusch 1993:361). In particular, future-tense markers are secondary throughout Northeast Coast Bantu: all are 'transparently innovations' (384), and none is reconstructible. The Swahili future tense/aspect-marker -ta- is derived from the verb stem -taka- 'want, like' (412), a connection still synchronically transparent in relative clauses, where the allomorph of the future marker is -taka-. The formative stage of this future marker can be seen explicitly in older Swahili texts (written in Arabic characters) reflecting the language of approx. 1700 A.D. (335). At this stage the transition from Aux+Infinitive ('want to Verb') to TnsMarker+VerbStem ('will-Verb') is still in progress, and future forms with -taka- still occur — e.g. (with 1-sg Subject indicated by aspiration of the initial t):

17) Ø-ṭhaka-ku-omba   'I will pray to you' (Miehe 1979:209).

The point to be emphasized here is that the reanalyzed helping verb -taka- is absorbed into the verb-word precisely in the Tns/Asp slot, that is, in such a way as to conform to the preexisting template. And the existence of the template was undoubtedly one of the factors shaping the reanalysis. Nurse and Hinnebusch make the point explicitly in their general discussion of the genesis of new Tns/Asp [TM] markers: 'Once the productivity of TM was thus established, it proved a magnet for other material' (458). A telling example of such attraction is provided by the category of aspect: while in morphologically more conservative Bantu languages only tense is coded in the Tns/Asp slot, with aspect expressed suffixally, in the Sabaki subgroup of Northeast Coast Bantu aspect (like tense) did come to be coded in the Tns/Asp slot (458).

Morphology is a much less flexible domain than syntax, and intraword bonds are much tighter and less permeable than interword bonds. In general, therefore, the process of paradigmatic expansion — whereby new morphemic material penetrates into a syntagm and generates a new paradigmatic substitution-element in a closed-class slot — should be much easier at the level of syntax than at the level of morphology. The kind of template-sensitive process just noted in Swahili at the word level should thus be expected to occur all the more readily at the clause level, in particular as regards the creation of new Aux elements in the syntagm S-Aux-O-V-Other.

I stated that the syntagm S-Aux-O-V-Other is to be reconstructed 'in some form or other'. But different branches of Niger-Congo, as seen in sec. 3, allow different material to count as 'Object' (only direct object; both direct and indirect object; also locative complement of verbs of motion). This presents a problem to which I have no very convincing solution: precisely what actually counted as 'Obj' in the protolanguage? Provisionally and tentatively, I reconstruct the variability as such. As remarked, a reconstructed protolanguage is a language like any language,

and as such it may perfectly well be not an ideal 'point source' with a unitary monolithic grammar, but a somewhat blurred constellation of dialects. In keeping with this view, Proto-Niger-Congo will be cast roughly in the image of Kru, as a group of dialects all of which had the syntagm S-Aux-O-V-Other but which showed various options regarding what could count as 'Object' and whether multiple objects could cooccur. The daughter languages have perpetuated sometimes this option, sometimes that. In protodialects where multiple objects occurred, I reconstruct the order seen today in Kru and in most of the other languages having multiple preverbal objects: indirect object before direct. Again as in Kru, I will allow locative complements of verbs of motion to be one possible realization of the notion 'Object', on a dialect-specific basis — certainly in Proto-Kru, quite likely in other protodialects as well. For those Kru languages where other adverbs may come between Aux and V, I assume (contra Marchese, 1986:251–54) a diachronic process of 'inbraciation', whereby the Adv moved from original postverbal to preverbal position. Here the already-preverbal position of one very prominent kind of adverb — the locative complement — would provide a model for a preverbal shift of other adverbs as well. Such inbraciation would be much more modest than that entertained (and rejected) by Marchese: only Adverbs would have to move leftward, since objects would be preverbal already.

More tentatively, I propose to assign to the proto-syntagm an additional feature found commonly in the modern languages. The reconstructed syntagm will not be 'flat' but will have internal structure, such that [O (O) V] forms a constituent over against S and Aux: S - Aux - [O - (O) - V].   The reason is simply that, in a great many Niger-Congo languages, the element symbolized 'V' in the syntagm is transparently a nominalized verb-noun (VN) form. Moreover, it is clear — as discussed at great length by Marchese and Claudi — that this stage, where V = VN, is a fundamental and recurrent phase in the cyclical life history of the construction. I interpret the nominalized verb straightforwardly as the head of a nominalized VP (i.e., V + its arguments), the latter thus constituting a subconstituent within the syntagm as a whole; the appropriate bracketing of the postverbal 'Other' is not clear to me, but at least the preverbal object(s) will be bound more tightly to the V than to S or Aux.

Further justification for these details of the reconstruction must be postponed until the end of sec. 8, when Bantu has been integrated into the argument.

7. Previous Analyses

The present paper does not focus on questions of mechanism, and I have little to add to earlier discussion of precisely how new Aux elements arise in the syntagm S-Aux-O-V-Other (or in other syntagms), and/or how Niger-Congo languages might have changed their word order over time. The dynamic of word order change in Kru is examined in minute detail in Marchese 1986. Various approaches to deriving SVO from an SOV proto-word order can be seen in (e.g.) Givón 1975, 1979 (serial verbs), Hyman 1975 (afterthought), and Williamson 1986 (verb copying), with critique in Claudi 1993. Mechanisms deriving S-Aux-O-V-Other (and OV in general) from original SVO are laid out in (e.g.) Heine 1980, Heine and Reh 1984, Claudi 1988, and especially Claudi 1993.

One point, however, bears brief discussion here. Heine (1980:104–6, cf. Heine and Reh 1984:187ff.) proposes a reanalysis of the syntagm S-Aux-O-V-Other as SVO. In languages like Ewe or Mande, the 'V' element shows clear

indications of being a verb-noun (VN), or at least of having originated in one. The 'O', in turn, is structurally the genitival dependent of this VN (like 'the shooting of the lions'). Crucially, in these languages Gen-HeadN order is the rule. If we now take Aux as the 'real' verb, with the lexical verb (V) as its complement, then transparently

$$
\begin{aligned}
& \text{S-Aux- O- V} \quad \text{- Other} \\
=\ & \text{S- V} \ \text{- [O- VN] - Other} \\
=\ & \text{S- V} \ \text{- [Gen-N] - Other} ,
\end{aligned}
$$

with [O-V] now recast as a nominalized constituent functioning as object of the Aux (recall sec. 6). The syntagm has thus metamorphosed into an instance of SVO.

This analysis embodies a presupposition: that the right way to conceive of the 'real' verb (and thus to select between the two candidates for the real verb, MainVerb and Aux) is in syntactic, not semantic, terms. It is not self-evident, however, that this is intrinsically any better (or worse) than specifying the 'real' verb on semantic grounds, an approach which would favor MainVerb over Aux. The two approaches simply clash; and the very fact that the clash exists, I would argue, is a manifestation of the inherent structural intermediacy of S-Aux-O-V-Other, as neither SVO nor SOV. Heine's analysis also has the effect of reducing the difference between   S-Aux-[O-V]-Other   and   S-Aux-[V-O]-Other   to something minor and secondary; the two become incidental variants on a single theme. Yet the former, unlike the latter, is typologically quite remarkable. Fortunately, these criticisms are largely terminological. The analysis per se is not affected by whether we choose to pigeonhole the syntagm S-Aux-O-V-Other as 'SVO' or not.

There are also more substantive criticisms which can be made on a language-specific basis. For one thing, the re-presentation of the syntagm as an instance of SVO requires the Aux to be morphosyntactically verb-like, something which holds only for some of the languages. Further, Williamson (1986:6–7) notes that the sequence [O-VN] also appears in languages which do not have Gen-HeadN order. And in languages like Kru, where *two* objects can precede the verb, the genitival reanalysis becomes very strained; nouns do not commonly take two genitive dependents (note that 'John's shooting of the lions', which does have a double genitive, does not involve two genitival *objects*). Mention should be made here of Claudi's 'Complement Serialization' proposal (1988:64ff.), which she offers as a solution to this problem in Kru.

If loosened slightly, however, Heine's analysis can apply even in a language like Kru, where genitive rection is ill-motivated. One need not, after all, insist on genitive rection in order to preserve the essential insight about the nominalized status of the V. It is not unusual crosslinguistically to find VPs nominalized while preserving the verb's case rection (including non-genitive rection of objects) more or less as in finite clauses; cf. infinitival constructions in the familiar languages of Western Europe. To be sure, the preverbal position of the object(s) now becomes largely an ad-hoc stipulation rather than a transparent consequence of the Gen-HeadN order; but this is in keeping with the general slant of this paper, which does not demand that S-Aux-O-V-Other should necessarily 'make sense' but simply accepts it as such, as a continually regenerating template.

8. Bantu

Little has been said about Bantu word order thus far except to remark that the construction S-Aux-O-V-Other occurs in only a very few of the languages. This is not surprising: in most of Bantu, the isolating, independent elements seen elsewhere in Niger-Congo have fused together to form an agglutinative verb-word, and there is no reason to expect that the new, post-agglutination sentence-level syntax (SVO) should mirror old, pre-agglutination patterns. However, *within* the Bantu verb-word, on the level of morphology, we find replicated exactly the same syntagmatic pattern seen elsewhere in Niger-Congo on the level of syntax:

```
         Niger-Congo  Subj  Aux    Obj   Verb
      =  Bantu        Subj-Tns/Asp-Obj - Stem
  e.g.   Swahili      ni -  li   - mw - ona   'I saw him'
```

It is thoroughly plausible to see the slot-sequence of the Bantu verb as a continuation of pre-Bantu syntactic patterns existing prior to agglutination.

This is internal reconstruction based on the 'Morphology as Frozen Syntax' (MAFS) principle, a hallmark of much of Givón's work. One should be cautious about appealing to this principle when doing syntactic reconstruction: not every sequence of morphemes automatically reflects earlier syntax, and applying the principle aprioristically and without independent corroborating evidence (as Givón often does) can be purely speculative, running the risk of reconstructing an artificially idealized, over-schematic protosyntax. In the present case, however, the Bantu internal reconstruction is bolstered by comparative data: precisely the right syntagm is attested, and vigorously attested, all over Niger-Congo, and even sporadically in Bantu itself.

The likelihood of the reconstruction is increased by the fact that the slot sequence of the Bantu verb, like the syntagm S-Aux-O-V-Other, is extremely rare globally, as shown by the same typological survey mentioned in sec. 4. In the typological sample the only good (though still not perfect) match to the Bantu verb was Tiwi, an isolate spoken off the north coast of Australia (Osborne 1974:37ff.). If we allow a string of clitics to count as a realization of 'inflection', then the French periphrastic future ('je-vais-le-voir') may count as a match, and perhaps the *'për të'* future in Albanian (Newmark et al. 1982: 26, 50); but in neither French nor Albanian is this the standard sequence of clitics with Auxes ('je-l'ai-vu' = Subj-Obj-Aux-V). The relevance of this to the reconstruction is straightforward. When two related languages agree in some configuration which is common globally, the agreement can easily reflect not a shared linguistic history but simply independent parallel development (e.g., in Semitic, the rise of SOV word order independently in Akkadian and Ethiopic). With a globally rare configuration, conversely, this is maximally unlikely.

Of course, this congruence between Bantu morphology and Niger-Congo syntax involves only the ordering of *bound pronominal* morphemes: in Bantu full-NPs, which are generally optional, adhere to SVO order, not to the ancestral S-(Aux)-O-V pattern. But this is hardly a counterargument against the kind of morphologization posited here, which in fact is quite common globally: old pronouns cliticize onto the verb and become agreement markers, doubling any full-NP arguments that may be present. Such morphologization is happening today in spoken French; within Niger-Congo it can be observed in-process in various

languages, e.g. Wolof, whose verbal complex involves an elaborate system of Subj and Obj clitics which generally double full-NP arguments, and also Mende (see sec. 3). In such cases, the ordering of full-NPs in the new, post-morphologization syntax may follow new principles having little to do with those governing the pre-morphologization syntax. Especially in the case of Bantu, where centuries or millennia have elapsed since the time when the agglutinated verb-word took form, there is no reason to expect the synchronic ordering of full-NPs to replicate old Niger-Congo patterns.

Not all scholars see the Bantu verb as representing a preservation (via MAFS) of earlier word order patterns. Heine (1980:103) notes that 'pre-verbal object pronouns are largely confined to the Congo branch of Bantu,' one of eight parallel branches in his subclassification of Bantu; this limited distribution, he argues, favors an analysis treating the preverbal Obj pronoun as 'a recent phenomenon confined to one Bantu branch,' and not as a survival. It is significant, then, that Tunen and Ewondo are spoken in Cameroon, a zone which not only is geographically distant from the territory of the Congo branch but is also the region of the proto-Bantu homeland (see Heine, Hoff, and Vossen 1977:60, 61, 71), precisely where maximal conservatism might be expected. These languages demonstrate that preverbal Obj markers are more widespread in Bantu than Heine indicates, strengthening the argument that the standard Bantu verbal slot sequence represents a retention and not an innovation.

The resemblance between the Bantu verb and the Niger-Congo syntagm S-Aux-O-V-Other goes beyond the general syntagmatic congruence just described, to encompass three points of detail. First, some Bantu languages allow multiple Object slots within the verb; thus in Haya (Duranti 1979:40):

> 18) a -ka   -mu -ku -leetela   'he brought him to you'   OR
>   he-PAST-him-you-brought.to   'he brought you to him'

Similarly, in Niger-Congo some families (e.g. Kru) have two Objects sandwiched between Aux and Verb, while others (e.g. Mande) have only one. The variability in number of Object-slots across Bantu languages can plausibly be seen as a synchronic reflex of inherited variability across Niger-Congo languages. The resemblance, to be sure, is not perfect. For one thing, languages like Kinyarwanda and Haya have more than two Obj slots in the verb (1979:34), while families like Kru are limited to two Objects. Moreover, the ordering constraints are different: non-Bantu Niger-Congo languages generally have indirect object preceding direct (see sec. 3), whereas in Bantu languages like Haya the indirect object often occurs next to the verb stem, with the ordering governed by a complex of interacting topicality constraints (per Duranti). Nonetheless, the pre-agglutinating Bantu languages Tunen and Ewondo do show the same IndObj-DirObj ordering seen in most of non-Bantu Niger-Congo; this argues that the more flexible ordering principles just mentioned are likely to represent a later innovation.

Secondly, the positioning of Locative complements between Aux and V in Kru, and their analysis as objects, also finds parallels in Bantu. Locatives are commonly marked on the Bantu verb; as such, they are recast as objects, with the locative coded as a locative (i.e., via a locative class-marker) in the Obj slot. Thus in Chichewa (Sam Mchombo, p.c.):

19) msika   ndi-na -ú - óna   'I saw the market'
     market   I-PAST-it -see   (Obj = class-marker *u*)
     [mu msika] ndi-na -mú -lówa   'I entered the market'
                         -enter   (Obj = Loc class-marker *mu*)
   * [mu msika] ndi-na -ú -lówa

Finally, an argument has been advanced by Myers (1992) for recognizing subconstituents within the Bantu verb; on his analysis, the major constituent break in the word is as follows: [Subj - Tns/Asp] # [Obj - Stem] ,   with the Obj bound more closely to the Stem than to the rest of the verb-word. This parallels arguments made above for several Niger-Congo families, whereby the sequence [O-V] in the syntagm Subj-Aux-O-V-Other is similarly treated as forming a constituent — in fact, a nominalized verb with 'O' as its dependent noun argument (see sec. 6).

The above similarities are not identities, and I do not mean to argue that the behaviors in Bantu and the rest of Niger-Congo are precisely parallel. My point is just that many of the issues which are characteristic and problematic for Niger-Congo clause-level syntax are likewise characteristic and problematic for the Bantu verb. This strengthens the thesis that the syntagmatics of the Bantu verb should be seen at bottom as a retention from Niger-Congo and not as an innovation. It also provides further support for the variable reconstruction tentatively put forward at the end of sec. 6: reconstructing Proto-Niger-Congo as a variable dialect cluster mirrors not only Kru but Bantu as well.

9. The Status of the Reconstruction

With the inclusion of Bantu, the syntagm S-Aux-O-V-Other is attested over the entire geographical and much of the genetic range of Niger-Congo, from Mande, Kru, and Atlantic in the north and west to Bantu in the south and east. Such breadth in itself would already argue that the phenomenon is old in Niger-Congo. The fact that the syntagm is found specifically in Bantu and in Mande has special methodological significance, and provides two additional arguments for its presence at an early stage in the history of Niger-Congo. Most concretely, its fossilized presence all across present-day Bantu on the morphological level — assuming the basic slot-sequence of the Bantu verb to have been more or less frozen at its inception and unchanged since — bears direct witness to its presence on the syntactic level (as S-Aux-O-V-Other) at the time of Proto-Bantu, or at least at that early stage in the history of Bantu when the characteristic agglutinative verb-structure took form. And the vigorous presence of the syntagm in Mande, the earliest branch-point on the Niger-Congo family tree (Williamson 1989:13), plausibly pushes it back to that stage in Niger-Congo when Mande split off from the rest of the family.

Are there alternatives to reconstructing the syntagm back to Proto-Niger-Congo? Yes; but they seem less tenable. One could, first of all, posit that that all (or many) of the occurrences in the various Niger-Congo subfamilies represent cases of totally independent parallel development — not merely independent genesis of concrete instantiations of S-Aux-O-V-Other, but multiple creation of the syntagm *ex nihilo* in languages which previously did not have it at all. This explanation is weakened by the great crosslinguistic rarity (see above) of both the syntagm S-(Aux)-O-V-Other and the verbal slot-sequence of Bantu. The more rare and

marked a phenomenon is globally, the harder it is to ascribe multiple occurrences in the same family simply to independent parallel development.

More plausibly, one might see the multiple occurrences of S-Aux-O-V-Other as an areal phenomenon (thus Carlson 1991:201–2; for Kisi see Childs 1988:34–35). In favor of this is the fact that the syntagm also occurs in dialects of Songhai, a non-Niger-Congo language spoken adjacent to Mande territory (1991:201). And indeed, within Niger-Congo several of the families where the syntagm occurs are situated more or less in the same general region, viz. western-most West Africa. However, the syntagm surely also existed in Proto-Bantu — which was spoken in Cameroon, many hundreds of miles to the east. Here an areal explanation is unmotivated: the relevant geographical territories are far apart and not contiguous. Any such area, clearly, would represent a reconstructed hypothetical entity rather than an empirical fact 'on the ground'. Moreover, it would have to have been extremely old, and essentially all the languages partaking in it would be Niger-Congo (assuming the Songhai phenomenon, which does not occur in all dialects, to be secondary). We would thus be dealing with an ancient Sprachbund of related languages. But the very concept of 'area' or 'Sprachbund' in such instances becomes conceptually more and more tenuous the further back in time we go: genetic inheritance and Sprachbund inter-influence become increasingly difficult to distinguish. In the limiting case, the separate 'languages' merge into dialects of the protolanguage, and the 'areal' commonality (in this instance S-Aux-O-V-Other) reverts to a feature of the protolanguage — the basic thesis of this paper.

Common genetic inheritance, then, would seem to provide a better explana-tion than either independent parallel development or areality.

One immediate advantage of reconstructing S-Aux-O-V-Other to Proto-Niger-Congo is that it goes far toward defusing the problem of the coexistence in Niger-Congo of SVO and SOV languages. Positing PNC as SOV will require an elaborate scenario to derive SVO languages (such as that found in Williamson 1986; see critique in Claudi 1993:35ff.); and conversely, for original SVO, with regard to modern-day SOV languages (Ijo). Positing S-Aux-O-V-Other, however, renders this problem much less acute, for two reasons: because the syntagm is structurally intermediate between SVO and SOV, and because it is a marked construction. Because it is structurally intermediate, it is naturally easier to chart a path from it to either SVO or SOV than it would be to go from one extreme to the other. Because it is distributionally and typologically marked, it embodies an inherent potential to change into something less marked, if not reinforced by strong template pressure from a whole paradigm of S-Aux-O-V-Other constructions. One can easily imagine circumstances where this reinforcement could dwindle and the construction thereby become vulnerable — where, e.g., phonetic erosion and fusion wiped out most of the Aux morphs at some period in a language's history, leaving the syntagm only weakly installed in speakers' competence and hence freer to develop in various ways into a less marked construction. The Aux might be totally lost so as to yield S-O-V-Other, itself a highly marked construction which would readily lend itself to abductive reanalysis either as 'orthodox' verb-final SOV (because O comes before V) or as SVO (because the verb is non-final). Which option was taken would depend largely upon other factors, both internal (e.g. the language's NP-level word order profile) and external (e.g. the linguistic area in which the language was located, the possible bilingualism of its speakers, and the word order of whatever other language(s) they may have spoken). And other paths of change are also imaginable.

As remarked, the logic behind the reconstruction is one which is standard in orthodox phonological/morphological reconstruction: a phenomenon that is widely attested in a family stands an excellent chance of reconstructing to the proto-language, and above all a marked or quirky phenomenon. But it would seem that this methodology, applied to Aux constructions in Indo-European, explicitly yields the wrong answer. Periphrastic tenses — formations involving an Aux word — are a commonplace in many branches of modern Indo-European (Romance, Indic, etc.), and occur in one ancient language (Hittite) as well. This data distribution would appear closely analogous to that found in Niger-Congo: here too the construction is widespread in the modern languages and can be assumed with reasonable confidence to have existed in one ancient language (Proto-Bantu). In Niger-Congo, I have argued that such a distribution warrants reconstruction of the Aux syntagm all the way back to the protolanguage. But for Indo-European, this argument fails: despite their great frequency in the modern languages, and despite their occurrence in ancient Hittite, Aux constructions are not a prominent feature of Latin, Sanskrit, etc., and do not reconstruct to Proto-Indo-European. Wherever they appear they are manifestly secondary. Is it not risky, then, to apply the analogous argument to Niger-Congo?

But the two cases are only superficially similar. The crucial difference is that in Indo-European, unlike Niger-Congo, there is no one Aux syntagm which is common and characteristic over much of the family. Rather, Indo-European periphrastics fit a variety of patterns, each fairly well constrained genetically and areally:

S Aux V X	French: (Romance)	l'homme a acheté le livre the man has bought the book 'The man (has) bought the book'
S Aux X V	German: (Germanic)	der Mann hat das Buch gekauft the man has the book bought 'The man (has) bought the book'
Aux S V X	Welsh: (Celtic)	mae 'r dyn yn prynu 'r llyfr is the man in buy(ing) the book 'The man buys/is buying the book'
S X V Aux	Urdu: (Indic)	ādmī kitāb xarīd-tā hai man book buy-PTCP is 'The man buys the book'
S X V Aux	Hittite: (Anatolian)	GIŠGIGIR turiyan ḫarweni (Friedrich chariot harnessed we.have 1974 I:137) 'We have harnessed the chariot'

(The Hittite and the Indic constructions represent independent historical developments.) The recurrence seen in Niger-Congo is thus much more precise than its Indo-European counterpart: it involves not just the general fact of a periphrastic construction, but the precise syntagmatic details of the construction itself. Moreover, none of the various Indo-European Aux syntagms appears particularly exotic crosslinguistically. By contrast, Niger-Congo S-Aux-O-V-Other is definitely

an exotic, and as such — as with any family-wide quirk — is a good candidate for reconstruction.

10. Conclusion

This paper has deliberately ignored the general question of the overall word order of Proto-Niger-Congo, focusing instead on one quite specific and highly characteristic syntagm. But reconstructing this syntagm to the protolanguage has consequences for our assessment of the 'basic' word order, too. Proto-Niger-Congo may or may not have had an Aux-less syntagm alongside S-Aux-O-V-Other (recall that Mandinka has no Aux-less syntagm); if such a syntagm did exist, it may have been SOV or SVO (though I find Claudi's (1993) case for SVO much stronger than the argument for SOV), or even S-O-V-Other as in Mande. Regardless of what such an Aux-less syntagm may have looked like, however, I have argued that Proto-Niger-Congo *also* had a syntagm, very sharply profiled and not at all marginal, which was fundamentally neither SVO nor SOV — a syntagm cast in an essentially intermediate mold all its own. If so, the dichotomy posed at the beginning of this paper, 'Niger-Congo: SVO or SOV?', is revealed as the wrong way to frame the problem. For it would appear that the single most reliable conclusion that can be drawn regarding word order in PNC, the reconstruction which is at once simplest, least speculative, and most in keeping with the comparative data and with standard comparative methodology, is that the protolanguage featured very prominently a syntagm which was neither straightforwardly SVO nor SOV.

The approach outlined here has at least one corollary which does bear directly on the dichotomous choice of SVO vs. SOV. The syntagm S-Aux-O-V-Other, which plays a central role in the body of evidence adduced by proponents of 'SOV Proto-Niger-Congo', can no longer be appealed to in this way: it has been recast as evidence for something else. The argument that PNC was SOV (verb-final) thus becomes correspondingly more difficult to make.

It is tempting in historical linguistics to always seek to trace things back to ultimate causes. Reconstructing the syntagm S-Aux-O-V-Other to Proto-Niger-Congo feels like begging the question: the syntagm must surely have come from *something*. Indeed; and many scholars have laid out plausible (and competing) scenarios for the recurrent development of the syntagm in particular Niger-Congo languages, and hence, in all likelihood, in Proto-Niger-Congo itself and even earlier. The purpose of this paper has not been to take a position for or against any of these scenarios, but simply to argue that S-Aux-O-V-Other itself shows every likelihood of having been a part of Niger-Congo as far back as we can reasonably project. Its history prior to PNC is unknown. Tautologically, some state of affairs preceded that of Proto-Niger-Congo; but (absent any clear picture of comparative Kordofanian syntax) we have no access to it via comparatively based reconstruction.

References

Aginsky, Ethel G. 1935. A grammar of the Mende language. Language Dissertation no. 20. Philadelphia: Linguistic Society of America.
Carlson, Robert. 1991. Grammaticalization of postpositions and word order in Senufo languages. In Approaches to grammaticalization, ed. Elizabeth Closs Traugott and Bernd Heine, II:201–23. Amsterdam: Benjamins.

Childs, Tucker. 1988. The phonology and morphology of Kisi. PhD dissertation, University of California at Berkeley.

Claudi, Ulrike. 1988. The development of tense/aspect marking in Kru languages [review article on Marchese 1986]. Journal of African Languages and Linguistics 10:53–77.

———. 1993. Die Stellung von Verb und Objekt in Niger-Kongo-Sprachen. Afrikanistische Monographien 1. Cologne: Institut für Afrikanistik, Universität zu Köln.

Creissels, Denis. 1983. Eléments de grammaire de la langue mandinka. Grenoble: Publications de l'Université des Langues et Lettres.

Davies, John. 1989 [1981]. Kobon. London & New York: Routledge.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. Language 68:81–138.

Dugast, Idelette. 1971. Grammaire du Tunen. Paris: Klincksieck.

Duranti, Alessandro. 1979. Object clitic pronouns in Bantu and the topicality hierarchy. Studies in African Linguistics 10:31–45.

Friedrich, Johannes. 1974. Hethitisches Elementarbuch. Heidelberg: Carl Winter.

Givón, Talmy. 1975. Serial verbs and syntactic change: Niger-Congo. In Word order and word order change, ed. Charles N. Li, 47–112. Austin: University of Texas Press.

———. 1979. Language typology in Africa: A critical review. Journal of African Languages and Linguistics 1:199–224.

Heine, Bernd. 1976. A typology of African languages based on the order of meaningful elements. Berlin: Dietrich Reimer.

———. 1980. Language typology and linguistic reconstruction: The Niger-Congo case. Journal of African Languages and Linguistics 2:95–112.

Heine, Bernd, Hans Hoff, and Rainer Vossen. 1977. Neuere Ergebnisse zur Territorialgeschichte der Bantu. In Zur Sprachgeschichte und Ethnohistorie in Afrika, ed. Wilhelm J. G. Möhlig et al., 57–72. Berlin: Dietrich Reimer.

Heine, Bernd, and Mechthild Reh. 1984. Grammaticalization and reanalysis in African languages. Hamburg: Helmut Buske Verlag.

Huehnergard, John. 1989. The Akkadian of Ugarit. Atlanta: Scholars Press.

Hyman, Larry. 1975. On the change from SOV to SVO: Evidence from Niger-Congo. In Word order and word order change, ed. Charles N. Li, 113–47. Austin: University of Texas Press.

Innes, Gordon. 1962. A Mende grammar. London: Macmillan and Company, Ltd.

———. 1963. The structure of sentences in Mende. London: School of Oriental and African Studies.

Kimball, Geoffrey D. 1991. Koasati grammar. Lincoln: University of Nebraska Press.

Li, Charles N., and Sandra A. Thompson. 1974. Historical change of word order: A case study in Chinese and its implications. In Historical linguistics, ed. John M. Anderson and Charles Jones, I:199–217. Amsterdam: North-Holland.

———. 1981. Mandarin Chinese: A functional reference grammar. Berkeley & Los Angeles: University of California Press.

Marchese, Lynell. 1986. Tense/aspect and the development of auxiliaries in Kru languages. Dallas: Summer Institute of Linguistics.

Meillet, Antoine. 1967. The comparative method in historical linguistics. Trans. Gordon B. Ford, Jr. Paris: Librairie Honoré Champion.

Miehe, Gudrun. 1979. Die Sprache der älteren Swahili-Dichtung. Berlin: Dietrich Reimer.

Migeod, F. W. H. 1908. The Mende language. London: Kegan Paul.

Myers, Scott. 1992. The morphology and phonology of INFL in Bantu. University of Texas ms.

Newmark, Leonard, Philip Hubbard, and Peter Prifti. 1982. Standard Albanian: A reference grammar for students. Stanford: Stanford University Press.

Nurse, Derek, and Thomas J. Hinnebusch. 1993. Swahili and Sabaki: A linguistic history. University of California Publications in Linguistics 121.

Osborne, C. R. 1974. The Tiwi language. Canberra: Australian Institute of Aboriginal Studies.

Redden, James. 1980. A descriptive grammar of Ewondo. Occasional Papers on Linguistics, No. 4. Carbondale, Ill.: Dept. of Linguistics, Southern Illinois University.

Welmers, William E. 1973. African language structures. Berkeley & Los Angeles: University of California Press.

Williamson, Kay. 1986. Niger-Congo: SVO or SOV? Journal of West African Languages 16:5–14.

———. 1989. Niger-Congo overview. In The Niger-Congo languages, ed. John Bendor-Samuel, 3–45. Lanham, MD: University Press of America.

A Diachronic Approach to Classes 10 and 11 in Bantu With Special Reference to North-Western Languages

Claire Grégoire
Royal Museum for Central Africa - University of Brussels

1. Introduction

1.1.   Some Bantu languages of zone B and zone C have a class 11 noun prefix whose vowel cannot be a regular reflex of the second degree rounded vowel which has been reconstructed in Proto-Bantu for this noun prefix and for its augment. Meeussen (1967) reconstructs *dó–dò– as the prefixal sequence for this class. Interestingly, in the languages in question, the vowel of the class 11 noun prefix is i, e, ɛ, ə or a, the consonant of the prefix being however a normal reflex of *d. Languages which exhibit this phenomenon are found in the area along the borders of Gabon, Congo-Brazzaville and Zaïre and belong to the subgroups B20, B50, B70, B80, C20, and C80 of Guthrie's classification. The class 11 noun prefix is in fact of the type:

Ca– in Kele, Ngom N, Ngom S, Mbaŋwe (B20), Mbuun (B80), Bushong (C80)
Ci– in Wumvu, Seki (B20), Duma (B50), Yaa, Fumu (B70), Ngongo, Tsong (B80)
Ce– in Ndasa N, Mahongwe (B20), Mbaama (B60), Koyo, Mboshi (C20)
Cɛ– in Sake (B20), Laale (B70)
Cə– in Nzebi (B50)

Depending on the noun considered, the prefix is li– or le– in Sigu (B20), and lɛ–, la–, or li– in Ndasa S (B20). Furthermore, it is interesting to notice that Lwel (B80) has lu–, lə–, or la–, whereas Yans (B80) has lo– or le– as a class 11 noun prefix, depending on the dialect considered. The same situation can be observed in a variant of Teke spoken in Gabon, where, according to Fontaney (1984), the alternation between the class 11 noun prefixes la– and le– is also due to dialectal variation.

The preceding list calls for a remark: it may be difficult to distinguish a noun prefix of class 11 li–, le–, lə– from a noun prefix of class 5, and it is true that both prefixes have merged in a number of languages. Therefore the list contains only languages in which the two prefixes are formally distinct or languages which, even though they have merged the two prefixes, have kept the pairings 5~11, 6 and 5~11, 10. The existence of the pairing 5~11, 10 precisely seems to show that the evolution entailed a formal confusion of the two classes rather than a substitution of class 11 by class 5 or a substitution of the pairing 11, 10 by the pairing 5, 6. In order to allow the reader to observe the aberrant class 11 prefixes, a few illustrative examples are given below:

B20	Kele	bàsúyì	súyì	hair	cl. 5 ɾè–
	Mahongwe	lèhùvé	màhùvé	hair	cl. 5 i–
B50	Nzebi	lənáːnga	náːnga	hair	cl. 5 lə–
B70	Fumu	likvi	nkvi	firewood	cl. 5 li–
	Laale	lɛlimɛ́	ŋlimɛ́	language	cl. 5 ∅–
B80	Yans	a. lekay	nkay	leaf	cl. 5 li– / i–
		b. lokay	nkay		according to dialects

The map on next page shows where the different forms are localized.

Aberrant vowels in the class 11 noun prefix.

Cl.11 NP Ca–
Cl.11 NP Ci–
Cl.11 NP Ce–
Cl.11 NP Cɛ–
Cl.11 NP Cɔ–
Cl.11 NP Cu–, Co–

0 50 100 150km
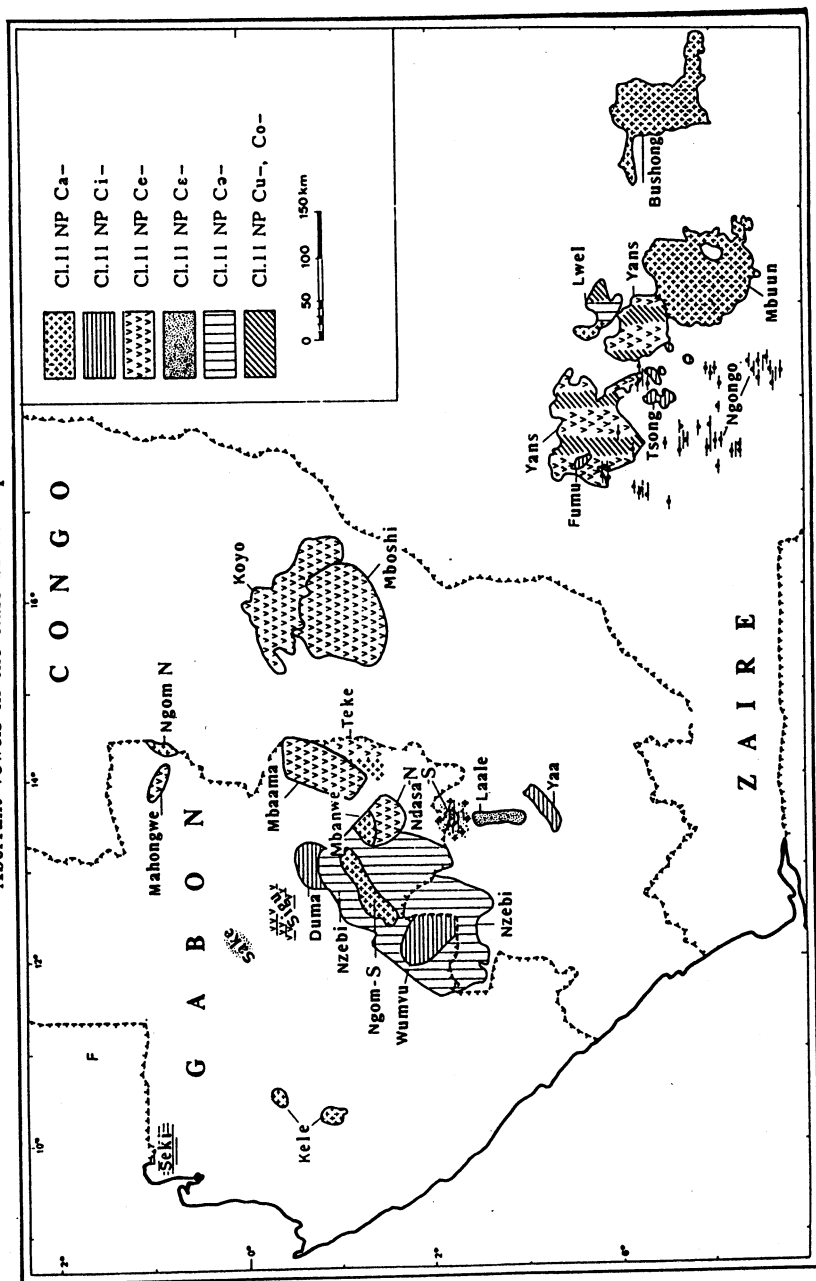
1.2. A comparative study should establish the causes of the observed vocalic mutations in the class 11 noun prefix. Apart from the developments which occurred later and probably determined the occurrence of vowels such as ǝ and ɛ, these mutations do not seem to be due to mere phonological change. Firstly, they did not affect all noun prefixes of the type Co-, Cu-. They sometimes appear in the noun prefix of class 13 (< *tɔ̀-) for which the forms ta- or t i- are found and, more rarely, in the locative noun prefix of class 17 (< *kɔ̀-) for which the form ka- is found; but, in all the languages investigated, the class 1 and 3 prefixes (< *mɔ̀-), or the prefix of class 14 (< *bɔ̀-) have kept a rounded vowel. Secondly, it is important to note that generally in the languages in question, the reflexes of the protovowel *o seem to be normal.

A full understanding of the mechanisms which have brought about the aberrant forms of class 11 noun prefix presupposes, without doubt, a thorough diachronic study of the various languages where these forms appear, and in particular a good knowledge of the phonological reflexes which characterize them as well as a good knowledge of the evolution undergone by their class systems. Such a study is being carried out but has not yet reached final results. However, a reconsideration of the restructurings characterizing the pairings 9, 10 and 11, 10 in certain Bantu zones allows the formulation of interesting hypotheses concerning the appearance of divergent vowels in class 11 noun prefix. The consequent tentative hypotheses are formulated hereunder, following a close study of the evolution of the pairings 9, 10 and 11, 10 in some particular languages.

2.   The pairings 9, 10 and 11, 10

2.1. First, it should be remembered that the prefixal sequences comprising the noun prefix and its augment are reconstructed as *jè-ǹ- for class 9 and *dí-ǹ- for class 10. These reconstructions are those proposed by Meeussen (1967) except for class 10 whose augment had been reconstructed as *jí- at that time. The author judged later that a reconstruction *dí- for this morpheme was more correct. We adopt this reconstruction, even if Meeussen did not defend it in any publication.

2.2. It is probably useful to reiterate some obvious facts concerning the morphemes in question. According to Meeussen, the reconstruction *jè- for the class 9 augment could easily be replaced by *è-, making class 9 distinct not only by the structure of its noun prefix but by its vocalic augment as well. The homophony of class 9 and 10 noun prefixes is entirely exceptional in Bantu within a singular / plural pairing.

In many current Bantu languages which have lost the augment, class 9 and 10 nouns are formally identical; the singular is no longer different from the plural. This homophony seems to be easily accepted by a number of linguistic systems, all the more as the singular / plural opposition can be maintained in the concords which frequently are differentiated by segments and / or by tones. The formal confusion of class 9 and 10 nouns can also occur in languages where the augment has generally become vocalic and where the shift from 7 to 5 vowels has merged in i - the second degree vowel which characterized the class 9 augment and the first degree vowel which characterized the class 10 augment. However, many systems disseminated in various zones have kept a formal distinction between corresponding singular class 9 nouns and plural class 10 nouns.

A few languages, even though they have acquired a vocalic augment in class 10, have retained the aperture difference inherent to the vowels of the two augments. The result is, for example:

C30	Doko	9, 10	éngúlú	ingúlú	pig
			émbûdù	imbûdù	bird
			ékókómbè	ikókómbè	sparrowhawk

Some languages have a vocalic augment in class 9 but a CV- augment in class 10, as in:

| P20 | Mabiha | 9, 10 | indimbe | dindimbe | antelope (sp.) |

Other languages, though they have on the whole lost the augments, have nonetheless maintained the CV- augment in class 10 whose only function is to oppose formally the plural of class 10 to the singular of class 9, as in:

K30	Luyana	9, 10	ndila	tindila	path
			ngombe	tingombe	cattle
N40	Nyungwe	9, 10	mvura	zimvura	rain

The type of formal opposition which characterizes such pairings as in Mabiha, Luyana and Nyungwe may suggest a reanalysis which changes the former class 10 augment into an additive noun prefix, preprefixed to that of class 9. This evolution has certainly been favoured by the CV- structure of class 10 augment, as it is identical to the canonical structure of noun prefixes in Bantu languages. Since many Bantu languages apply rules which delete the nasal prefix in certain environments depending on the initial consonants of the noun stem, it is worth noticing that the ancient class 10 augment may be the only class morpheme which is represented in the noun. If examples such as

| S30 | Lozi | 9, 10 | tau | litau | lion |
| | | | kuhu | likuhu | hen |

become numerous in gender 9, 10, the language will tend to reanalyse the morpheme li < *dí no longer as an augment but as a noun prefix of class 10.

It may happen then, that this new class 10 prefix acquires by analogy a vocalic augment i– and, possibly, becomes an autonomous noun prefix of a plural class included in a new pairing. In Zulu, for instance, the pairing 9, 10 uses the sequences i–n+ and i–zi–n+ with readdition of an augment i– to the ancient class 10 augment zi–, now considered to be a noun prefix. The sequence i–zi– makes the plural of class 7, with no following nasal except when the latter is found in the singular[1]. The result is:

S40	Zulu	9, 10	imbuzi	izimbuzi	goat
			indlu	izindlu	house
		7, 8	isilo	izilo	thing
			isifuɓa	izifuɓa	chest
			isinkwa	izinkwa	bread

The same evolution can be observed in Tswana for example, where the nasal of the class 9, 10 prefixes is directly represented with monosyllabic stems only, though it is indirectly represented with disyllabic stems through the particular nature

of the initial consonants. As in Zulu, the noun prefix d i - of class 10 is used without a following nasal consonant to form the plural of class 7. The results are, for example: ·

S30	Tswana	9, 10	ŋkwè	dìŋkwè	leopard		
			m̀phɔ́	dìm̀phɔ́	gift	cf. +fá	to give
			qhósí	dìqhósí	chief		
			pɔ́tsɔ́	dìpɔ́tsɔ́	question	cf. +bɔ́tsá	to ask
		7, 10	sɩ̀lɔ̀	dìlɔ̀	thing		
			sɩ̀fòfù	dìfòfù	blind		
			sɩ̀dìbà	dìdìbà	well		

We have therefore what some Bantuists call a class 10b, that is a class whose concords are those of class 10 but which differs from this class by the prefixal sequence used in nominal forms. In Zulu or in Tswana, these sequences are distinguished only by the presence versus the absence of the nasal consonant. But Myene-Nkomi (B11e), for example, has a prefixal sequence i -n- in class 10 plural of class 9, but a sequence i -d i - in class 10b plural of a class which results from the confusion of the former classes 11, 14 and 15. Besides, it is particularly interesting to observe that a similar restructuring has occurred in a number of languages of zone S and in the languages which belong to the subgroups B10 and B30. Concerning zone B languages, authors such as Blanchon (1987) insist on the fact that the pairing 11, 10 could have played an important role in the process which led to the appearance of class 10b. The existence of a (V-)CV- sequence in class 11 could have brought about the generation, by analogy, of a prefixal sequence such as i -d i - in the plural. This is quite possible but it must also be remarked that, in some languages, the pairing 9, 10 may contain in itself all the elements apt to give way to such a restructuring. This aspect cannot be neglected.

It is true that various analogical restructuring procedures happen within the pair 11, 10, according to two opposing tendencies which coexist more or less frequently in the same linguistic system, where they can moreover determine variants between which the speaker hesitates. The first of the two tendencies leads to the elision of the nasal prefix which characterizes the plural noun, and consequently this noun shifts to class 10b. This can be observed in a pairing like:

| S40 | Zulu | 11, 10 | ululimi (uːlimi) | izindimi | language |
| | | | | or izilimi | |

The second tendency on the contrary, leads to the inclusion of a nasal prefix in the singular noun, between the class 11 noun prefix and the nominal stem. This is apparent, for example, in pairings like:

| D10 | Lengola | 11, 10 | lundelu | ndelu | beard |
| | | | lungbíli | ngbíli | hair |

This last case is interesting because it shows how class 11 can tend to become an autonomous class whose additive prefix is preposed to that of another class, in a complex prefixal sequence of the type CV-n+. The present paper comes back to this point at a later stage.

2.3.   It is important to observe the procedures which have been used to maintain a formal difference between class 9 singular nouns and class 10 plural nouns in

languages in which the augments merged into a unique vowel which is identical in all classes (or in almost all classes).

2.3.1. Such a situation is observed, for instance, in Nyanga and in Hunde (D40 and J50). In both languages, the augment is á – in all classes, with the exception of class 5 in Nyanga. This augment still has a grammatical function since its absence versus its presence expresses an opposition like 'indefinite versus definite'. The analysis of examples such as:

	Class 9		Class 10		Gloss
	Indefinite	Definite	Indefinite	Definite	
Nyanga	mbу̰ra	ámbу̰ra	mbу̰ra	ásámbу̰ra	rain
	ncangi	áncangi	ncangi	ásáncangi	cloth
Hunde	ngóko	ángóko	ngóko	ásángóko	hen

shows that class 9 and class 10 nouns are identical when indefinite, but that they are formally distinct when definite. It shows particularly that in the definite plural, the two languages use what can be diachronically analysed as an accumulation of augments since the prefixal sequence comprises a first unified augment á –, a class 10 augment s í – and a second unified augment á – which precedes the prefixal nasal. In Ronga (S50), which also has a unified augment a –, the augment of the nasal prefix does not appear in the class 10 form and we have for instance a t i m b u t i 'goats'. Conversely, in Nyanga and in Hunde, the class 9 augment á – is maintained after the ancient class 10 augment. This last morpheme is, to some extent, analysed as an additive prefix since it is preceded by another augment á –. Interestingly enough, Nyanga and Hunde maintain also the á – augment of noun prefixes when these morphemes are preceded by diminutive prefixes of classes 12, 13 or by augmentative prefixes of classes 7, 8. As a result, we obtain:

D40	Nyanga	kámutí	twámutí	small tree
		cámutí	byámutí	big tree
J50	Hunde	kákúbóko	twámábóko	small arm
		kyákúbóko	byámábóko	big arm

One could argue that the maintenance of the augment after an additive prefix is not frequent in Bantu languages. However, we showed elsewhere (Grégoire 1975) that these languages, in a former stage of their evolution, probably retained the augment of the noun prefix after an additive locative prefix. Moreover, in a number of languages, additive prefixes (diminutive or augmentative) have a long vowel. Consequently, a thorough study may well prove that, if in some languages the class 13 noun prefix has a vowel which is not a first or second degree rounded vowel as would be expected from the reconstruction * t ò –, it is because it has incorporated into its own structure the vocalic augment which it preceded in a previous stage of the evolution. This hypothesis seems to be backed up by alternations such as:

K50	Phende	12, 13	gáákází	túúkází	small wife
			gáámulúmé	téémilúmé	small male

in which the class 13 noun prefix has a normal vowel when it is not followed by another prefix but can have an aberrant vowel in the presence of a following prefix.

2.3.2. It is also important to observe how a number of languages in zones R, K and H have reacted to the necessity of maintaining a formal distinction between class 9 nouns and class 10 nouns. In Herero, Nkhumbi, Nyaneka, Mbundu S (R), Kwangali (K) and in a particular variant of Mbundu N (H) spoken near Amboïm, a vocalic unified augment o- is used in almost all classes. Entirely frozen, the augment o- is included in the noun prefix and has no grammatical function. In Kwangali (K), Nyaneka and Nkhumbi (R), the following forms are found:

K30	Kwangali	9, 10	nzira	nonzira	path
			mbwa	nombwa	dog
			hũhwã	nohũhwã	hen
R10	Nyaneka	9, 10	ongombe	onongombe	cattle
			ofufua	onofufua	hen
	Nkhumbi	9, 10	ongolo	onongolo	knee
			osefe	onosefe	trial

None of these languages has a reflex *d > n and the prefixal sequence attested in class 10 cannot therefore contain a reflex of the ancient augment *dí- proper to this class. It seems rather that it is the sequence Augment + class 10 NP o-n- which could have been treated as additive with regard to the class 9 homophone sequence. In Kwangali, where, unlike what happens in zone R, the general augment o- is latent and manifests itself only in particular syntactic contexts, we effectively obtain nⁱonzira 'with the path (cl. 9)' and eyuvi limwe lyⁱonomfi 'a thousand fishes (cl.10)'. However, the augment generally does not appear after additive prefixes as we see in: kangombe, tungombe 'small cow(s)', kambwa, tumbwa 'small dog(s)'. We may thus suppose that the representation of the class 9 augment o- within the class 10 sequence (o)non- is due to the consonantal structure of the two noun prefixes and hence to syllabification constraints. Furthermore, it will be noted that examples such as

K30	Kwangali	9, 10	hũhwã	nohũhwã	hen
R10	Nyaneka	9, 10	ofufua	onofufua	hen
		11, 10	oluhuki	onohuki	hair

show that the class 9 augment o- is kept when a rule deletes the nasal consonant which was its original prefixal support. This procedure tends to create a class 10 prefixal sequence (o-)no-where this augment is integrated into a (V-)CV structure.

In Herero and in Mbundu S, the pairings 9, 10 appear as follows:

R30	Herero	9, 10	ongombe	ozongombe	cattle
R10	Mbundu S	9, 10	ónjíla	ólonjíla	path
			ómbwá	ólombwá	dog
			ongombe	olongombe	cattle

It can be observed that in these languages, the restructuring process which reaches the class 10 prefixal sequence is similar to the one observed in Zulu or in the definite forms of Nyanga and Hunde. Once again, class 10 is characterized by what can be historically analysed as an accumulation of augments. Its prefixal sequence seems in fact to include a unified augment o-, an ancient class 10 augment *dí-, another unified augment o- and a nasal prefix. Concerning Mbundu S for instance, the following restructuring process is suggested:

```
                    d í  –        ǹ +
          o  –  l i  –  o  –  n +
          o  –  l    –  o  –  n +
                 └──────┬──────┘
          o  –      l o     –  n +
```

Along with this restructuring process, an additive class 10 prefix o l o – is created. It will be noted that in Mbundu S, the class 11 noun prefix is o l u – so that only the aperture degree of the second vowel allows one to distinguish the singular from the plural in pairings 11, 10 such as:

R10	Mbundu S	o l uma t i	o l oma t i	rib
		o l unh i h i	o l onh i h i	bee
		o l uhwĩ	o l ohwĩ	firewood

The dialectal variant of Mbundu N (H) spoken near Amboïm has a class 10 prefixal sequence which is interesting due to the nature of its second vowel. The following examples are quoted from da Silva Maia (1964):

H20	Mbuim	9, 10	mbu l i	o l ambu l i	goat
			ng omb e	o l ang omb e	cattle
		11, 10	l uhwĩ	o l ahwĩ	firewood
			o l upa j i	o l apa j i	rib
			l unzwa na	l anzwa na	nail, claw
			l ukamba	l akamba	bird of prey (sp.)

In this prefixal sequence, the vowel a – recalls what we observed in the prefixal sequences á s á n+ of Nyanga and Hunde. Yet, these two languages have a generalized augment á –, which is found neither in Mbundu N nor in the variant of this language spoken in Amboim. The ultimate problem is thus to determine the origin of the prefixal sequence (o) l a (N)+ which tends to produce a class 10 prefix morpheme l a –, as it can be observed in: l ukamba pl. l̲a̲kamba 'bird of prey (sp)' and k a n j i l a pl. ka l̲a̲n j i l á 'small bird', k a s a n j i pl. ka l̲a̲ s a n j i 'small hen'.

Two remarks found in the descriptions lead to an interesting hypothesis. Authors point out that in Kwangali (K) and in Mbuim (H), there are two alternating plural forms for class 9 nouns (or for some of them). They give the following examples:

K30	Kwangali	mp huku	nomp huku	or	vamp huku	rat
		nz ovu	non z ovu	or	banzovu[2]	elephant
		ngwe	nongwe	or	bangwe	leopard
H20	Mbuim	ng omb e	o l ang omb e	or	ang omb e	cattle

The second of these alternating plural forms has a more collective meaning and uses a prefixal sequence which comprises an additive class 2 prefix represented by a CV-form derived from *b à – in Kwangali and by a vocalic form which resembles an augment in Mbuim. Da Silva Maia explicitly points out that there is a clear-cut tendency to replace the plural prefix l a – by a –. However, the mere existence of the sequence o l a – indicates, on the contrary, that the class 2+9 plural form is probably ancient and that the vowel a – is in fact included in a prefixal sequence which multiplies additional augments and which is hence similar to those observed in zones K and R. Therefore, the prefixal sequence attested here probably comprises a

unified augment ɔ–, a reflex of the ancient class 10 augment *dí–, an ancient class 2 augment now considered to be a prefix, and a nasal prefix. It seems to be the result of a restructuring such as:

```
                         d i  –  n  +  ɡ o m b e


  ɔ – l i – ɔ – n + ɡ o m b e    or    a  –  n  +  ɡ o m b e



                  ɔ – l i – a – n + ɡ o m b e
                  ɔ – l   – a – n + ɡ o m b e
                      └────┬────┘

                  ɔ –     l a   – n + ɡ o m b e
```
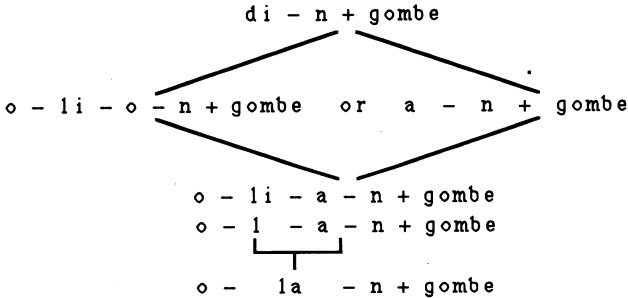
As in other languages, the nasal prefix can be deleted in some nouns. The same holds for the augment ɔ–, even if da Silva's description does not allow a clear statement of the conditions in which this morpheme is deleted. Hence the class 10 prefixal sequence can be reduced to la–, as can be observed for example in the pairing lukamba, lakamba 11, 10 'bird of prey (sp.)'.

2.4. To summarize, the data examined up to now show, on the one hand, how a class 10b - whose noun prefix is the ancient CV- augment of the reconstructed class 10 - can be progressively created. The existing asymmetry between the singular in class 9 and the plural in class 10 together with the total homophony which characterizes the prefixes of the two classes can lead to the treatment of the class 10 augment as an additive noun prefix to which a new vocalic augment.has been preprefixed As a result we have, (where = marks the boundary of what is reanalyzed as an additive noun prefix) :

cl. 9   n + CVCV   cl. 10   d i – n + CVCV
                       cl. 10 + 9   d i = n + CVCV
                                      i – d i = n + CVCV

On the other hand, when the language has a unified vocalic augment, the pre-prefixation of the new V-CV- sequence of class 10 can occur before a prefixal nasal of class 9 with its augment or before an alternating plural of class 2+9, the prefixal sequence of which is a–n+. The result is:

cl. 10 Aug.   cl. 10 NP   NP or Augment   cl. 9 NP
ɔ–   l i–   ɔ–, a–   n+

The additive class 10 noun prefix can then incorporate into its own structure the vocalic prefix or the vocalic augment before which it appears, to the detriment of its original vowel, as in the following:

V_1 - CV_2 - V_1 - n +   or V_1 - CV_2 - V_3 - n +
V_1 - C - V_1 - n + V_1 - C - V_3 - n +
V_1 - CV_1 - n + V_1 - CV_3 - n +
–>   ɔ l ɔ (n) +   –>   ɔ l a (n) +
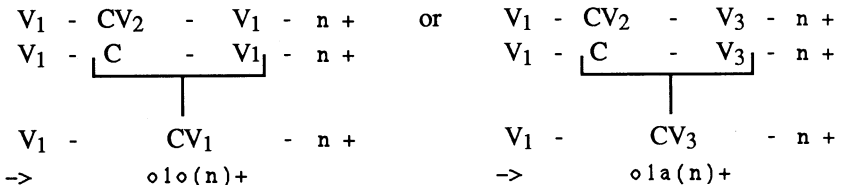
3. The aberrant forms of the class 11 noun prefix

3.1. Languages in which class 11 noun prefixes are characterized by an aberrant vowel no longer have an augment as such. Generally, they nonetheless contain numerous traces of this morpheme. For instance, they not infrequently have an alternation between a CV- form and a V- form of the noun prefix, the choice between these two allomorphs depending on the structure of the following stem and on the nature of its initial segment. The alternation between them seems to be due to the fact that the former V-CV- prefixal sequence has been reduced to the vocalic augment when it preceded a stem with an initial consonant, but to the CV- prefix when it preceded a stem with an initial vowel or, in certain languages, a monosyllabic stem. In such systems where the choice between the V- and the CV-allomorphs of the noun prefixes depends on the nature of the following segment, one might expect that an additive prefix, which is most of the time preposed to the V-allomorph of the noun prefix, would be frozen in a CV- form which is then generalized. In Koyo for instance, where all noun prefixes other than n+ have a vocalic representation before stems with an initial consonant and before nasal prefixes, it is remarkable that the class 11 noun prefix has a constant form 1 e – and that it is the only noun prefix in the system which has no vocalic allomorph. So we have:

C20	Koyo	1, 2	mo r o	ba r o	person
			mwá na	bá na	child
			okond z i	akond z i	chief
			ng ubú	angubú	hippopotamus
		11, 10	1ekó ɲi	kó ɲi	firewood
			1embánd z i	mbánd z i	rib

Similarly in the B20 languages, where the vocalic representations of noun prefixes are not rare in singular classes, the class 11 noun prefix has a V- structure only in those languages where it is characterized by a rounded vowel reflecting the reconstructed vowel. But in all the languages in which its vowel is i, e, ɛ or a, it has the CV- structure which can be expected if it were previously preposed to a vowel. Thus we have[3]:

B20	Kota	ɔ́húwὲ	húwὲ	hair
	Pove	ɔ̀1émὲ	1émὲ	language
but:				
	Kele	ὂàsúyὶ	súyὶ	hair
	Ngom N	1ásúὶ	súὶ	hair
	Ngom S	yàsúyὶ	súyὶ	hair
	Wumvu	1èsùyí	sùyí	hair
	Sigu	1íkúɲὶ	kúúɲὶ	firewood
	Ndasa N	1è1ímὶ	màlímὶ	language
	Ndasa S	1ὲ1ímὶ	màlímὶ	language
	Mahongwe	1èhùwé	màhùwé	hair
	Sake	1ὲdémὶ	mὲdémὶ	language

3.2. Another general observation concerns the languages in which the class 11 prefix has a vowel which is not a regular reflex of *ɔ-. Most of these languages have a tendency to redistribute in gender 1, 2 the nouns which historically belong to gender 9, 10, especially when they designate animals. Class 9 retains only nouns designating objects and, not infrequently, these nouns make their plural with an additive prefix preposed to the nasal prefix. Languages such as Koyo, Mboshi, Mbaama, Fumu and all the B20 languages have completely lost the pairing 9, 10 and make the plural of class 9 nouns in class 6+n. As examples, we have:

C20	Koyo	ndziá	andziá	path
		mbóga	ambóga	country
B60	Mbaama	ntulu	antulu	chest
B70	Fumu	nzo	manzo	house
B20	Mahongwe	mbókà	màmbókà	village
	Sake	ŋkúúrù	mɛ̀ŋkúúrù	strength

In all these languages, class 10 is only associated with nouns which appear in class 11 as well, but the same tendency is also manifest in other languages where the class 11 noun prefix is aberrant, even when the 9, 10 pairing has not yet completely disappeared.

On the other hand, it is important to point out that all the reviewed languages, without exception, do have the same prefixal sequence in class 9 and in class 10 (associated with class 11). This sequence is reduced to a nasal consonant which is deleted by rule in a high number of cases, in particular when the stem begins with a voiceless consonant. Such rules have probably favoured the redistribution of many class 9 nouns and, for the nouns remaining in this class, the generalisation of plural forms with an additive class 6 prefix.

3.3. It is also clear that in zones H, B or C, class 11 has a singulative semantic value. This class is used to derive nouns designating single objects, the collection of which is named as such by a corresponding noun in class 10 or in some cases in class 6(+n). Various authors note explicitly that the so-called pairing 11, 10 (or 11, 6+n) does not constitute a pairing as such, in which a singular is opposed to a plural. For instance, when observing the situation in Mboshi where the pairing 9, 10 no longer exists, Prat (1917) writes: 'Avec certains substantifs, le préfixe le- opère une véritable soustraction pour distinguer séparément un des éléments d'une combinaison complexe ou d'une collectivité marquée par la forme du pluriel, plus usitée ici que le singulier'. The following translations, some of them given by the author, can be proposed:

		Cl. 11		Cl. 10	
C20	Mboshi	lekoɲi	a log	koɲi	firewood
		lesue	one hair	sue	hair
		lepanda	one scabie's pimple	panda	scabies
		lendugu	a grain of pepper	ndugu	pepper
		lendeli	one hair of a beard	ndeli	a beard
		lengunga	one hair	ngunga	fur

In having a collective meaning, class 10 nouns are semantically similar to singulars. Moreover, they are formally identical to class 9 nouns and, as gender 9, 10 has disappeared, there exist in the languages no more class 10 nouns designating countable objects with a plural meaning. The examples above show clearly that

class 11 might easily have been reinterpreted as an autonomous class with an additive prefix, apt to be preposed to a class 10 noun prefix, then to a class 9 noun prefix, and later to other singular prefixes. If this first hypothesis is correct, it is then imaginable that the additive class 11 noun prefix could have incorporated into its own structure either the vocalic augments of class 10 and 9, or the vocalic noun prefixes proper to other singular classes.

In order to determine the exact origin of the incorporated vowels, one should consider the facts inherent to each particular system, which is not an easy task. According to the languages, the same incorporated vowel can have different origins and it should be recalled that, in zone B, the reduction of noun prefixes to simple vowel markers has created ambiguities which favoured the migration of nouns from one to another class or pairing. For instance, a class 11 noun prefix such as l i – can possibly result from the incorporation of a class 10 or 9 augment, but also from the incorporation of a vocalic prefix belonging either to class 5, 7, 19, or even 10b. A lot of work is thus still to be done in order to check the relevance of the above hypotheses. But these are credible and two observations from the B 20 languages seem to reinforce their plausibility. First, two languages in this group, Sigu and Ndasa S, have multiple allomorphs for class 11 noun prefix, which differ only by their vowel as we see in:

B20	Sigu	11, 6	l é k á á y ì	m á k á á y ì	leaf
		11, 8	l ì n j é d ì	b ì n j é d ì	beard
		11, 10	l í k ú ɲ ì	k ú ú ɲ ì	firewood
B20	Ndasa S	11, 6	l à k á y ì	m à k á y ì	leaf
			l ê l í m ì	m à l í m ì	language
		11, 10	l ê s ù y í	s ù y í	hair
			l ì y ì s í	b ì y ì s í	bone

It should be noted that in Sigu, l i – is also the class 5 prefix. However, in other languages, l ì n j é d ì and l í k ú ɲ ì belong mostly to class 11 and in Sigu, they make their plural in class 8 and 10 respectively, which indicates rather a formal confusion between the noun prefixes of classes 5 and 11 (or perhaps 11 + 7 i –), and not a mere reclassification of the two nouns in class 5. In Ndasa S, the prefix of class 5 is r i –. Prefixes l à – and l ì – are analysed by Jacquot (1983) as class 13 and class 19 respectively, but this analysis seems to be misleading since the reflex *t > l is not observed in the language and since a change *p > l is quite impossible.

The second observation is that those B20 languages which have a class 11 noun prefix with a non-alternating a vowel are also the only ones which have kept tangible traces of class 12 *k à –. In those languages, class 12 merged with class 7 whose prefix is a – before stems with an initial consonant and g y – before stems with an initial vowel. It is clear that the first allomorph belongs historically to class 12, whereas the second one belongs historically to class 7.

Such a coincidence is not likely to have occurred by mere chance. On the contrary the coexistence of an a – prefix in class 7 and a Ca – prefix in class 11 is explainable if these languages were using class 11 as additive with regard to class 12 to express such meanings as 'only one little piece of..' and if this additive use led later to the appearance of a frozen Ca – form. Consistently the Ca – form is not attested in those B20 languages which have completely lost class 12. Furthermore, the existence of a link between class 11 and the merger of classes 7 and 12 seems to be confirmed by the Teke variant described by Fontaney (1984) where the class 7 prefix is alternatively g a – or g e – and the class 11 prefix is l a – or l e –.

4.   Conclusions

So far we have shown that :
1. The structural asymmetry between both terms of a pairing offers the possibility of a reanalysis according to which the CV- prefix appearing initially within one of the two terms is interpreted as the prefix of an autonomous class.
2. An additive prefix (which might have had this status originally or have acquired it by restructuring) can be preposed to a noun whose prefix has maintained its vocalic augment or has itself acquired a V- structure.
3. The additive prefix can then allow incorporation of the following vowel (let it be an augment or a prefix) into its own structure, to the detriment of its original vowel.
4. By a restructuring which occurs in the pairing 11, 10, the class 11 noun prefix can be reanalysed as an additive prefix which is preposed to the class 10 prefix.
5. In languages where the pairing 9, 10 has disappeared or has a tendency to disappear, a class 10 noun associated with a class 11 noun acquires a collective meaning, similar to the meaning of a class 9 singular noun, from which it is no longer formally distinct.
6. The singulative meaning of class 11 agrees easily with the status of an autonomous class, whose prefix is additive with regard to class 10 collective nouns.
7. Many of the languages, which show an aberrant vowel in their class 11 prefix have two allomorphs for their noun prefixes: a V- allomorph which is represented before the nasal prefix or before consonant-initial stems, and a CV- allomorph which is represented before vowel initial stems and, in some languages, before monosyllabic stems.
8. In these languages, the class 11 noun prefix has a constant CV- form and may be the only one which has no V- allomorph.

Synchronically, in the languages where its vowel is aberrant, there seems to be no reason to consider the class 11 noun prefix as an additive one. Nevertheless it is reasonable to suppose that, in a former stage of the evolution, it has been reanalysed as such and preposed first to the prefix of class 10, then to that of class 9, and later to those of other singular classes.

Being mostly used in prevocalic position, it could have retained its CV-structure and could have incorporated the following vowel. This process could have created frozen representations which are characterized by variable vowels and may not be unaccounted by a regular phonological change.

Notes

1. In Zulu, z i – is the regular reflex of *d i –. It cannot b e a reflex the class 8 noun prefix *b ì – since *b > z only if followed by a sequence ..i V.. . So we have: *+b i a d –'to give birth' > +za 1 –, but *+b im– 'to swell' > +v im–.

2. The class 2 noun prefix of Kwangali is noted to be va– by Dammann (1957) but ba– by Westphal (1958).

3. All the cited forms are from Jacquot (1983). They are reproduced as such, even if the word for 'language' in Pove is to be corrected as o l éme (R. Mickala, personal communication). In addition, the alternation between ú and úú in the Sigu words for 'firewood' does not seem to be normal. It will be noticed that in Ngom S, y seems to be one of the reflexes for *d.

References

Blanchon, J.A. 1987. 'Les classes 9, 10 et 11 dans le groupe bantou B40', *Pholia* 2, 5-22, Lyon: Université Lumière-Lyon II.

Blanchon, J.A. 1991. 'Le pounou (B43), le mpongwè (B11 a) et l'hypothèse fortis / lenis', *Pholia* 6, 49-83, Lyon: Université Lumière-Lyon II.

Dammann, E. 1957. *Studien zum Kwangali - Grammatik, Texte, Glossar*, Hamburg: Cram-De Gruyter.

da Silva Maia, A. 1964. *Liçoes de grammatica de Quimbundo (Português e Banto). Dialecto Omumbuim*, Luanda.

Fontaney, V.L. 1984. 'Notes towards a description of Teke (Gabon)', *Pholia* 1, 47-70, Lyon: Université Lumière-Lyon II.

Grégoire, Cl. 1975. *Les locatifs en bantou*, Tervuren: *Annales du Musée royal de l'Afrique Centrale* n° 83.

Herbert, R.K. 1977. 'Prefix restructuring, lexical representation and the Bantu noun', *Studies in African Linguistics*, Supplement 7, 105-111, Los Angeles: University of California.

Hyman, L.M. 1980. 'Reflections on the nasal classes in Bantu', in L.M. Hyman (éd.):*Noun classes in the Grassfields Bantu Borderland*, 179-210,*Southern California Occasional Papers in Linguistics*, n° 8, Los Angeles: University of Southern California.

Jacquot, A. 1983. *Les classes nominales dans les langues bantoues des groupes B10, B20, B30 (Gabon, Congo)*, Paris: ORSTOM.

Meeussen, A.E. 1967. 'Bantu Grammatical Reconstructions', *Africana Linguistica* III, 79-121, Tervuren: Musée royal de l'Afrique Centrale.

Prat, J. 1917. *Grammaire mbochie et dictionnaire. Rivière Alima, Congo français*, Brazzaville: Mission Catholique.

Westphal, E.O.J. 1958. *Kwangari : an index of lexical types*, London: School of Oriental and African Studies.

On the Genesis of Aspect in African Languages:
The Proximative

Bernd Heine
University of Cologne

1. Introduction

Studies carried out in the course of the last decade have revealed a number of regularities in the evolution of linguistic categories in African languages, especially for categories belonging to such domains as tense and aspect. For example, there are essentially only six different sources for progressives (Heine 1994:269). In a similar way, other aspectual notions, such as perfect, perfective, or habitual, each have a limited range of conceptual sources (cf. Heine et al. 1993; Bybee, Perkins & Pagliuca 1994). Observations like these allow us to propose a number of probabilistic predictions (where "prediction" not only refers to conceivable future situations but also to unknown synchronic states, as well as to reconstructed past events), such as the following: (i) If we find a progressive construction in a given African language, then most likely it is derived from a locative or comitative schema; (ii) If we find a perfect construction then it is likely to involve a verb meaning 'finish' or 'end'.

In the present paper, an aspectual notion is looked at that does not figure in standard grammars of African languages although in many of these languages there exists a conventionalized expression for it.

2. The Proximative

In a number of African languages there is a grammatical construction involving a marker that is "homophonous" with a verb meaning 'want', 'desire', 'seek', or 'look for'. The following are a few examples.

(1) Swahili (Bantu, Niger-Congo)

   a. *Ni-li-taka   ku-m- piga.*
       I-PAST-want   INF-him-hit
     'I wanted to hit him.'

   b. *Mvua   i-li-taka   ku-nyesha.*
       rain   it-PAST-want INF-rain
     'It was about to rain.'

(2) Ewe (Kwa, Niger-Congo; Ameka 1990:145)

   a. *Kofí   dí   bé   ye-a   kpɔ́ wò.*
       Kofi   want that   LOG-IRR see you
     'Kofi wants to see you.'

b. *tsi   dí   bé   ye- a   dza.*
   water   want   that   LOG-IRR   fall
   'It is about to rain.' (Lit.: 'Water wants to fall.')

(3) Chamus (Maa, Nilo-Saharan; Heine 1992)

  a. *k-e-yyéú   m-partút.*
     k-3.SG-want   F-woman
     'He wants a woman/wife.'

  b. *k-é-yyéú   l-cáni   n-éuróri.*
     k-3.SG-want M-tree   NAR-fall
     'The tree almost fell.'

While in the (a) examples we are dealing with a verb of volition, the 'homonymous' counterpart in the (b) sentences expresses a grammatical function. This function is described in various ways by the authors concerned. In most works on languages having a similar distinction, the meaning of the corresponding item in the (b) sentences is rendered as either 'nearly' or 'almost'. Doke (1930:212) observes that the Zulu 'deficient verb' *-cishe* is used of past time to indicate 'being nearly but never quite', while another deficient Zulu verb, *-funa*, indicates 'to be about', 'to be on the point of'. Similarly, Cole notes that the Tswana verb *-batla* 'want, look for, desire', when used deficiently, expresses 'almost', 'nearly', or 'on the point of, but never quite doing' (Cole 1955:292), and Brown (1875:17) says that the Tswana verb *-batla* 'want', when 'used adverbially,' expresses 'the nearness to completion of an action, hence, nearly; almost; about.' In more general terms, the function concerned may be said **to define a temporal phase located close to the initial boundary of the situation described by the main verb**. We are dealing essentially with an aspectual notion, one that is referred to as the ALMOST-aspect by Heine (1992:339). König (1993:85) proposes the term proximative instead and I will adopt this term here. In the present paper, some regularities to be observed in the behavior of proximative patterns in African languages are highlighted.

3. Models

          How is this case of 'homonymy' between verbs of volition and proximative markers to be accounted for? There is an obvious answer: The two are diachronically related in that the proximative marker is historically derived from a verb meaning 'want, desire'. The process is one of auxiliation: The verb comes to serve as an auxiliary and acquires a grammatical function. But this does not explain the process.
          One way of accounting for the process is by using a metaphor model: The shift from a proposition like (4a) below to (4b) can be described as being metaphorically structured, for the following reasons. First, one might say that the propositional meaning *X wanted to do Y* provides a structural template, or a vehicle, for expressing the meaning *X nearly did Y*. Second, the shift can be understood as involving a transfer between two domains of conceptualization: (4a) is suggestive of a domain of willful human beings who act, and who have wishes: The subject in (4a) is typically human. In (4b) on the other hand, the subject is

inanimate and the proposition implies a domain where human categories are irrelevant. One may thus interpret the process involved as an instance of what Lakoff and Johnson (1980:134) refer to as personification, or Heine, Claudi and Hünnemeyer (1991:169ff.) as the Person-to-Object metaphor. Third, metaphor is frequently described as 'deviant behavior' which conflicts with our expectations and has to do with a violation of existing rules, or that metaphor involves a statement that, if taken literally, is false (Heine, Claudi & Hünnemeyer 1991:207-8). If we assume that a proposition like (4a) provided the template for developing a structure like (4b) then one may say that the latter is a 'literally false' proposition since the concept 'want' is not normally compatible with inanimate subjects. Fourth, metaphor typically has to do with a transfer from concrete to less concrete expressions. Now, it is obvious that a grammatical function as expressed by *wants* in (4b) is more abstract than the lexical use of *want* in (4a). To summarize, there are good reasons to argue that the metaphor model is adequate to account for the shift from (4a) to (4b).

(4) a Person X wanted to do Y
    b ('Object X wanted to do Y' >) Object X was about to enter situation Y

There are, however, also observations that suggest that the metaphor model does not take care of all the facts that have to be considered. One instance of a proximative marker has been discussed in Heine (1992), where this marker is described as being part of a grammaticalization chain and, in fact, the use patterns of proximative constructions in other languages as well are suggestive of a chain-like structure or a continuum, rather than of discrete and segmentable categories. Along this continuum, a number of salient points or stages can be distinguished. I will now briefly look at these stages in turn.

At Stage 0, there is a verb of volition where the subject is a human being and the object some concrete item, as in (5).

(5) Southern Sotho (Bantu, Niger-Congo; Doke & Mofokeng 1957:292; Guma 1971:188)

Ke-batla nama.
'I want meat.'

At Stage I, instead of a concrete item, the object refers to a dynamic situation, typically encoded as a verb in a nominalized/infinitival form, as in (6).

(6) Swahili (Bantu, Niger-Congo)

Wa-na-taka   ku- lima.
3.PL-PRES-want   INF-farm
'They want to farm.'

In the following sentences of (7), a different situation emerges.

(7) Swahili (Bantu, Niger-Congo; Hassan Adam, p.c.)

Ni-li-taka   ku-fa.
I- PAST-want   INF-die
'I nearly died; I narrowly escaped death.'

Tswana (Bantu, Niger-Congo; Cole 1955:292-3)

Kenê ka-batla goikgaola monwana kathipa.
'I nearly cut off my finger with a knife.'

Chamus (Maa, Nilo-Saharan; Heine 1992:338)

k-é-yyéú   l-páyyan   n-é-rriá.
k-3.SG.-want   M-elder   NAR-3.SG-fall
'The old man nearly fell.'

The item *-batla* in (7) is a manifestation of the lexical verb 'want' in Sotho and the closely related Tswana language.
    The sentences in (7) are compatible with both (4a) and (4b): They have a human subject and, hence, imply a world of typically human experiences; at the same time, the content of the sentences is at variance with this interpretation: One does not normally expect, as a native speaker of English, Tswana, or Chamus, that dying or falling are situations that one really wants to happen. Thus, (7) can be interpreted both with reference to the lexical and the aspectual meaning of 'want'. The Chamus sentence in (7), for example, can be interpreted alternatively as meaning 'The old man wanted to fall' or 'The old man nearly fell'. What this suggests is that these sentences are ambiguous. I will refer to such situations as Stage II-situations, in contrast with Stage I-situations, where 'want' functions exclusively as a verb of volition, being associated with meanings such as 'want', 'desire', 'search', 'seek', or 'look for'.
    A different stage is reached when instead of a human subject, an inanimate one is used. In this case, the lexical meaning of volition is ruled out and the aspectual meaning survives. This stage, which I will call the Stage III-situation, is exemplified in (8).

(8) Swahili (Bantu, Niger-Congo; Hassan Adam, p.c.)

Gari   li-li-taka   ku-pinduka.
car   it-PAST-want INF-overturn
'The car nearly overturned.'

Chamus (Eastern Nilotic, Nilo-Saharan; Heine 1992)

k-é-yyéú   l-cáni   n-é-uróri
k-3.SG-want   M-tree.NOMIN   NAR-3.SG-fall
'The tree almost fell.' (Lit.: 'The tree wanted to fall.')

    A new stage, called Stage IV, is reached when instead of inanimate subjects, animate ones may be used, as in (9). What distinguishes this stage from Stage II is the fact that a lexical interpretation is no longer possible, that is, in spite

of the fact that (9) has a human subject, a volitional meaning appears to be ruled out.

(9) Zulu (Bantu, Niger-Congo; Doke 1930:203, 212)

Ba:cishe bangibulala.
  .want.
'They nearly killed me.'

West African Pidgin English (Agheyisi 1971:144, 149)

Tif	*wan*	*tek*	*wi*	*moni.*
thief	want	take	our	money

'A thief is/was about to take our money.'

As we will see below, this stage tends to be associated with a reanalysis of the morphosyntactic structure involved.
    Another stage, let us call it Stage V, is reached when the complement of the erstwhile verb of volition refers to a quality or quantity, rather than to a dynamic situation, as in (10).

(10) West African Pidgin English (Agheyisi 1971:144, 149)

Dis	*pikin*	*wan*	*tol*	*pas*	*dat*	*wan.*
this	child	want	tall	surpass	that	one

'This child is almost taller than that one.'

Southern Sotho (Bantu, Niger-Congo; Doke & Mofokeng 1957:292; Guma 1971:188)

khomʊ tse-batlang hʊ-ba lɪshʊmɛ
            -want.
'about ten cattle'

    Finally, at Stage VI, the use of the erstwhile verb of volition refers to past events, where the meaning 'having almost reached the situation described by the main verb' triggers a context-induced reinterpretation to the effect that that situation is actually not reached, that is, where the proximative marker primarily expresses negation, as in the following example.

(11) Chamus (Eastern Nilotic, Nilo-Saharan; Heine 1992:340)

i-túm-o	*m-partút?*	*(k)éyyeu*	*a-túm.*
2.SG-get-PFV	F-woman	almost	1.SG-get

'Did you get a wife?'   'No (but I almost did).'

    To summarize, rather than with a discrete jump from a lexical to a grammatical/schematic semantics, we are dealing with a continuous transition. The main semantic stages that may be distinguished in the course of this transition are sketched in Table 1.

Table 1. Some stages in the semantic transition from volition to proximative.

Stage	Type of event schema	Contextual attributes
0	Person X wants item Y	[Y refers to a concrete item]
I	Person X wants to do Y	[Y refers to a dynamic situation]
II	Person X is about to undergo Y	[X can be assumed not to want Y]
III	Object X is about to become Y	[X is close to entering situation Y]
IV	Object/Person X is about to become Y	[Instead of an inanimate X, X may now be human]
V	X is only a little less than Y	[Instead of a dynamic situation Y is a quantity or quality]
VI	X has not attained state Y	[X has not reached situation Y]

4. Grammaticalization

In Swati (Swazi), the proximative items -*cishe* and -*phoshe*, which appear to be used interchangeably and whose meaning is rendered as 'nearly', are described as 'verbs with an infinitive complement,' even though they 'have no corresponding independent verbs' (Ziervogel & Mabuza 1985:151, 158). The proximative markers -*cishe* and -*funa* of Zulu (see 2 above) are called 'deficient verbs followed by the infinitive' by Doke (1930:211-12). In Venda, the proximative marker -*thodhou*- ('almost, nearly') has developed into a verbal aspect prefix (Poulos 1990:354).

The problems that students of African languages have when dealing with the categorial status of proximative markers are reflected in inconsistent taxonomic treatments of these markers. The proximative marker -*cishe* of Zulu, for example, is described by Doke (1930:211-12) as a verb but by Taljaard and Bosch (1988:160) as an auxiliary. The morphosyntactic process involved in the development of proximatives is essentially the same as that observed in other kinds of auxiliation (see Heine 1993:55ff.), it involves on the one hand the decategorialization of the verb 'want, desire', 'choose', 'look for', or 'seek', and on the other hand the shift from a verbal morphosyntax to that of a verbal affix. The main stages to be distinguished are summarized in Table 2.

Table 2. Some stages in the morphosyntactic transition from volition to
proximative

0   The verb exhibits a fully verbal morphosyntax and the complement has a
    noun phrase or a subordinate clause as its nucleus.

I   Instead of a noun, the complement consists of a nominalized/non-finite
    verb.

II The verb loses verbal properties, and the nominalized verb in nominal
    properties.

III The verb becomes a clitic and loses the ability to be negated separately,
    and the complement loses nominal properties.

IV The verb loses virtually all remaining verbal properties and becomes a
    verbal affix, and the complement acquires the morphosyntax of a main
    verb.

This scenario is a simplified account of what is likely to happen (for example, I
have conflated Heine's parameters of decategorialization and cliticization; for a more
detailed treatment, see Heine 1993). A few examples may suffice to illustrate the
process concerned. An example of Stage II is found e.g. in Chamus, where the
verb *-yyéú* 'want' became an invariable grammatical marker when used as a
proximative marker: It may no longer be inflected for person and number.
        An example of Stage III, which tends to coincide with the semantic Stage IV
(see above), is also provided by the Chamus verb form *k-é-yyéú* 's/he wants',
which has given rise to a clitic *(k-)eyyéú* 'nearly, almost', as illustrated in (12).

(12) Chamus (Maa, Nilo-Saharan; Heine 1992:339)

    (k-)eyyéú   a-ók   nánʋ kʋlé.
    PROX   INF-drink I   milk
    'I was about to drink milk.'

        Stage IV appears to have been reached in Venda, where the verb *thodha* (the
graphs *th* and *dh* stand for dental stops) 'want' has been grammaticalized to a verbal
prefix *-thodhou-* (Poulos 1990:332). I will return to the morphosyntactic process
concerned below. Poulos observes, for example, that, while both verbs and
auxiliaries are followed by 'full complements,' this is not so in the case of the
prefix *-thodhou-*.
        Finally, the grammaticalization of the erstwhile verb to a grammatical
element also involves an erosion of the phonetic substance employed for the
expression of the proximative. Part of this process is what Heine and Reh (1984)
call adaptation, that is, the phonetic assimilation of the verb to its complement. We
may illustrate this process with the following example from Venda. Compare the
following examples:

(13) Venda (Bantu, Niger-Congo; Poulos 1990:332)

a. *Ndo   thodha   u   mu   rwa .*   b. *Ndo thodhou mu rwa.*
   I   want   INF   him   hit   I   almost him hit
   'I wanted to hit him.'   'I nearly hit him.'

Example (13a) represents Stage II of both the semantic and the morphosyntactic chains: the verb *thodha* exhibits a largely verbal behavior and the complement is an infinitival (INF) verb having an object prefix. (13b) is characteristic of Stage IV of the morphosyntactic chain: *-thodhou-* is a verbal prefix (even if Venda orthography treats it as a separate word). The change from *thodha + u* to *thodhou* is due to the merger of the verb *thodha* 'want' with the infinitive marker into a grammatical item, whereby the infinitive marker *u-* became part of the preceding item, whose final vowel *a* assimilated in vowel height to *o*. Thus, a kind of 'reanalysis' or 'boundary shift' took place, in that the infinitive marker was detached from the complement and became part of the erstwhile main verb, which now has the status of a verbal prefix.

5. Proximative vs. Future

The proximative exhibits a number of structural similarities with another grammatical category, the future tense, to the extent that the two are considered the same by some authors. The similarities are in particular the following:

(a) Both correspond to the notion of 'irrealis' markers: They refer to situations that may be expected to occur but have not yet occurred (Givón 1984:285). The following example from Zulu, involving the 'deficient verb' *-cishe* ('nearly') illustrates this property of proximatives.

(14) Zulu (Bantu, Niger-Congo; Doke 1930:203)

Wa:-cishe wafa, wabuye wabangcono.
'He was on the point of death when he recovered.'

(b) Not infrequently, both are similar in form. In Maasai, for example, both have the form *-un* (König 1993).

(c) The fact that the two are similar in form can be explained with reference to their diachronic development. Thus, both the proximative and the future markers of Swahili are historically derived from the verb *-taka* 'want,' and both the proximative and the future functions of the Maasai suffix *-un* can be traced back to the verb **-buon* 'come' (Heine & Claudi 1986:72).

(d) As these observations indicate, proximative and future markers are derived from the same pool of source concepts, i.e., from verbs of volition and verbs of motion.

(e) Furthermore, the grammaticalization of both proximative and future markers involves the same general process, which can be sketched like this: Originally, the verb of volition is confined to human subjects. In the course of grammaticalization, the use of the verb is extended to inanimate subjects and the verb assumes the function of an auxiliary while the verbal complement turns into the main verb.

Nevertheless, there are also a number of differences, such as the following:

(a) The proximative is an aspect while future is a tense. This means, for example, that proximatives are not sensitive to deictic time: they may refer to past, present, or future situations, as example (15) from Zulu, involving the 'deficient verb' -*cishe* ('nearly'), indicates.

(15) Zulu (Bantu, Niger-Congo; Doke 1930:203)

Nga:cishe ngawa .	'I nearly fell.'
Ngicishe ngiwe .	'I almost fall.'
Ngiyokucishe ngifike.	'I shall be on the point of arriving.'

Future tenses on the other hand, are incompatible with present or past situations.

(b) While futures have to do with prediction, in that the speaker predicts that the situation in the proposition will hold (Bybee, Pagliuca & Perkins 1991:54), proximatives do not involve prediction.

(c) While the cognitive domains giving rise to both kinds of categories are essentially the same, the salience patterns are different: Proximative markers are mostly derived from verbs of volition, whereas future tenses are more likely to have verbs of motion as their source.

The following question arises in this connection: Why does one and the same source, say, a verb of volition, give rise to two contrasting grammatical notions? One way of looking for an answer is by reference to the metaphor model mentioned earlier: One might say that one and the same event schema (*X wants to do Y*) may serve as a vehicle for two contrasting metaphors, roughly as conceived in (16).

(16)	X wants	> (a) X is about to do Y	[Proximative]
	to do Y	(b) X can be predicted to do Y	[Future]

This raises the following question: Is it possible to determine under what conditions (16a), as opposed to (16b), is chosen? For example, what was responsible the development of English *will* into a future marker rather than a proximative marker?

It would seem that an answer is not possible without reference to the contextual background in which new grammatical meanings arise. As the work of Bybee, Pagliuca and Perkins (1991) suggests, the development from volition to future involves an intermediate stage where volition (e.g. *X wants to do Y*) is reinterpreted as denoting intention (*X intends to do Y*). Thus, this development requires that 'want' be used in contexts that induce an inferential reasoning of the following kind: *Since X wants to do Y s/he also has the intention to do it.* Note that such an intermediate stage is not confined to the Volition Schema (Heine 1993), that is, to the development from volition to future; it can also be observed in the case of the Motion Schema, that is, when the directional verbs 'go to' or 'come to' are involved (cf. Bybee, Pagliuca & Perkins 1991; Pérez 1990).

The rise of proximative markers on the other hand, appears to involve an inferential mechanism that has to do with what I referred to earlier as Stage

II-situations: The verb of volition is employed in contexts which are not in accordance with normal expectations of volition: If someone says: 'The old man "wants" to die', then one reasonable inference would be that 'The old man is close to death'. Such an inference is even more obvious when the subject is inanimate, e.g., when I say 'The tree *wants* to fall', where perhaps the most immediate interpretation would be that the tree is about to fall. To summarize, it is the specific contextual frame and the expectations associated with it that appear to be responsible for the fact that the Volition Schema gives rise to proximative constructions in some cases but to futures in others.

6. Conclusions

What this would seem to suggest is that the metaphor model provides a useful descriptive device but does not explain the process concerned. In addition, another model is required, one that I have referred to elsewhere as the context model. This model accounts for conceptual reinterpretation with reference to a mechanism of contextual expansion, roughly as outlined above.

The dynamics of the process from verb to proximative marker is manifested in the fact that in some languages, the process has happened repeatedly. Thus, we observed that in Tswana, three different verbs, *-batla* 'want, look for, desire', *-rata* 'love, like, want', and *-senka* 'seek, look for', are used for this purpose. Similarly, in Zulu the verb *-cish-* was grammaticalized to a proximative category: The 'deficient verb' *-cishe* expresses the aspectual notion 'to be on the point of doing, but never quite doing' (Doke 1930:203, 212). At the same time, the verb *-funa* 'want, look for' has also entered the pathway of becoming a proximative marker: While its behavior is still more verb-like than that of *-cishe*, its meaning is already largely that of a grammatical marker ('to be about to, to be on the point of').

As I observed in the introduction with reference to other aspectual notions, there is usually a pool of different conceptual sources that serve as structural templates for a given grammatical category. The same applies to the proximative. In addition to verbs of volition, proximatives may be derived from verbs of motion, and perhaps an even more common source is provided by a locative schema of the form *X is near to Y*. At the present stage of research, no generalizations on the various conceptual sources seem possible. What is obvious, however, is that it is the same general range of propositional structures employed for other aspectual categories that is also recruited for the development of proximatives.

Abbreviations

ABS	absolute case	PL	plural
F	feminine gender	PRES	present tense
INF	infinitive marker	PROX	proximative
IRR	irrealis marker	SG	singular
LOG	logophoric pronoun	1	first person
M	masculine gender	2	second person
NAR	narrative marker	3	third person
NOMIN	nominative case		
PAST	past tense		
PFV	perfective		

References

Agheyisi, Rebecca Nogieru. 1971. *West African Pidgin English: Simplification and simplicity*. Ph.D. Dissertation, Stanford. Ann Arbor, Michigan: University Microfilms.

Ameka, Felix. 1990. The grammatical packaging of experiencers in Ewe: A study in the semantics of syntax. *Australian Journal of Linguistics* 10:139-81.

Brown, J. Tom. 1875 [1988]. *Setswana dictionary: Setswana - English and English - Setswana*. Braamfontein: Pula Press.

Bybee, Joan L., William Pagliuca & Revere D. Perkins. 1991. Back to the future. In Traugott & Heine 1991b. Pp. 17-58.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Cole, Desmond T. 1955. *An introduction to Tswana grammar*. Cape Town: Longman Penguin South Africa.

Doke, Clement M. 1930. *Textbook of Zulu grammar*. Cape Town: Maskew Miller Longman.

Doke, Clement M. & S. M. Mofokeng. 1957. *Textbook of Southern Sotho grammar*. Cape Town: Maskew Miller Longman.

Givón, T. 1984. *Syntax: A functional-typological introduction*. Volume 1. Amsterdam, Philadelphia: John Benjamins.

Guma, S. M. 1971. *An outline of Southern Sotho*. Pietermaritzburg: Shuter & Shooter.

Heine, Bernd. 1992. Grammaticalization chains. *Studies in Language* 16.2:335-68.

----- 1993. *Auxiliaries: Cognitive sources and grammaticalization*. New York, Oxford: Oxford University Press.

----- 1994. Grammaticalization as an explanatory parameter. In Pagliuca, William (ed.), *Perspectives of grammaticalization*. Amsterdam, Philadelphia: John Benjamins.

Heine, Bernd & Ulrike Claudi. 1986. *On the rise of grammatical categories: Some examples from Maa*. (Kölner Beiträge zur Afrikanistik, 13.) Berlin: Dietrich Reimer.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framewok*. Chicago: University of Chicago Press.

Heine, Bernd, Tom Gueldemann, Christa Kilian-Hatz, Donald A. Lessau, Heinz Roberg, Mathias Schladt, and Thomas Stolz. 1993. *Conceptual shift: A lexicon of grammaticalization processes in African languages* (Special Issue of AAP, 34/35). Cologne: University of Cologne.

Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske Verlag.

Hopper, Paul & Elizabeth C. Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

König, Christa. 1993. *Aspekt im Maa*. (AMO, Afrikanistische Monographien, 3.) Cologne: University of Cologne.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Pérez, Aveline. 1990. Time in motion: Grammaticalization of the *be going to* construction in English. *La Trobe University Working Papers in Linguistics* 3:49-64.

Poulos, George. 1990. *A linguistic analysis of Venda.* Pretoria: Via Africa Limited. Maskew Miller Longman.

Taljaard, P. C. & S. E. Bosch . 1958? *Handbook of Isizulu.* Pretoria: J. L. van Schalk.

Traugott, Elizabeth C. & Bernd Heine. 1991a, b. *Approaches to grammaticalization.* (Typological Studies in Language, 19,1; 19,2.) Amsterdam, Philadelphia: John Benjamins.

Ziervogel, D. & E. J. Mabuza. 1985. *A grammar of the Swati language (siSwati).* Second impression. Pretoria: J. L. van Schalk.

A Neglected Ethiopian Contribution to Semitic and Afroasiatic Reconstruction

Grover Hudson
Michigan State University

0. Introduction

In section 1 of this paper I will describe a prominent feature of Ethiopian Semitic languages, their lexical classes of verb roots, termed A-type and B-type, which make up the large majority of verb roots of the languages. Perhaps more than any other features of Semitic and Afroasiatic languages, this pair of verb classes appears to present evidence for two somewhat controversial features of the reconstructed Semitic and Afroasiatic verbal system: a formally expressed distinction of active or transitive verbs versus stative or intransitive verbs, and a present/imperfect verb stem marked by consonant length. In section 2 I will briefly review the basic evidence for these two features, and show how, nevertheless, because the Ethiopian Semitic (henceforward ES) B-type has been considered to be a development of the Semitic derived intensive formation seen in Arabic, Afroasiatic and Semitic, reconstruction has proceeded without appreciation of the Ethiopian Semitic A/B-type dichotomy. Finally, in the third section, I will suggest reasons why this very prominent ES characteristic has been ignored, and will show some ways that it can contribute importantly to reconstruction.

1. A and B-type in ES

Following are 3m.sg. forms of a pair of typical A and B-type triconsonantal verbs, in the past, present, and jussive conjugations in three ES languages: Ge'ez, Amharic, and Harari.

(1) A and B-type verbs in Ethiopian Semitic (3m.sg. forms)

Language	Root		Type	Past	Present	Jussive
Ge'ez	*sbr*	'break'	A	*səbərə*	*yɨ-səbbɨr*	*yɨ-sbər*
	mzn	'weigh'	B	*məzzənə*	*yɨ-mezzɨn*	*yɨ-mezzɨn*
Amharic	*sbr*	'break'	A	*səbbərə*	*yɨ-səbr*	*yɨ-sbər*
	mzn	'weigh'	B	*məzzənə*	*yɨ-məzzɨn*	*yɨ-məzzɨn*
Harari	*sbr*	'break'	A	*səbəra*	*yi-səbri*	*yə-sbər*
	mzn	'weigh'	B	*meezəna*	*yi-miizni*	*yə-meezni*

The A and B-type verb classes appear in all the ES languages (eleven to thirteen, depending on how dialect continua within the so-called Gurage languages are counted). Two formal characteristics distinguish the types: a front vowel *e* (in Harari *ee >ii*) after the first root consonant of some verb forms, and lengthening of the second consonant in all forms of languages that have long consonants. The vowel characteristic appears in only some forms of some ES languages (for example, Ge'ez and Chaha), appears uniformly in some (including Tigrinya and Harari), and is uniformly absent in others (including Amharic and Tigre). The characteristic appears to have been original in the present (or imperfect) stems, from

which it was extended to past (or perfect) forms in some of the languages (Hetzron 1972: 23-26). The long consonant characteristic appears throughout ES except in those languages which have undergone a sound change of degemination. (Since Ge'ez has been equated with Proto-ES (Ullendorff 1960: 129), it should be mentioned that this language, while the earliest recorded ES language, by a thousand years, has apparently no descendant among the modern languages, and deserves only slight priority as evidence in reconstruction (Hetzron 1972: 19-21).)

Notice in (1) that A and B types are merged in the Amharic perfect, in which roots of both types form stems of the pattern CVC:VC, without the front vowel and with a long second C. In the other forms and in the perfect of most of the other languages, the A/B-type dichotomy is prominent throughout ES, and always there is no apparent correlation of meaning with the formal distinction of types, which is thus a strictly <u>lexical</u> distinction --that is, one which must be marked in the lexicon for each simple triconsonantal verb root (by far the most common root-type in ES as throughout Semitic), and also for biconsonantal roots with triconsonantal etymologies in which the second root consonant survives. Being fully present and prominent throughout ES, the lexical characteristic of A and B types certainly must be reconstructed for the group.

The absence of a meaning correlate of the B-type is not recognized or admitted by earlier grammars of ES languages. Dillmann (1907: 143-6) equated the Ge'ez B-type to the Arabic derived so-called form 2, generally an intensive or causative. That there are almost no Ge'ez basic (A-type) verbs also forming B-types Dillman attributed to 'frugality displayed in the housekeeping of forms'. Conti Rossini (1941 [1967: 40]), in the Ge'ez grammar perhaps most commonly used as a textbook for the language, says the B-type indicates 'intensity of action' and the transitive of an A-type, and he mentions some of the very rare cases for which this seems so, including root *msl* A-type 'seem' and B-type 'compare', and root *mtr* A-type 'amputate' and B-type 'amputate completely'. Concerning the A/B-type dichotomy in Amharic, Cohen (1936: 201) says 'the ancient intensive value of the B type is reflected in diverse facts': A-types are 'generally active and "neuter"', while B-types are 'generally active', and he too mentions some of the rare roots which form both types, such as *t'bq* , with A-type 'be tight' and B-type 'guard, protect'. Perhaps Leslau (1936: 93) was first to correctly identify the B-type throughout ES as a lexical variant without regular semantic value; but he too considered it historically a derived form, cognate with Arabic form 2.

2. Semitic and Afroasiatic Reconstruction

Following is a presentation of Arabic and Akkadian comparisons with the data of (1), followed by a survey, in broad terms, of present reconstruction of the Proto-Semitic and Afroasiatic verbal systems.

2.1. Arabic

The structure of classical or standard Arabic is so relatively regular and elegant that early in the history of comparative linguistics this Semitic language strongly influenced or even determined the European conception of the Semitic language type. Fifty years ago Goetze (1942: 2) wrote that:

'In earlier days of Semitic research the belief was universal that Arabic, in which both vowels and consonants are so well preserved, is the most archaic among the

Semitic languages and particularly close to the common mother tongue. In such circumstances it was a foregone conclusion that not only the individual forms ...but also the general scheme of which they form part was inherited from Primitive Semitic'.

This tendency, if it persists, is not so strong in Semitic historical linguistics today, and certainly the trend has long been to broaden the basis of comparison within Semitic to include ES and, more recently, even, to consider non-Semitic Afroasiatic evidence.

The essential Arabic comparisons with (1) are presented in (2), for active and stative root 3m.sg. forms of the basic (form 1) suffixed past and prefixed present, and form 2 also in the past and present.

(2) Arabic basic verb morphology (3m.sg. forms)

Root	Form 1 perfect	Form 1 imperfect	Form 2 perfect	Form 2 imperfect
ktb 'write'	*katab-a*	*ya-ktub-u*	*kattib-a*	*yu-kattib-u*
krm 'be generous'	*karum-a*	*ya-krum-u*	*karrim-a*	*yu-karrim-u*

Notice the characteristic long consonant of form 2 stems.

Form 1 is basic and form 2 is considered derived. Often termed an intensive, as in *kassara* 'he smashed' from *kasara* 'he broke', form 2 as often yields a simple causative, as in *ʕallama* 'he taught' from *ʕalama* 'he learned'. In fact, the tendency is slight for a productive regular derivative to be expressed by the form (Goetze 1942: 2-3).

Formally speaking, the ES perfect is a good comparison with the Arabic basic perfect, with stem of form CVCVC, and the ES jussive with the Arabic present of form *y* V-CCVC. The ES B-type has always been compared with Arabic form 2 (Hebrew *piel*), both with imperfect form *y* V-CVC:VC. I have mentioned how earlier ES grammarians have supposed the B-type to be a derived intensive or development of a derived intensive. It seems apparent that this notion is the result entirely of an attempt to make sense of the ES facts in terms of the Arabic (or Hebrew) model.

But an opposite interpretation would seem just as reasonable --and moreso, perhaps, if the crusades had been fought in Ethiopia instead of the Arab Mediterranean world. That is, Arabic form 2 may be understood as a development of a lexical type still prominent in ES, reduced by analogical leveling in favor of the more basic type, and reinterpreted, presumably on the basis of a resulting semantic tendency or on the basis of a few prominent verbs in which the relation could be seen as derived, often as an intensive, from the basic type.

2.2. Akkadian

When Akkadian came to European awareness as the result of archaeological discoveries and the eventual decipherment of its cuneiform writing system after about 1875, this Semitic language would at first have little initial effect on the Arabicized conception already fixed. Though undeniably old (documentation is extensive from as early as 4000 years ago), Akkadian must at first have been accorded minimal consideration as evidence for Proto-Semitic because of, first, the

established belief, based upon the Bible stories so influential then and reinforced by the Akkadian documents themselves, that Assyria-Babylonia had been a land of extreme multilingualism, 'a great mixture of races and languages [where] it was accordingly that the scene of the confusion of tongues was laid' (Sayce 1880: 3), and second, the fact that its writing system was not Semitic but Sumerian.

The verbal system of Akkadian showed an imperfect set of comparisons with Arabic, as in (3) (from Rössler 1950 [1981: 696-98]), which shows 3m.sg. forms of one root in the so-called stative (with subject suffixes in other persons), the prefixed present, the prefixed preterite, and the prefixed present intensive 'D-form'.

(3) Akkadian basic verb morphology (3m.sg. forms)

Root	Stative	Present	Preterite	Present intensive
kšd 'conquer'	*kašid*	*i-kaššad*	*i-kšud*	*u-kaššad*

Here, a number of details must be omitted concerning such barely sufficient examples. But the essential formal comparisons are clear, up to a point: the ES A-type perfect and Arabic perfect with the Akkadian stative (CVCVC); ES A-type jussive and Arabic present with Akkadian preterite (*(y)*V-CCVC); the ES present with the Akkadian present, with their characteristic long consonants (*(y)*V-CVC:VC); and the Akkadian intensive with the Arabic intensive. Both the simple present stem and the intensive show the long root-second consonant.

An important point not apparent in (3) is that 'the stative remains outside any system that makes action and tense its basic categories' (Goetze 1942: 4). While an Akkadian root may appear in both the stative and preterite, like *kšd* 'conquer' in (3), this is often not the case (Goetze 1942: 4; Rössler 1950 [1981]: 698). It is convenient in (3) that a single root exemplify the forms, as typically and equally misleadingly often in the grammars, but this result is untypical. Stative roots have stative forms in preference to present and preterite forms (as seems reasonable, since a state perfected in the past is also present), whereas active roots have both the prefixed present and preterite (both are needed since an action in the past, being punctual, is not present). The Akkadian stative appears to be an essentially nominal conjugation, with its pronominal suffixes perhaps originally a nominal sentence (Buccellati 1968, 1988, Huehnergard 1987).

When Akkadian cuneiform was barely deciphered and Akkadian grammar barely understood, already Haupt (1878) rejected the conventional wisdom of his time 'that Assyrian is ... a corrupt branch of the Semitic family of speech' (245). Haupt claimed priority of age for the Akkadian-type verbal system over the Arabic: 'The Assyrian [Akkadian] Present (iqátal) and the Ethiopic Imperfect (yeqátel) are no new formations, but the oldest verbal forms of the Semitic family of speech' (246). Notice that Haupt was unaware, even, of the long second consonant shared by Ge'ez and Akkadian in this form. His argument, it seems, was not widely accepted (cf. Moscati, et al.: 1964) and, generally, the issue lay dormant until the fifties, when Rössler (1950) presented his reconstruction based on a survey of Afroasiatic as well as Semitic verbal morphology, again attributing shared archaicness to Akkadian and Ge'ez. Rössler emphasized (i) the formal identity of Akkadian and Ge'ez presents with long second consonant, (ii) the two 'past' conjugations of Akkadian, prefixed and suffixed, which represent conjugations of different roots, so that this system shows a lexical dichotomy not present in Arabic

and not derivable from it, and (iii) apparent cognates of this dichotomy in South Arabian, Berber, and Cushitic languages. Later, Greenberg (1952) too argued for the Afroasiatic basis and historical identity of the Akkadian and ES A-type present of form *(y)*V-CVC:VC, and of the Akkadian preterite, Arabic present, and Ge'ez jussive of form *(y)*V-CCVC. Leslau (1953) countered-argued, noting that gemination is typically absent in the A-type of Amharic and other southern ES languages, and might be secondary elsewhere in ES. (Hetzron (1972: 23) reviews the ES evidence and arguments.)

2.3. Ethiopian Semitic

Even more than Akkadian, Ethiopian Semitic languages other than Ge'ez were latecomers to the consciousness of European Semiticists. Ge'ez, indeed, was known to Semiticists before Akkadian, but the early standard Ge'ez grammar of Dillmann (1857 [1907: 3]) suggested and probably fixed in the minds of generations of European Semiticists the notion that migrations of Semites from Yemen had brought Semitic languages to Ethiopia, a speculation which became thoroughly established when Sabaen South Arabian inscriptions in northern Ethiopia were published in the 1920s. As Akkadian was assumed to be corrupted by Sumerian, so the ES languages were assumed to be corrupted by Cushitic languages. (Indeed, while Cushitic influence may be apparent in ES syntax and lexicon, evidence of morphological influence is less certain, and to interpret seeming Cushiticisms in ES as representing persistence of Afroasiatic features, in both groups, seems to me often as reasonable.)

The essential facts of ES verbal morphology presented in (1) and comparable with (2) and (3) lack a semantic parallel to the Arabic and Akkadian so-called intensives. In fact, there is a reduplicative derived intensive verb in ES formed in triconsonantals by copying the medial consonant of the root forward, followed by the vowel *a* (though the formation is not known in Ge'ez), as in the following Amharic examples:

(4) Amharic intensive (3m.sg. forms)

Root	Type	Past	Present
sbr 'break'	A	*səbabbər-ə*	*yɨ-səbabbɨr*
flg 'seek'	B	*fəlalləg-ə*	*yɨ-fəlallɨg*

The meanings of the forms of (4) are 'smash' from 'break' and 'seek here and there' from 'seek(/want)'. The form is sometimes called a 'frequentative' (Leslau 1936: 111). Such a universally expected sort of derivation of an intensive by reduplication is, indeed, prevalent in Cushitic Ethiopian languages, including Kemant Agaw: *käl-* 'break' / *käläkäl-* 'shatter' (Appleyard 1975: 330), and Sidamo: *kad-* 'kick'/*kakkad-* (<*kadkad-*) 'kick repeatedly' (Hudson 1976: 272). The apparent absence of this form in Ge'ez is interesting, and might argue that the form is not, indeed, Semitic, but borrowed in other ES languages from Cushitic, an interpretation countered by the fact that the particular pattern of reduplication of (4) is otherwise common and reconstructible ES, and is not in any straightforward or borrowable way related to the simpler Cushitic pattern of doubling of a biconsonantal stem.

2.4. Non-Semitic Afroasiatic Languages

Other Afroasiatic languages contribute considerable evidence for characteristics of consonant length in the present-tense verb and a lexical dichotomy of stative and active (or intransitive and transitive: Diakonoff (1965: 78ff). Berber and Cushitic, particularly, illustrate the former characteristic, and Cushitic languages the latter. Rössler (1950), Greenberg (1952), and Diakonoff (1965) presented such evidence in support of their respective conclusions that the characteristics are not just Semitic but Afroasiatic.

Greenberg mentioned an array of evidence from Berber, Cushitic, Chadic and Egyptian for an Afroasiatic present tense characteristic of gemination, and Rössler mentioned evidence from these same groups (just Beja of Cushitic) for two conjugations distinguishing two semantic groups of verbs. Recently, Voigt (1987) added new comparisons to the Cushitic and Chadic evidence for an Afroasiatic formal distinction between geminating and nongeminating prefixed conjugations. Cushitic languages often show relics of the Afroasiatic prefix conjugation (Beja, Somali, Saho, Southern Agaw). Verbs of a closed class are conjugated with prefixes, and verbs of an open class are conjugated with suffixes deriving from an auxiliary bearing the Afroasiatic subject prefixes seen in the closed class. In (5) is a small sample of this evidence: two Tamazight Berber verbs each in the perfect and in the geminated imperfect so-called 'intensive' (Pencheon 1973: 34), and three forms each of two Saho verbs, one of the suffixing and one of the prefixing verb classes (Welmers 1952: 236).

(5) a. Tamazight basic and intensive verbs (3m.sg. forms)

Perfect	Present ('Intensive')	
i-kroz	*i-korroz*	'plough'
i-zrəy	*i-zərrəy*	'pass'

b. Saho suffixing and prefixing verbs

Sg.1	*ḥab-e* 'I left'	*u-bl-e* 'I saw'
Sg.2	*ḥab-te*	*tu-bl-e*
Pl.3	*ḥab-en*	*yu-bl-in*

Among the evidence that the Afroasiatic formal verb dichotomy had a semantic basis is that of the 100 or so Saho prefixing-type verbs about two-thirds are transitive, but of the open-class suffixing verbs only about half are transitive. The Tamazight 'intensive' is basically just a present or imperfect, and would no doubt be so called if it were not for the Arabic-influenced grammatical tradition for these languages, which, as for ES, interprets second-consonant gemination as formation of an intensive.

2.5. Reconstruction

The few Semitic comparisons of (1), (2), (3) and (4), and our very brief consideration of other Afroasiatic languages of (5), are not intended to establish a reconstruction of Proto-Semitic or certainly Afroasiatic verbal morphology, but

only to show the range of data and opinion for the likelihood of (6), argued, generally, by the mentioned authors and others.

(6) Proto-Semitic and/or Afroasiatic basic verb morphology

      a. *CVCVC(-V), a stative
      b. *(y)V-CCVC, a past or perfect active
      c. *(y)V-CVC:VC, a present or imperfect active

3. The ES A/B-type Dichotomy

The purpose of this paper is to take note that concerning (6), probably the most researched issue in Afroasiatic linguistics, the most prominent characteristic of ES languages, the A/B-type dichotomy, has been almost completely ignored, except uncritically as support for a geminating intensive for which the B-type is the assumed reflex. Reasonably important questions about the history of the two types have never been asked: (i) How could they have evolved from a distinction of basic and derived forms? (ii) Why is the semantic evidence of such a history so lacking? (iii) If the B-type is an old intensive, what is the role of the common ES reduplicated intensive?

It seems apparent that the questions were not considered important since the forms only, not form and meaning correspondences, were compared to those of other Semitic languages. The initial associations were made before the Akkadian evidence came along, and form-wise the B-type goes alongside the Arabic form II. Even when evidence has been raised for consonant length in the reconstructed present stem (by Rössler, Greenberg, and Voigt), and about the contribution of stative/active meanings in the proto-system (Goetze, Rössler, Buccellati), the ES evidence has not been reevaluated. Difficulty or obscurity of the data cannot explain this result, which, however, can be readily attributed to a combination of factors:

a. The classical 'analogist' tendency in historical linguistics to ignore irregularity or lexical characteristics in both data and reconstruction, which resulted in suppression of recognition in the ES grammars of the plainly lexical status of the B-type;

b. Unreasonable priority in reconstruction given to Arabic, which focused attention on Arabic and Arabic-like phenomena at the simple level of form, a tendency generally recognized and corrected, it now seems, within Semitic linguistics;

c. Perhaps an element of racial prejudice in the original discounting of the Ethiopian data, as also seen, for example, in the employment and persistence of the notion 'Hamitic language', according to Greenberg (1955: 55 [1963: 51]).

But objective evaluation of the ES A/B-dichotomy must support reconstruction of the Akkadian-type over the Arabic-type verbal system, including the existence of a formally expressed distinction of stative and active verbs.

1. My main point: since the A/B-type dichotomy must be reconstructed for Proto-ES, there is unquestionably an ES cognate to the Akkadian present in the Proto-ES B-type, which is also characterized by gemination of the second consonant.

2. The source of the A/B-type dichotomy is not reasonably in a distinction of basic and derived intensive (or causative) stem, since the breakdown of such a distinction would favor persistent analogical extension of the basic stem, and

survival of this stem only, with perhaps only lexicalized and, at least generally, semantically transparent relics of the earlier intensive or other derived meanings, and not the actual ES state of competing lexical types. Furthermore, as we have seen, the category of reduplicated derived intensive is alive and well in ES.

3. A semantic and lexical dichotomy reflected in a difference of forms, such as in Akkadian, of stative vs. active (or intransitive vs. transitive, durative vs. punctual), however, would reasonably break down in a competition between the forms which, with crossing analogies of form, could yield the variety of A/B-type outcomes in ES as seen in (1). Starting earlier or in slightly different circumstances, the active/transitives could have been reinterpreted as an intensive or causative formation in some languages (Arabic, Hebrew), have survived as prefixing relics only in some (Saho, Beja, Somali, Agaw), and have been fully replaced by leveling in other languages (Oromo, Sidamo, etc.).

4. Most importantly, there is very good ES evidence that the A/B-types reflect an earlier contrast of meaning, along the lines of stative vs. active or intransitive vs. transitive. I have discussed some of this evidence elsewhere (Hudson 1979, 1991).

a. Numerically, B-types tend significantly to be transitives (Cowley 1969: 5; Bender and Fulass 1978: 78).

b. Throughout ES there survives, alive and well, a formally-reflected distinction of transitive and intransitive: in most of the languages there are two causative-forming prefixes, one for use with intransitive roots and one for use with transitives. Furthermore the two causatives give evidence for the basically transitive origin of the B-types, since typically causatives of transitives, whether A or B type, are conjugated with the B-type characteristic of long second consonant or, in the languages that have lost gemination, such as Harari, the B-type's front vowel. Some of these data are presented in (7), from Amharic, which has causative prefixes *a-* and *as-*, the latter for transitives, and Harari, which has causative prefixes *a-* and *at-*, the latter for transitives (Leslau 1958: 32).

(7) ES Causatives (A-type 3m.sg. present forms)

	Root		Causative
Amharic	*dkm*	'tire' (vi.)	*y-a-dǝ kɨ m*
' '	*sbr*	'break' (vt.)	*y-as-sǝ bbɨ r*
Harari	*bsl*	'ripen' (vi.)	*y-a-bǝ sli*
' '	*sbr*	'break' (vt.)	*y-at-seebri*

c. Passive/reflexive derivatives, throughout ES marked by the prefix *t-* (perhaps basically reflexive in Semitic), are necessarily based on transitive rather than intransitive verbs. Interestingly, in ES and often elsewhere in Semitic, passive/reflexives are derived with a long second consonant or other B-type characteristic (in Harari, stem-shape CVCVC, discussed below), even when formed on A-type roots like *sbr* 'break', as in (8).

(8) ES passives (3m.sg. present)

Ge'ez	*yɨ-t-səbbər*
Amharic	*yɨ-s-səbbər* (*<yɨ-t-səbbər*)
Harari	*yɨ-t-səbər*

Arabic form 5, generally a reflexive, has gemination in both perfect *taqattal-u* and imperfect *ya-taqattal-u.*

    d. Present conjugation stems of ES reflexive/passive verbs, in addition to B-type gemination, have not only the stem shape but the vocalization of the past conjugation, not of the present. See in (8) the Ge'ez and Amharic passive present stem *səbbər* vs. Ge'ez basic present *səbbər* , Amh. *səbr*, and Harari *səbər* vs. *səbri* of the basic present. This unexpected pattern, also seen in the Arabic form 5 (see above), can be understood in light of the present hypothesis, in which the present passive would be a form/meaning hybrid: in meaning active by virtue of basic or root transitivity and stative by virtue of passive derivation. The form of the present passive may be understood as an extension of the prefixes of the present active conjugation to a stative stem CVC(:)əC, perhaps with gemination extended from the basic present stem CVC:ɨC. Such an understanding may seem rather complex, but such an outcome would seem reasonable in the breakdown of the formal distinction of stative and active, which resulted also, I have argued above, in the ES A/B-type dichotomy itself.

    e. Perhaps the front vowel characteristic of the B-type may receive explanation on the hypothesis that the B-type was the active, originally prefixing, verb (and the A-type a later extension of the stative stem into the present (earlier active) conjugation by the addition of the prefixes). The first stem-vowel of the Proto-ES (present/active) B-type *yi -Ca C:VC (Akkadian i-Ca C:{i,u,a}C), would have found itself in a closed syllable preceded and often followed by high vowels, and so could naturally have yielded *yi- Ce C:VC by umlauting.

References

Appleyard, D. L. 1975. A descriptive outline of Kemant. BSOAS 38:316-350.

Bender, L. and Hailu Fulass. 1978. Amharic verb morphology. East Lansing, MI: African Studies Center, Michigan State University.

Buccellati, Giorgio. 1968. An interpretation of the Akkadian stative as a nominal sentence. Journal of Near Eastern Studies 27:1-12.

Buccellati, Giorgio. 1988. The state of the stative. Fucus: A Semitic/Afrasian gathering in remembrance of Albert Ehrman, Yoël L. Arbeitman, ed., 153-189. Amsterdam: John Benjamins.

Cohen, Marcel. 1936. Traité de langue amharique. Paris.

Conti Rossini, C. 1941 [1967]. Grammatica elementare della lingua etiopica. Rome: Istituto per l'Oriente.

Cowley, R. 1969. A and B verbal stem type in Amharic. Journal of Ethiopian Studies 7:1-13.

Diakonoff, I.M. 1965. Semito-Hamitic languages: An essay in classification. Moscow: Nauka.

Dillmann, August. 1907. Ethiopic grammar. Carl Bezold, ed., James A. Crichton, trans. London: Williams & Norgate.

Goetze, Albrecht. 1942. The so-called intensive of the Semitic languages. JAOS 62:1-8.

Greenberg, Joseph. 1952. The Afro-Asiatic (Hamito-Semitic) present. JAOS 72:1-9.

Greenberg, Joseph. 1955. Studies in African linguistic classification. Branford, CN: Compass [revised version 1963, The languages of Africa, IJAL 29.1, part II].

Haupt, Paul. 1878. The oldest Semitic verb-form. Journal of the Royal Asiatic Society of Great Britain and Ireland 10:244-257.

Hetzron, Robert. 1972. Ethiopian Semitic: Studies in classification. Manchester: Manchester University Press [Journal of Semitic Studies, Monograph no. 2].

Hudson, Grover. 1976. Highland East Cushitic. The non-Semitic languages of Ethiopia, M. L. Bender, ed., 232-277. East Lansing, MI: African Studies Center, Michigan State University.

Hudson, Grover. 1979. The Ethiopian Semitic B-type. Proceedings of the Fifth International Conference of Ethiopian Studies, R. Hess, ed., 101-110. Chicago: University of Illinois Press.

Hudson, Grover. 1991. A and B-type verbs in Ethiopian and Proto-Semitic. Semitic studies in honor of Wolf Leslau, Alan S. Kaye, ed., 679-689. Wiesbaden: Otto Harrassowitz.

Huehnergard, John. 1987. 'Stative,' predicative form, pseudo-verb. Journal of Near Eastern Studies 46:215-232.

Leslau, Wolf. 1936. Étude descriptive et comparative du gafat. Paris: Librairie Klincksieck [Collection linguistique, LVII, Société de Linguistique de Paris].

Leslau, Wolf. 1953. The imperfect in South-East Semitic. JAOS 73:164-168.

Leslau, Wolf. 1958. The verb in Harari. Los Angeles: University of California Press [University of California Publications in Semitic Philology, vol. 21].

Moscati, Sabatino, et al. 1964. Introduction to the comparative grammar of the Semitic languages. Wiesbaden: Otto Harrassowitz.

Pencheon, Thomas G. 1973. Tamazight of the Ayt Ndhir. Los Angeles: Undena [Afroasiatic Dialects, 1].

Rössler, Otto. 1950. Verbalbau und Verbalflexion in den Semitohamitischen Sprachen: Vorstudien zu einer vergleichenden Semitohamitischen Grammatik, ZDMG 100.2:461-514 [1981. Bono Homini Donum: Essays in historical linguistics, in memory of J. Alexander Kerns, Yoël L. Arbeitman and Allan R. Bomhard, eds. 679-748. Amsterdam: John Benjamins].

Sayce, A.H. 1880. An introduction to the science of language, Vol. I. London: Kegan Paul.

Ullendorff, Edward. 1960. The Ethiopians, an introduction to country and people. London: Oxford University Press.

Welmers, William. 1952. Notes on the structure of Saho, 2 parts, Word 8:145-162, 236-251.

Voigt, Rainer M. 1987. The two prefix-conjugations in East Cushitic, East Semitic, and Chadic. BSOAS 50 (part 2):330-345.

From Focus Marker to Copula in Swahili

John McWhorter
University of California, Berkeley

1. Introduction

This paper addresses the diachronic development of the copula *ni* in Swahili. Although today this item links subject and predicate in all persons and most predication types, historical and synchronic evidence indicate that this situation has only gradually arisen over the past few centuries. As recently as a century ago, the usage of *ni* was more constrained; it shared the copular domain with various other strategies. Diachronic developments since then have created a copula system anomalous among the Eastern Bantu languages, and I will show that while certain aspects of this development follow established patterns of syntactic change, other aspects, as well as the rate of the aforementioned changes, resulted from the widespread use of Swahili as a second language.

2. The Copula in Modern Swahili

It is extremely common for languages to morphologically encode a distinction between two copular subdomains: the equative, as in the Ewe *É-nye fia* 'I am the chief', and the locative, as in the Ewe *Me-lè xo me* 'I am in the house' (Kofi Agawu, p.c.). Swahili makes such a distinction, in that the copula *ni* is used in the equative, while in nonverbal locative predications, the copula is expressed with zero:

(1) Hamisi **ni** mpishi.
    Hamisi COP cook
    'Hamisi is a cook.'

(2) Hamisi **yu-ko** nyumba-ni.
    Hamisi AGR-LOC house-LOC
    'Hamisi is in the house.'

Ni is used in all persons and numbers. What is interesting is that in the spoken language, the omission of *ni* is grammatical (3). Informants suggest that this is particularly common after independent pronouns (4) and determined NPs (5).

(3) Hamisi ø mpishi.
    Hamisi cook
    'Hamisi is a cook.'

(4) Sisi ø wa-pagazi.
    we PL-porter
    'We are porters.'

(5) Mi-zigo hii ø mi-zito.
    PL-load these AGR-heavy
    'These loads are heavy.'

Informants universally concur that this omission is considered substandard; it is exceedingly rare in written texts. However, the fact that this construction occurs is a vital clue that zero copula was the original construction, given that a change from state A to state B in a grammar is generally mediated by a stage where A and B coincide, in combination with the rarity of attestations of the disappearance of copulas in natural change.

3. The Copula in Early Modern Swahili

Indeed, in Swahili texts of the seventeenth and eighteenth centuries[1], the language of which I will henceforth designate Early Modern Swahili (EMS), equative sentences quite regularly have zero copula. It is especially significant that we encounter this in writing, which suggests that the construction had yet to acquire a sociolinguistically pejorative connotation:[2]

(6) *Pili   ø Rahamani Muwawazi wema na   wawi.*
    second-word   merciful merciful   good and bad
    'The second word is the merciful, merciful for the good and the bad.'
    (Knappert 1969:4)

(7) *Moli wetu ø Rabbi tangu insi   na ajinani.*
    Lord our   Lord since mankind and demons
    'Our Lord is the Lord of men and demons since time immemorial.' (ibid. 14)

Meanwhile, in locative sentences, the verb *li* is used, in contrast to the modern (2):

(8) *U-li-we   mbali haka...*
    you-be-you far   very
    'You are quite far removed...' (Miehe 1979:216)

(9) *Tu-li-po.*
    we-be-LOC
    'Here we are.' (Knappert 1969:8)

In non-present tenses, *li* is used in equative sentences as well:

(10) *A-a-li   mtoto tifuli.*
     she-IMP-be child   still
     'She was still a small child.' (Miehe 1979:216)

We see a cognate configuration in modern ChiBemba, in which zero copula is still grammatical in many equative constructions, although there is a *ni* cognate as well, and in the meantime, *li* is used in locative predications:

ChiBemba:
(11) a. Ma-kòndé.
       they-nets
       'They are nets.' (Sadler 1964:467)

b. Ishìna liakue **ni**   Càali.
   name his   COP Caali
   'His name is Caali.' (ibid. 155)

c. Tù **lì**   kuno.
   we COP here
   'We are here.' (ibid. 156)

4. The Development of *Ni*

Evidence of all kinds suggests that the copula *ni* is derived from the presentative reflex of *ni* used as a third person deictic in subject position:

(12) **Ni**   mamoja   kwangu.
    it-is   the-same   to-me
    'It's all the same to me.'

This usage would have been reanalyzed as a copula as a result of usage as a resumptive in topic-comment constructions. Such constructions, when used in contrast to subject-predicate constructions, highlight the identity of the predicate to the topic:

(13) [*Vita*]   [*ni*]   ø [*taabu*].
    topic   subject   predicate
    TOPIC COMMENT
    'War, that's trouble.'/ 'What war is is trouble.'

Over time, the force of the highlighting would have eroded via usage. Concomitantly, the topic would come to be processed as a subject, a common tendency in language change, forcing a reanalysis of *ni* as a copula as it necessarily moved into the copula slot after vacating the subject position for the erstwhile topic. Thus such sentences are now processed as subject-predicate sentences:

(14) [Vita]   [ni]   [taabu].
    subject copula predicate
    'War is trouble.'

Support for this scenario comes from various quarters. First, this type of development has been richly documented cross-linguistically. For example, Archaic Chinese had zero copula sentences (15) which contrasted with topic-comment constructions using a proximal demonstrative (16):

Archaic Chinese:
(15) Wáng-Tái ø wù   zhe   ye.
    Wang-Tai   outstanding person DEC
    'Wang-Tai is an outstanding person.' (Li & Thompson 1977:421)

(16) Jì   yù   qí   sheng yoù yù   qí   si, **shì** huò   ye.
    already wish him live   also wish him die this indecision DECL
    'Wishing him to live while wishing him to die, that is indecision.' (ibid. 424)

By the first century, *shì* appears in unambiguously copular usages, and this usage is regular today:

First century Chinese:
(17) Yú **shì** suo   jià   fu-ren zhi fù   ye.
    I   be NOM marry woman GEN father DECL
    'I am the married woman's father.' (ibid. 426)

A similar process has been documented for languages as diverse as Hebrew (Li & Thompson 1977:427-431), Panare (Gildea 1993), and Saramaccan Creole (McWhorter 1993).
    We would expect that between the stage where *ni* served to highlight identity and its unambiguous processing and the stage where *ni* came to be processed as a copula, sentences with *ni* would retain a degree of the highlighting semantics despite having advanced significantly towards copular processing. The grammar by Delaunay (1885) apparently captures this stage in describing *ni* as a copula serving to lend attention to the predicate (161):

> '[To] answer this question: "Who is your brother?", if I say "My brother is a sultan", then it is the notion of the sultan that carries the weight of the sentence, and I say *Ndugu yangu ni sultani* [My brother is a sultan]. On the other hand, to answer this other question: "Is everybody ready?", if I say "Us, we are ready...", it is as if I said: "As to us, we are ready," and I will say: *Tu tayari* [We are ready]'.

We find further support for the pathway described above in another non-copular synchronic reflex of *ni*. The presentative *ni* can be analyzed as one manifestation of an element of more general syntactic movement which can be characterized as a general focus marker, occurring before a constituent which the speaker desires to call particular attention to:

(18) a. Momela Farm **ni**   i-me-fany-wa   National Park.
    Momela Farm FOC AGR-PAST-make-PASS National Park
    'It is Momela Farm that was made into a National Park.'
    (Closs, Kondo & Mbaye 1967:24)

    b. Momela Farm i-me-fany-wa   **ni**   National Park.
    Momela Farm AGR-PAST-make-PASS FOC National Park
    'What Momela Farm was made into is a National Park.' (ibid.)

It is clear that this usage is not copular, in that it can occur not only with full verbs as above, but also with verbs 'to be':

(19) Juma a-li-kuwa   **ni**   mkurugenzi.
    Juma AGR-PAST-be FOC director
    'Juma was the *real* director.' (Wesana-Chomi 1978)

However, the semantics of a sentence like (19) are significant in that we see *ni* again used in a deictically highlighting function. The difference between (19) and sentences like (13), which produced the copula *ni*, is simply that in (19) the full verb 'to be', required in a non-present sentence, blocks *ni* from reanalysis as a

copula. Thus (19) reflects the semantic contribution that *ni* presumably made to present-tense nonverbal sentences before its reanalysis as a copula:

(20) a. EMS:
  [*Vita*] [*ni*] ∅ [*taabu*].
  topic subject predicate
  TOPIC COMMENT
  'War, that's trouble.'

 b. Juma a-li-kuwa **ni** mkurugenzi.
  Juma AGR-PAST-be FOC director
  'Juma was the *real* director.'

Moreover, the historical documentation in combination with grammars over the past one hundred years support the hypothesis in rather neatly indicating a progression from zero copula as regular to *ni* as regular. The earliest text extant, the Hamziya, has no occurrences of *ni* as a copula. Although this hardly confirms that *ni* had not begun the process of reanalysis, its absence here is nevertheless suggestive. In 18th century texts, copular *ni* alternates with zero, with the latter somewhat more frequent. As late as 1900, Seidel describes zero copula as the default condition, then proceeding to outline alternative strategies such as *ni*. By the mid-twentieth century, grammarians such as Ashton in 1944 and Perrott in 1951 describe zero copula as secondary to the usage of *ni* and restrict it to certain grammtical contexts; by the late twentieth century grammarians describe it as marginal if at all.

 A final piece of evidence supporting my claim is that according to the scenario I have sketched, *ni* would have been first reanalyzed as a copula in third-person sentences, given the third-person domain of the presentative *ni* used as in (12). This is confirmed by Ashton, who describes *ni* in 1944 as used only in the third person. Equally important is that cognates of *ni* in related languages are typically restricted to third person, as we see in ChiBemba and Kikuyu:

ChiBemba:
(21) a. Ishìna liakue **ni** Càali.
  name his COP Caali
  'His name is Caali.' (Sadler 1964:155)

 b. Tù **lì** bàana be sùkuulù.
  we COP children of school
  'We are students.' (ibid. 156)

Kikuyu:
(22) a. Muti **ni** mwega.
  tree COP good
  'The tree is good.' (McGregor 1905:121)

 b. N-**de** mwega.
  I-COP good
  'I am good.' (ibid.)

Thus a rich array of evidence demonstrates that the copular *ni* emerged as a result of its usage as a resumptive subject in topic-comment constructions.

5. The Evolution of the Locative Construction

On the other hand, locative constructions underwent a completely separate restructuring at the same time. As we have seen, the locative copula in EMS was the verb *li*:

(23) *N-a-li-mo   usindizi.*
    I-IMP-be-LOC sleep
    'I was asleep.' (Miehe 1979:216)

Today, however, *li* serves a past tense marker:

(24) Ni-**li**-nunua kitabu.
    I-PAST-buy book
    'I bought a book.'

The reanalysis of *li* as a past tense marker most likely occurred within the context of the EMS imperfect, which was encoded with the infix *-a-*:

(25) *A-a-li   mtoto tifuli.*
    she-IMP-be child   still
    'She was still a small child.' (Miehe 1979:216)

The reanalysis was spurred by the co-existence of two phenomena. First of all, the third person pronominal was (and is) *a*, creating in the spoken language a potentially ambiguous portmanteau morpheme of frequent occurrence in third person imperfect sentences as in (25). Concomitantly, it would appear that the verb *li* was also regularly used in imperfect constructions, a situation made likely by the fact that *li* is regularly used in this fashion in sister languages such as ChiBemba:

ChiBemba:
(26) Tù **à**   **lî** mona màayo.
    we PAST be see   mother-my
    'We saw my mother.' (Sadler 1964:55)

The result would have been that in the EMS equivalent of (26), shown in (27), the stage would have been set for the reanalysis of *li* as a past marker, with new generations processing the initial *a* as simply a pronominal, with no imperfect semantics.

Reconstructed EMS sentence:
(27) *A-a-li   ona mtu.*
    he-IMP-be see man
    'He saw a man.'

When *li* was reanalyzed as a past marker, its semantic contribution to present-tense locative predications would have become opaque, and to maximize transparency, it dropped out of such constructions. Thus emerged the modern locative construction as in (2), which had already been regularized by the late nineteenth century, as we see in (28):

Late nineteenth century:
(28) Sasa **tu-ko**   katika nchi   nyingine.
      now we-LOC in   country different
      'Now we were in another country.' (Harries 1965:124)

It is highly likely that it was indeed the reanalysis of *li* as a past marker which spurred the disappearance of *li* from locative constructions, given that *li* retains its locative function in the sister languages which have not incorporated *li* into the tense-affix paradigm, as we see in sentences such as the ChiBemba (11c).

   Thus we see how two independent processes have neatly reversed the division of labor in the Swahili copula scenario: where in EMS, equative predication was expressed with zero while an overt element was used with locatives, today an overt element is regularly used in equative predications while zero is the rule with locative ones.

EMS: MODERN:

Hamisi ø mpishi.   Hamisi **ni** mpishi.
Hamisi   cook   Hamisi COP cook
'Hamisi is a cook.'   'Hamisi is a cook.'

Yeye yu-li-ko   *nyumba-ni.*   Yeye **yu-ko**   nyumba-ni.
he   AGR-COP-LOC house-LOC   he   AGR-LOC house-LOC
'He is in the house.'   'He is in the house.'

6. Swahili Diachrony and Second-Language Acquisition

Standard Swahili is generally spoken only in official contexts, except by a minority of coastal Muslims. It is a second language for most of its speakers, serving as the primary vehicle of interethnic contact in East Africa. I will finally discuss how the phenomena described above are manifestations not only of explainable tendencies in syntactic and semantic change, but also of the effects of second-language acquisition upon the Swahili grammar.

   In the process of its evolution, *ni* has succeeded in marginalizing what was a competing copular strategy in earlier stages of the grammar. In EMS and up until about the turn of the century, free-standing subject prefixes, reflecting the earlier status of those prefixes as independent pronouns, could be used to associate subject and predicate as well:

(29) Hamisi **yu** mpishi.
      Hamisi he cook
      'Hamisi is a cook.'

(30) *Uwe kifungo-ni u   ø kijuwo muutasima.*
      you knot-LOC you   book precious
      'Thou art in that unison the well-guarded book.' (Knappert 1968:64)

This is another common source of copulas cross-linguistically, currently occurring, for example, in Hebrew:

Hebrew:
(31) David **hu** ha-ganav.
 David he/COP the-thief
 'David is the thief.' (Li & Thompson 1977:429)

Grammars throughout the century describe this construction as increasingly marginal; in 1967 Closs, Kondo & Mbaye found that out of twelve informants two rejected it altogether while the other ten found it archaic or substandard. By this time, the construction had apparently narrowed in its semantic scope to indicating particularly strong association between subject and predicate, but even here, it shared that domain with the dominant strategy, *ni*, as we see in (32).

(32) a. Wewe **u** mtoto mbaya.
 you you bad child
 'You are a bad child (by nature).'

 b. Wewe **ni** mtoto mbaya.
 you COP child bad
 'You are a bad child (today or by nature).' (Closs, Kondo & Mbaye 1967:7)

Given that equivalents of this strategy are more vigorous in many of Swahili's sister languages, it is possible that the swift and unequivocal generalization of *ni* was a function of it being more easily manipulated by second language speakers with only partial command of the morphology, given that the subject prefix strategy required command of six different prefixes. In support of this account, similar generalization of single copula morphemes is typical of reduced and restructured registers of Swahili. For example, urban Kenyans have generalized the locative construction we saw in (2) as a stative marker, such that it is used even with descriptive predicates:

(33) Vipindi vya televisheni **vi-ko** boring.
 shows of television AGR-LOC boring
 'The television shows are boring.' (Myers-Scotton 1984:8)

Similarly, while the construction in (33) retains allomorphic variation, there are pidginized registers of Swahili in which a generalized item *iko* has been similarly overgeneralized but has even lost all vestiges of agreement, invariant across the board:

(34) Yeye **iko** tayari.
 he COP ready
 'He is ready.' (Duran 1979:148)

Heine (1973:95-6) has described a register where *iko* is part of a new copula paradigm in which *ndiyo* is particularly emphatic, *ni* is neutral, and *iko* falls in between. A final indicator that second-language acquisition is a significant factor in the changes I have outlined is that in more isolated Northern dialects of Swahili, *li* is still used as a verb 'to be', in contrast to its incorporation as a past marker in the standard (Wald 1987:1005).

7. Conclusion

I have intended to demonstrate that the current copula configuration in Swahili is a recent development, representing a significant departure from the configuration more typical of languages of its family. The differences that Swahili displays are indeed due in part to established patterns of syntactic change. The development of resumptive deictic subjects in topic-comment constructions into copulas is well-attested cross-linguistically, while the incorporation of the verb *li* into the verbal affix paradigm is a predictable outcome of its usage in an imperfect construction which already contained an ambiguous portmanteau morpheme encoding both the third person singular and imperfect aspect. However, the speed and extent of these changes, which have left Swahili an anomaly amidst its more conservative sisters, are manifestations of the effects of centuries of second-language acquisition upon the grammar.

Notes

1. Sources: seventeenth-eighteenth century - Poetic Documents: *the Hamziya* (seventeenth), *the Kasida* (early eighteenth), *the Herekali* (late eighteenth) (as edited and translated by Knappert 1967, 1968, 1969 and Miehe 1979); nineteenth century: "Safari Yangu ya Urusi na ya Siberia" (1896), by Salim bin Abakari (Harries 1965).
2. Early Modern Swahili citations will be given in italics.

References

Ashton, Ethel O. 1944. Swahili grammar (including intonation). London: Longman.

Closs [Traugott], Elizabeth, Ali Kondo, and Salma Mbaye. 1967. Some copula patterns in Standard Swahili. Swahili 37:5-25.

Delaunay, (Le Père). 1885. Grammaire Kiswahili. Paris: F. Levé.

Duran, James J. 1979. Non-standard forms in Swahili in West-Central Kenya. Readings in Creole studies, ed. by Ian F. Hancock, 129-51. Ghent: E. Story-Scientia.

Gildea, Spike. 1993. The development of tense markers from demonstrative pronouns in Panare (Cariban). Studies in Language 17:53-74.

Harries, Lyndon. 1965. Swahili prose texts. London: Oxford University Press.

Heine, Bernd. 1973. Pidgin-Sprachen im Bantu-Bereich. Berlin: Dietrich Reimer.

Knappert, Jan. 1967. Traditional Swahili poetry. Leiden: E.J. Brill.

_____. 1968. The Hamziya deciphered. African Language Studies 9:52-81.

_____. 1969. The discovery of a lost Swahili manuscript from the eighteenth century. African Language Studies 10:1-30.

Li, Charles N. and Sandra A. Thompson. 1977. A mechanism for the development of copula morphemes. Mechanisms of syntactic change, ed. by Charles N. Li, 419-44. Austin: University of Texas Press.

McGregor, A. Wallace. 1905. A grammar of the Kikuyu language. London: Richard Clay & Sons, Ltd.

McWhorter, John H. 1993. Towards a new model of genesis: Competing processes in the birth of Saramaccan Creole. PhD Dissertation, Stanford University.

Miehe, Gudrun. 1979. Die Sprache der Älteren Swahili-Dichtung. Berlin: Dietrich Reimer.

Myers-Scotton, Carol. 1984. Roots in Swahili? A locative copula becomes a stative marker. Michigan State University Ms.

Perrott, Daisy Valerie. 1951. Teach yourself Swahili. London: The English University Press.

Sadler, Wesley. 1964. Untangled CiBemba. Kitwe, Northern Rhodesia: The United Church of Central Africa in Rhodesia.

Seidel, August. 1900. Swahili Konversations-Grammatik. Heidelberg: Julius Groos.

Wald, Benji. 1987. Swahili and the Bantu languages. The world's major languages, ed. by Bernard Comrie, 991-1014. New York: Oxford University Press.

Wesana-Chomi, E. 1978. A note on the apparent problem in the description of copula constructions in Kiswahili. Kiswahili 48:38-40.

Restructuring, Feature Selection, and Markedness: From Kimanyanga to Kituba[*]

Salikoko S. Mufwene
University of Chicago

1. Introduction

Kituba in this paper is short for *Kikongo-Kituba*, also known as *Kikongo ya leta, Kileta, Kikongo ya bula-matadi, Kibula-matadi*, or *Mono-kutuba*, whence *Kikongo-Kituba*, among other names listed in Heine (1970) and discussed in Mufwene (to appear-a). It is spoken primarily in the Southwestern part of Zaire and the Southern part of the Congo, although it is apparently also used in the Northern part of Angola. It is better known as a creole, having been discussed more in creole studies than in African linguistics (see also Part 4). It must have developed toward the end of the nineteenth century associated with the formation of the Congo Free State in 1885, which officially marks the beginning of the colonization of what is now known as Zaire.

An important question in this and other papers of mine is: why does Kituba seem to diverge from the Bantu morphosyntactic canon regarding some features considered typical of this genetic group of languages? The question becomes more important as we consider the fact that, unlike Atlantic creoles lexified by European languages, Kituba falls in the category of what Chaudenson (1979) calls 'endogenous' creole, for two reasons. First, it started in Bas-Zaire (the Lower Zaire Region), the Kikongo-speaking area, in which speakers of the lexifier and related languages have always been indigenous and the majority. The odds for the formation of a more Bantu-like Kituba, similar to Lingala, seem to have been as great as those which led to the formation of Tok Pisin and Bislama, reflecting strong influence from Melanesian and other Austronesian languages.

Second, even after spreading eastward, outside Bas-Zaire and into the Bandundu Region (east of the Kwango River), it has been used still by Bantu speakers, even though of languages not belonging in what Hinnebusch (1989) calls the 'Kikongo Group.' Since influence from these other languages was just as likely in the earlier stages of Kituba's emergence in Bas-Zaire as in the later stages of its spread (see below), we may treat the creole as a primarily Bantu contact phenomenon whose usage has remained within the relevant Bantu territory. Thus it may be considered more endogenous than some creoles such as Krio and Nigerian Pidgin English, which have a non-African lexifier.[1]

Despite some structural differences among Bantu languages, the spread of Kituba both within and outside Bas-Zaire, in the Bantu territory, raises the question of why its morphosyntax and its tonological systems in particular have turned out so different from those of Lingala (Mufwene 1989b), another Bantu contact

phenomenon that developed around the same time, at the end of the nineteenth century and under very similar colonial circumstances, with the West Africans still involved in the colonial work force—at least according to Samarin (1982, 1989, 1990). For instance, Lingala has subject-verb agreement (albeit restructured and simplified) and an incorporated reflexive marker, but Kituba does not.

I thus decided to examine morphosyntactic features of some of the Bantu languages of the Bandundu Region and compare them with the alternatives available in Kimanyanga, one of the Kikongo Group varieties, which has been assumed since Fehderau (1966) to be Kituba's lexifier (see below). Without discounting the element of chance, I wished to determine whether either the lexifier or some of the other Bantu languages in contact with it may have features that influenced the development of Kituba in the direction of divergence from the Bantu canon presented in much of the literature. I intended to focus on Kiteke (B75, according to Guthrie 1953), because its speakers had been involved in the precolonial trade and served in the colonial system (Samarin 1989). However, lack of accessible information on it led me to start by checking a related language, Kiyansi (Bantu B85, according to Guthrie 1953), which I speak natively. The Bayansi, sometimes confused with the Babangi (Whitehead 1899), were also involved in the precolonial trade and live in the area east of the Kwango River, where Kituba has spread as a major lingua franca and as an urban vernacular. A sketch of that comparison, partial to date, follows in Part 2, after which I raise some issues regarding Kituba's genesis in Part 3 and some others on Bantu genetic linguistics in Part 4.

2. Some Morphosyntactic Characteristics of Kituba: A Genetic Perspective

One of the most noticeable features of Kituba in comparison with Bantu languages is absence of subject verb-agreement, as illustrated below:[2]

(1) *Móno 0+mon+áka   yándi   ye yándi   0+mon+áka   móno.*
     me AGR+see+ANTER him/her and him/her AGR+see+ANTER me
     ' I saw him/her and he/she saw me.'

I was curious whether this apparent idiosyncrasy was a confirmation of the kinds of innovations—virtually ex nihilo—claimed by Bickerton (1981, 1984) and in later work, or something that Kituba may have selected from any of the languages in contact. As I show in Part 3, I also had doubts, contrary to Samarin (1982, 1990), about the extent of the influence of West African colonial escorts in the restructuring of Kimanyanga into Kituba. So I started with Kiyansi, at least my Kibwaal dialect. A startling finding in this project was that Kiyansi diverges from the Bantu canon, as illustrated below:

(2) Sample Verb Conjugation in Kiyansi:

mɛ mǎ dia	*nzɛ́ mǎ dia*	*ndɪ́ mǎ dia*	*bɪ́ mǎ dia*
me NPERF eat	you-SG NPERF eat	he/she NPERF eat	we NPERF eat
'I have eaten'	'You have eaten'	'He/She has eaten'	'We have eaten'

mɛ ɛ+dɪ́	*nzɛ́ ɛ+dɪ́*	*ndɪ́ ɛ+dɪ̆*	*bɪ́ ɛ+dɪ́*
I RPERF eat	You RPERF eat	He/She RPERF eat	We RPERF eat
'I ate (already)'	'You ate'	'He/she ate'	'We ate'

mɛ ǎyɪ́: ´dia	*nzɛ́ ǎyɪ́: ´dia*	*ndɪ́ áyɪ́: ´dia*	*bɪ́ ǎyɪ̄: ´dia*
me FUT eat	you-SG FUT eat	he/she FUT eat	we FUT eat
'I will eat'	'You will eat'	'He/She will eat'	'We will eat'

mɛ a+di+áná	*nzɛ́ a+di+áná*	*ndɪ́ á+di+áná*	*bɪ́ a+di+áná*
'I eat [HABIT]'	'You eat'	'He/She eats'	'We eat'

A similar pattern without a canonical Bantu subject-verb agreement obtains also in the Kiwumbu dialect of Kiteke, even though it has a singular/plural distinction and clear third person plural marker in the verb form:[3]

(3) Sample Conjugation in Kiwumbu (Kiteke):

mè (à)lyɛ:l	*nzè àlyɛ:l*	*ndɛ àlyɛ:l*
I cry	you-SG cry	he/she cries

bɪ́ lìlyɛ:l	*byɛn lìlyɛ:l*	*bo balyɛ:l*
we cry	you-PL cry	they cry

A closer examination of facts in Kimanyanga reveals that Kiyansi and Kiwumbu-Kiteke are not so unique. Kimanyanga does not follow the Bantu canon so strictly either in regard to subject-verb agreement, as illustrated below. There are a few cases where there is a null marker, which may well be interpreted as absence of agreement, at least from the point of view of a Kiyansi speaker:

(4) Illustration of Subject-Verb Agreement in Kimanyanga:

    a. *Mbuta muuntu w+e+zieet+a*   *kuku,*
       Old   man   AGR+ANTER+travel+Ending here
       ∅+mwen+i   *ki+ina*   *ki+ndumba-ndumba.*
       AGR+see+NEAR PERF. AGR7+that CL.7+girl-girl
       'My old man traveled this way [and] saw that young woman over there.'

b. *Ma+lavu ma+tanu Ɵ+bong+ele me+eso*
   CL6+wine CL6+five take+NPERF CL6+eye
   '[Those] five [bottles of] wine are before your eyes.'

One of the characteristics of Kituba, which has no incorporated object pronominal markers either, is that it uses the same free pronominals in both subject and nonsubject functions. One depends thus on the syntactic position of the pronominal to determine its function, as below:

(5) Distribution of Pronominals in Kituba:

    a. *móno mé(n)e móna yô.* 'I have seen it.'
       me   PERF see it

    b. *yó búl+a móno.* 'It hit me.'
       it hit+NARR me

    c. *béno món+a béto.* 'You saw us.'
       you-PL see+NARR we-PL

    d. *béto món+a bô.* 'We saw them.'
       you-PL see+NARR them

Kiyansi is also like Kituba in some related respects, especially in also allowing usage of independent object pronouns in several contexts, even though in some of these the pronouns cooccur with an incorporated object pronominal marker, as below:

(6) Object Pronominals in Kiyansi:

    a. *mɛ m+ǒ+sámí.* 'I have told him/you.'
       I NPERF+you/him+tell

    a'. *mɛ nzɛ́ m+ǒ+sámí.* 'I have told you.'
       I you-SG NPERF+you+tell

    a". *mɛ ndí m+ǒ+sámí.* 'I have told him/her.'
       I him/her NPERF+him/her+tell

    b. *mɛ (bɛa:g),*[4] *m+ɛ̌+sámí.* 'I have told them.'
       I (them) NPERF+them+tell

c. *mɛ (bɛ́) m+ɛ̌+sámí.* 'I have told you.'
   I (you-PL) NPERF+you+tell

d. *bɛ́ (mɛ́) ma+n+sámí.* 'You have told me.'
   you-PL (me) NPERF+me+tell

e. *nzɛ́ (bí) m+ɛ̌+sámí.* 'You have told us.'
   you-SG (we) NPERF+we+tell

f. *bí (ndí) ma+k+ō+sámí.* [k+ < kʸɛ 'go']
   We (he/she) NPERF+go+him+tell
   'We are leaving to tell him/her.'

g. *ndí (nzɛ́) ma+y+ō+sámí* [y+ < ya 'go']
   he/she (you) NPERF-come-you/him-tell
   'He/She has come to tell you/him.'

Note that in Kiyansi the free object pronominal plays an important role in clarifying reference, since incorporated pronominals are distinguished only in terms of first person singular vs. non-first person singular vs. plural. Because reliance on context alone may be quite confusing, it helps to disambiguate reference with concomitant use of free pronominals. By the same token, the free pronominals make the incorporated ones superfluous. Although Kimanyanga has a fully articulated system of distinctions for incorporated pronominals, it also seems to allow occasional usage (based on the Matuka corpus) of free object pronouns, whose usage, I suspect, may have increased in the contact settings that prompted and nurtured the development of Kituba:[5]

(7) Free and Incorporated Object Pronominals in Kimanyanga:

a. *bu n+tel+ele mono.* 'This is what I personally said.'
   this I+tell+NPERF me-EMPH

b. *mo+mo xu+n+tel+ele mono* 'what you told me'
   PL+what you+me+tell+NPERF me-EMPH

c. *u+m+bak+idi.* 'You got me/him/her.'
   you+me/him/her+get+NPERF

d. *mono mpe n+kaanda y+e+taanga wo.* 'The letter, I read it too.'
   me too letter I+PAST+read it

e. *mono bu yi+zol+anga.* 'I like [it] this way.'
   me this-way I+like+HABIT it

f. *tu+mwiin+i yo.* 'We have seen it.'
   we+see+NPERF

g. *mono mpe n+kutu we xoxa.* 'I too was about to say it.'
   me too AGR+about it say

To be sure, there is not as much flexibility in Kimanyanga as in Kiyansi in the use of free object pronouns. However, compared to how creoles lexified by European languages have selected most of their grammatical features (Mufwene 1989a, 1991), one cannot deny the significance of the option to use free object pronouns in Kiyansi, in Kimanyanga (although it seems more constrained here), and perhaps in other Bantu languages of the region.[6] As a matter of fact, the selection of free object pronouns over their agglutinated counterparts may have helped solve the apparent homonymy between, for instance, *n-* in (7b) and *m-* in (7c). Since the nasal is not syllabic, the phonological variation in this case may be considered no more than homorganic assimilation triggered by the following consonant.[7] Note also that the absence of an incorporated object in (7a) might be due to a constraint against doubling the same morpheme, *n-* 'I' and 'me', before the verb stem. Thus, syntactic marking eliminated some complications in the grammar.

Another respect in which Kituba diverges from the Bantu canon is lack of an incorporated reflexive form, as illustrated below. Because Lingala has one, it seems natural to wonder whether this peculiarity of Kituba's is a Bickertonian innovation.

(8) Reflexivization in Kituba:

a. *Yándi lwad+ís+a yándi mósi.* 'He/She cut him-/herself.'
   he/she cut+CAUS+NARR he/she one

b. *Bó lwad+ís+a bó mósi.* 'They cut themselves.'
   they cut+CAUS+NARR them one

(9) Reflexivization in Lingala:

a. *A+mi+tumb+í móto.* 'He/She burned him-/herself.'
   he/she+REFL+burn+NPERF fire

b. *Ba+mi+tumb+í móto.* 'They burned themselves.'
   they-+REFL+burn+NPERF fire

Kituba's periphrastic reflexive marker is a grammaticization of the quantifier 'one', which is sometimes used also with the meaning 'alone' and, in addition, serves as an emphasis marker, as in (10):[8]

(10)a. *Béto mon+áka mu+ntu mósi.* 'We saw one person.'
     we see+ANTER CL1+person one

   b. *Móno/Béto vand+áka móno/béto mósi.* 'I was/We were alone.'
      me/we LOC-be+ANTER me/we one

   c. *Móno/Béto mósi mu+ntu/ba+ntu sal+áka yo.* 'I/We-EMPH did it.'
      me/we   CL1+person/CL2+person do+ANTER it

While varieties of ethnic Kikongo are usually reported to have an incorporated reflexive marker, such as illustrated in (11) from Mufwene (1988), I have not identified any yet in Matuka's Kimanyanga text. On the other hand, I have noticed that Kiyansi-Kibwaal does not have one at all, as illustrated in (12). The pattern in the latter language is not significantly different from that in Kituba, though it makes a morphological distinction between the reflexive and the 'alone' functions. Note incidentally that the emphatic and reflexive functions are performed by the same morpheme *ŋgáàkwɛ̄ag* similar to pronominal compounds with *-self* in English for that matter:

(11) Reflexivization in Kikongo:

   a. *Yandi ka+ki+zol+ele.* 'He/She likes him-/herself.'
      he/she AGR+REFL+like+NPERF

   b. *B+ana ba+ki+zol+ele.* 'The children like themselves.'
      CL2+child AGR+REFL+like+NPERF

(12) Reflexivization in Kiyansi:

   a. *ndí á+tyɛ̌:n ι ndí ŋgáàkwɛ̄ag.* 'He/She is talking to him-/herself.'
      he/she NARR+talk CONN he/she EMPH.SELF

   b. *ndí ŋgáàkwɛ̄ag y+ǐ:.* 'He/She him-/herself came.'
      he/she EMPH-SELF come+NPERF

   c. *ndí á+tyɛ̌:n ι ndí ntu.* 'He/She is talking alone.'
      he/she NARR+talk CONN he/she alone

In a somewhat different vein, I have shown elsewhere (especially Mufwene 1988, 1989b, 1991) that Kituba's time reference distinctions are semantically more or less the same as in varieties of ethnic Kikongo.[9] The same applies to mood distinctions. As noted much earlier by Fehderau (1966), the tense-mood-aspect (TMA) markers themselves have generally come from Kimanyanga, although it is still not clear how to account for the change from the suffixal form -*anga* to -*áka* for the habitual aspect.[10] Because it is commonly assumed that Bantu languages typically mark time reference by agglutination, it seems very appealing to invoke some grammaticization processes, which must have been concomitants of the development of Kituba, to account for the current grammatical functions of the auxiliary verbs. Without rejecting offhand this explanation (which is consistent with work by especially Heine and his associates at the University of Cologne, e.g., Heine et al. 1991), grammaticization would have to be interpreted loosely here, in the sense of extended or regularized usage of these verbs for grammatical functions that were not unattested in the lexifier, as well as in the sense of their selection over their agglutinated alternatives, but not in the sense of newly formed grammatical patterns. More or less the same verbs in more or less the same periphrastic constructions are available in Kimanyanga, and/or in other varieties of ethnic Kikongo, for more or less the same functions, as below. Moreover, Kituba has followed their pattern in not marking the main verb in the infinitive as one of the options (Mufwene to appear-a).[11]

(13) Periphrastic Verbal Delimitative Constructions in Kimanyanga:

    a. *Mono ng+yena fila nkaanda.* 'I am <u>going to send</u>/sending a letter.'
       me AGR+go-PRES send letter

    b. *Nduumba yiyi y+ina teetuka.* 'This young woman is blooming.'
       young-woman this AGR+be-PRES come-out-of-shell

    c. *W+eta ku+n+tuma.* 'He is sending me.'
       he+be INFIN+me+send

    d. *K+eta pelesa.* 'He is rushing.'
       he+be rush

    e. *K+i+leend+i teka mw+ana ko.* 'I dare not sell [my] child.'
       NEG+I+dare+NPERF sell CL1+child NEG

    f. *Mono ng+ina kala toko dy+ani* 'I am (going to be) her husband.'
       me AGR+be-PRES be husband AGR+POSS

The following options were cited in Mufwene (1991) from Fehderau (1966):

(14) More Periphrastic Verbal Delimitative Constructions in Kimanyanga:

a. *W+eti dia.* 'He/She is (busy) eating.'
   he/she+be-PRES eat

b. *Ø+kedi dia.* 'He/She was (busy) eating.'
   +be-NPAST

c. *W+a+kedi dia.* 'He/She was indeed (busy) eating.' RPAST

d. *U+mana dia.* 'He/She has eaten.'
   he/she+finish eat

e. *Ø+meni dia.* 'He/She has eaten.' NPAST

f. *W+a+meni dia.* 'He/She has eaten.' RPAST

(15) Periphrastic Verbal Delimitations in Kituba:

a. *Móno mé(n)e 'dia.* 'I have eaten.'

b. *Móno ké(l)e 'dia.* 'I am eating.'

c. *Móno vand+á(k)a 'dia.* 'I was eating.'[12] [*vánda* = 'sit']

d. *Móno lénda 'dia.* 'I can eat.' [*lénda* = 'be able to']

e. *Móno tond+á(k)a 'dia.* 'I would have eaten.' [*tónda* = 'want']

Regarding time reference and mood markers, Kiyansi uses a system that is as heterogenous as in ethnic Kikongo, resorting both to agglutination and to periphrasis. This may be noted from the examples cited in (2). The patterns are very similar in both languages, even with respect to the variable position of tense markers (see Mufwene 1988, 1989b, and, to appear-a, for ethnic Kikongo). The division of labor that now obtains in Kituba, with aspectual and mood markers being free and preverbal while the tense markers are suffixes, should not be surprising and must have been a happy solution, producing more regularity.

I could go on showing how perhaps most morphosyntactic features of Kituba are developments from features of Kimanyanga, other varieties of ethnic Kikongo, or other Bantu languages involved in the ethnolinguistic contacts that produced it. However, space is limited and I should hasten to highlight the relevance of these data to the debate on the development of Kituba, at the intersection of history and language contact, and to genetic aspects of Bantu linguistics. Before I

do this, however, it will help to show that not all selections made by Kituba originated from the lexifier or its closest kin alone. Examples of influence from outside the Kikongo group may be cited in Head + Modifier constructions. Unlike in ethnic Kikongo, the form of the connective in Kituba does not vary much in such constructions, as illustrated below:

(16) Head + Modifier Relation in Kimanyanga's NP:

    a. *bi+bundi by+a bi+oole* 'two six-yard cuts of cloth'
       CL8+six+yard cut CL8+CONN CL8+two

    b. *mi+inda mi+e Koleman* 'Coleman lamps'
       CL4+lamp CL4+CONN

    c. *fuunda dy+e n+zitu* 'in-law's packet'
       CL5a+packet CL5+CONN in-law

(17) Head + Modifier Relation in Kituba's NP:

    a. *mw+ána ya 0+máma* 'mother's child'
       CL1+child CONN CL1a+mother

    b. *b+ána ya 0+máma* 'mother's children'
       CL2

    c. *m+bísi ya 0+máma* 'mother's fish'
       CL9/1a+fish

    d. *di+kulu ya 0+nkénto* 'left leg'
       CL5+leg CONN CL1a+woman

    e. *ma+kulu ya 0+nkénto* 'left legs'
       CL6

    f. *mw+ána na béto/béno* 'our/your child'[13]
       CL1+child CONN we/you-PL

This development in Kituba is the kind of simplification that would dispute the kind of explanation I summarize below, from especially Mufwene (1991), if all Bantu languages followed the Bantu morphosyntactic canon presented in much of the literature. However, Kiyansi diverges again from this misleading stereotype, as shown below:

(18) Head + Modifier Relation in Kiyansi's NP:

    a. *mw+á:n í mâ:*   'mother's child'
       CL1+child CONN mother

    b. *b+á:n í mâ:*   'mother's children'
       CL2+child CONN mother

    c. *m+biy í mâ:*   'mother's fish'
       CL9/10+fish CONN mother

    d. *∅+kwū:l í ká:r*   'left foot'
       CL5a+foot CONN female

    e. *m+yě:l í ká:r*   'left feet'
       CL6+foot CONN female

I argue in Mufwene (1990) that, while the invariance of the connective may have been selected from outside the Kikongo group of languages, the form *ya* itself seems to have come from the second half of the variable connective in the lexifier itself and its immediate kin. While the consonantal beginning of the connective varies a lot, reflecting agreement, the segment following it is [y] (represented graphically as *y* or *i*), and the following/final vowel is predominantly [a].

I could thus argue safely that, like in other creoles (or call them simply contact-induced language varieties), very few morphosyntactic features of Kituba, if any, are innovated in the sense suggested by, for instance, Bickerton (1984, 1988, 1992). There are undoubtedly innovations in the traditional historical linguistic sense, according to which, in the case of the aspectual markers of Kituba discussed here, the systemic distribution of some constructions has been extended and their function generalized, but nothing in the Bickertonian sense of creation virtually ex-nihilo.[14]

The position I present here is based on the assumption that creoles are mixed languages, as was advocated before me by, for instance, Schuchardt (1909), Hjelmslev (1938), and, recently, Thomason and Kaufman (1988). Thus their grammatical features need not have come from one language variety only, even though their morphemes, especially the grammatical ones, may have a predominant common source, as argued in Mufwene (1986). Creoles qua restructured varieties need not be any more homogeneous than other languages, which have often been restructured in various ways and to various extents in the course of their histories, Hjelmslev (1938) argues quite convincingly.[15] Having articulated this position, we may now re-examine the facts with regard to the contact-based development of Kituba and, later, relative to genetic aspects of African linguistics.

3. On the Development of Kituba

There are two competing hypotheses on the development of Kituba. According to the first, assumed by Fehderau (1966), Kituba started in the sixteenth century as a koiné out of diverse mutually unintelligible language varieties of the Kikongo Group. Although this position may seem plausible, given the existence of the Kongo Kingdom at that time, Samarin (1982, 1990) has undermined it successfully with two main objections: 1) It was not necessary for a koiné of any sort to develop in order for different African ethnolinguistic groups to communicate with each other. According to him, Africans have generally been associated with the facility to learn their neighbor's language, a disposition that has resulted in widespread individual multilingualism on their continent. Further evidence for this observation lies in the fact that in the course of their colonial expeditions Europeans recruited Africans as their interpreters, apparently not in the least concerned by geographical distance and concomitant systemic differences.

2) If there had been a koiné since the sixteenth century, this would have been reported in historical accounts, like several other facts about human contact in precolonial Africa. For instance, colonial agents typically identified those ethnic languages that were widely spoken as lingua francas and adopted them, as in the case of Bobangi, the primary lexifier of Lingala, according to Samarin. In the same way, Kimanyanga, the language variety of Manyanga, an important trade post inland on the Zaire River rapids, rather than any of the coastal language varieties, was adopted as a lingua franca for the Western part of Zaire, then the Congo Free State.[16]

I would like to add a third objection: 3) if koineization or pidginization was a common solution in precolonial Africa, there would have been many more cases similar to the central African contact phenomena represented by Kituba, Lingala, and Sango, certainly given the precolonial existence of several noteworthy empires and kingdoms on the continent. Nurse (1985) argues convincingly that even Swahili, which is assumed by some scholars (such as Ohly 1982) to have had a pidgin/creole origin, did not start as one and was only adopted by the Arabs, without grammatical restructuring, for trade in East Africa.

According to the second hypothesis, assumed by Samarin (1982, 1989, 1990), the development of Kituba was a concomitant of colonization, dating from the end of the nineteenth century, with the Berlin Treaty in 1885 and the formation of the Congo Free State serving as a convenient historical milepost. To paraphrase him, the ethnographic circumstances of the formation of this creole, originally a pidgin, involved the encounters of, on the one hand, West African languages with local Bantu languages, and, on the other, Bantu languages among themselves in colonial posts, originally in Bas-Zaire and later, as colonization spread toward the hinterland, in the Bandundu Region south of the Kasai River.

As reported by Samarin (1982, 1989, 1990), the West Africans, recruited as militia men, clerks, interpreters, and porters for the Belgian colonial agents, in-

cluded Senegalese (speakers of West Atlantic and Mande languages), Kru, Hausa, and several speakers of Kwa languages.[17] The Bantu populations not speaking any of the ethnic Kikongo varieties[18] came from as far East as Zanzibar, brought over especially to build the railroad from the Atlantic Ocean to Kinshasa, or to serve as militia men, cooks or porters. According to Samarin, 'they were the first most important workers of [Sir Henry Morton] Stanley and his successors [in the nineteenth century] and for quite some time they were the principal recruits of the "force publique"' (p.c. 1994). Many words of Arabic origin may have found their way into Kituba through these East African recruits!

I conjecture that although Kituba's development started in colonial posts in the Lower Zaire Region, where language varieties of the Kikongo Group are indigenous, the process itself entailed selection from among competing grammatical structures. The colonial *centres extra-coutumiers*, beginnings of present-day cities, started with the multilingual indigenous populations forming the overwhelming majorities. In these new ethnographic settings, the local Bantu languages must have led the way in the competition of features, with their speakers becoming the models for the West African colonial escorts. According to the ecological-ethnographic approach to creole genesis that I have advocated over the past few years (e.g., Mufwene 1991, 1992a, 1993, to appear-a), factors such as the ethnolinguistic composition of the colonial posts and the demographic proportions of speakers of the different languages (on which there is very little information in the case of Kituba) are as significant as both the systemic options available in the lexifier and the alternatives provided by the non-indigenous languages. In other words, markedness values are not determined by structural considerations alone, even though these are heavily weighted.

These considerations suggest that, because speakers of the local Bantu languages were likely to be in the majority, their features were just as likely to prevail as any others that may appear to be less marked from a purely structural point of view. We are reminded of this possibility especially by Melanesian pidgins in which features such as the inclusive/exclusive distinction in the pronominal system have been retained from the convergent substrate languages.[19] However, since, as shown above, the indigenous Bantu languages in contact do not always follow the Bantu canon suggested in the literature, selection must have been more manifold than traditionally assumed. Both purely structural factors and the demographic-ethnographic factors involving speakers only may have converged variously to produce the Kituba phenomenon. Still it may not be necessary to invoke languages from outside the Bantu area to account for Kituba's selection of exclusive periphrastic delimitation for aspect, mood, and pronominal arguments.

Contrary to what might be concluded from the above, I disagree only partly with Samarin's (1990) position that the West African colonial workers were the primary agents of Kituba's formation. Although I questioned this view in Mufwene (1989b), I entertain it in Mufwene (to appear-a), arguing that the West Africans may be considered the initiators of the restructuring. As the Bantu-

speakers adopted the West Africans' second-language variety of Kimanyanga (probably assuming in areas outside the Kongo territory that it was the colonists' language), they must have played an important role in making this ancestor of Kituba as Bantu-like as possible with respect to, for instance, TMA semantic distinctions, noun class morphological distinctions, and verbal extensions. They must have done this in the same way that Melanesians restructured English, not only omitting some native distinctions but also assigning it new ones.

Assuming that the Kru and other West Africans from the coast must have spoken some pidgin English, which may have served as the means of communication between them and the Belgian colonial agents in lieu of English,[20] Kituba would probably have been much less agglutinating than it is, had it not been for the intervention of Bantu-speakers in its development. The contribution of Bantu-speakers must have become more important during the formation of colonial posts, in which they came to work, having been brought from diverse parts of Zaire east of the Kwango River and from Zanzibar. In Mufwene (to appear-a), I reinterpret Samarin's position as follows:

'Samarin's hypothesis is thus not as implausible as it may seem. The colonial posts, the factory towns, and the missions, with which Kitúba's development has been associated, provide the socio-historical setting. The West Africans as interpreters between the Europeans and the Bantu populations were instrumental in the development of especially the colonial posts and factory towns. Their association with colonial power must have made the contact language they initiated a viable alternative for communication once identified by the indigenes as a lingua franca. Its usage in all the colonial posts where the West Africans and others carried it must have given it some relative universal usefulness that later on relegated the Bantu languages that hitherto had functioned as lingua francas to a lower ethnographic status, especially in the colonial posts.

It may be assumed that, despite the similarity, at least lexical, of the initial-stage Kitúba—as used by the West Africans—to Kikóngo and to some of the languages around the area, the Bantu populations, including the Bakóngo, just assumed that it was colonial speech and largely accommodated the powerful foreigners. On the other hand, they must have not been passive language learners in this kind of setting. While accommodating the colonial, European and West African, foreigners, they must have brought the new language closer back to the Bantu patterns. In turn, the West Africans must have adjusted back to the variety spoken by the Bantu speakers. Whether or not things worked out this way remains a speculation difficult to verify beyond doubt in the absence of a sample of West Africans' Kitúba and, generally, of early linguistic records. One particular factor encouraging the above division of labor between the West Africans and the Bantu is that, despite its development toward an isolating

morphosyntax (Mufwene 1988), Kitúba is still in several ways more different from Kwa, or West Atlantic, or Mande languages than it is from Kimanyánga or any of the surrounding Bantu language varieties, e.g., regarding the noun class and time reference system (Mufwene 1988, 1990a).'

One of the points I wish to make with the data in Part 2 is that with or without the initial instrumental role of West Africans in the colonization of the Congo Basin at the turn of the century, a Kituba-like creole could still have developed, provided the other ethnographic conditions remained the same. Every morphosyntactic feature of Kituba discussed here and in Mufwene (1988, 1989b, 1990b, 1991, to appear-a) can be related to some option in Kimanyanga or some other indigenous Bantu language(s), represented here partly by Kiyansi. Note that the selection is hardly a perfect copy, as it is generally also accompanied by some partial innovation in the traditional historical linguistic sense (Boretzky 1993). I will go as far as claiming, as in Mufwene (1991), that the selections are generally in favor of some unmarked option, with 'unmarked' being a descriptive shorthand for a variety of more explanatory factors, such as perceptual salience, semantic transparency, morphological uniformity, morphosyntactic regularity, convergence between the lexifier and other languages in contact, among other factors.

Markedness values in this model are determined relative to other competing parametric options in the specific ethnographic setting of language contact rather than in Universal Grammar. There are cases where these factors may be in conflict with each other, suggesting different options as unmarked. Such conflicts may be resolved either by weighting the factors in ways that still remain to be worked out or by considering how they cluster to favor some particular options.

The latter part of this statement explains, for instance, why Kituba has selected the free-pronoun option for anaphoric objects over the incorporated alternative, especially since Kiyansi and probably most of the other local Bantu languages do have the latter option, which may be claimed to favor Hagège's (1992) 'ease of production' principle. As noted above, the incorporated alternative was likely to refer ambiguously in some cases, due to homonymy of some forms. This state of affairs is likely to favor alternatives which are both salient, unambiguous, and uniform, all of which are embodied in the free-pronoun option. In any event, an important advantage of this approach is that it accounts for cases where more than one option is selected, reflecting occasional natural inability to resolve a structural-option conflict among those who developed the new language.

My hunch is that this model accounts adequately for how creoles have selected their structural features from competing alternatives in the contact settings in which they developed. In a way that is quite consistent with the history of contacts, it also marginalizes the role of children in the vernacularization and normalization of such varieties. One would have to be familiar with some of the competing options to play a role in the selection of features!

There are also several established assumptions about language that one would have to abandon or modify in order to adopt the above approach. We may start with the realization that while language is transmitted from one generation to the other, it is also constantly being made by its speakers (Hagège 1992). Language is an exponential working construct; the reality is speech, from which the construct 'idiolect' was formed. In a contact setting, one can imagine several individual speech varieties competing with each other, up to the point where patterns of the most successful attempts prevail. For instance, construction patterns that reflect more regularity, transparency, and relative simplicity are likely to be selected over other alternatives.

Some creolists, such as Thomason and Kaufman (1988), have advocated accommodation as an important explanation. I do not see accommodation and selection as mutually exclusive. In fact I see them as working together, with selection probably accounting for the direction of accommodation processes. Note that accommodation is an interactional notion that does not explain which structural options get selected and for what structural or ethnographic reasons.

4. Implications for Genetic Aspects of African Linguistics

We may now address the question of how relevant this discussion of the development of Kituba is to genetic aspects of Bantu linguistics. Let us start with the fact that Kituba has generally been omitted in classifications of Bantu languages. For instance, neither Bendor-Samuel (1989) nor Bright (1992), two of the latest major references in which one might want to check how Kituba is classified genetically, lists it among the Bantu languages under any of its many names.[21]

There are several conceivable reasons for the above state of affairs, one of which is that Kituba may simply have been confused with other language varieties subsumed by the name *Kikongo Group*. This interpretation is consistent with Hjelmslev's (1938) position that creoles are dialects of their lexifiers despite the putatively extensive restructuring that has produced them and the fact that their systems have mixed sources.[22] This particular view of the facts makes it acceptable to disregard some of the respects in which Kituba differs from other members of the Kikongo Group, for instance, in not having much of the inflectional agglutination (including subject-verb and head-modifier agreement) typically associated with Bantu languages. However, it does not explain the inconsistency in several genetic classifications which do not list Kituba along with other varieties of Kikongo such as Kiyombe, Kifioti, Kintandu, or Kimanyanga, its lexifier.

Things become complicated if one considers the fact that Kituba speakers do not generally perceive much kinship, if any, between their language and those of the Kikongo Group. It is not even certain that the Bakongo themselves consider Kituba to be one of the related language varieties identified by outsiders indiscriminately as Kikongo. We must thus wonder why a language variety recognized

as a major urban vernacular and as a lingua franca for millions of speakers in rural areas of Bas-Zaire and the Bandundu Region south of the Kasai River has been overlooked in genetic classifications.

A second and most likely reason for not listing Kituba seems to lie in the tradition of genetic linguistics since the nineteenth century to focus on what may essentially be claimed as pure parent-to-offspring relations, rather than in mixed languages. Thomason and Kaufman (1988:3) in fact reiterate the nineteenth-century position in claiming that:

> **mixed languages** [such as pidgins and creoles] do not fit within the genetic model and therefore **cannot be classified genetically at all; but most languages are not mixed,** and the traditional family tree model of diversification and genetic relationship remains the main reference point of comparative-historical linguistics owing to the fact that it is usually possible (except in relatively rare borderline cases) to distinguish mixed languages, whose origins are nongenetic, from languages whose development has followed the much more common genetic line (T&K's emphases).

Like most genetic linguists, they perpetuate an inadequate analogy from biological taxonomies, regardless of the fact that one of the concerns of biology itself is how to account for features of hybrid species, which are in fact considered a normal occurrence. For most species, the normal offspring is a hybrid, which makes the genetic linguistics tradition inadequate. Schuchardt (1909) and Hjelmslev (1938) were justified in disputing the treatment of pidgins and creoles as abnormal cases, and in arguing that all languages are mixed to some extent.

I should clarify that Schuchardt's and Hjelmslev's position that creoles are mixed languages, only to a larger extent than their non-creole counterparts, is not quite the same as Mufwene's (1992b) arguments that linguistic systems are not monolithic. While Mufwene highlights inconsistencies and the fact that rules overlap, independently of the histories of individual languages, Schuchardt and Hjelmslev meant what is suggested by the history of contacts that are characteristic of the developments of, e.g., the Romance languages. This sense can also be extended to other contacts in Europe that have affected the structures of several languages, as well as to the migrations that account for the current geographic distribution of Bantu languages in sub-Saharan Africa. Their position refers to the diverse sources of structural features in particular languages.

In the case of the Bantu languages, their speakers moved into territories that were inhabited by other people. They did come in contact with these populations while pushing them away. During the migrations, the Bantu (sub-)groups must also have come into contact with one another and thus must have influenced each other. Assuming that present-day differences are not simply the result of their divergent migrations, variation within Proto-Bantu and selection in the direction

of present-day crosslinguistic variation are factors that must be taken as seriously in Bantu genetic studies as the features of the substrate languages themselves.[23]

So, while it is convenient to assume that all the Bantu languages stem from the same ultimate Proto-Bantu, through the misguided model of one parent for any number of offspring, we cannot ignore the likelihood that the proto-language must have allowed internal variation, another important assumption for the evolution theory in population genetics. One of the consequences of this assumption is that some of the features that now distinguish Bantu subgroups from one another may follow from Proto-Bantu itself, by a process similar to speciation (involving selection, with the helping hand of ecology[24]) in biology. Contact among the Bantu languages themselves after their respective contacts with the substrate languages complicates the speciation patterns. Could the variation and systemic inconsistencies shown above in Kiyansi and Kimanyanga follow from such manifold contacts? Whatever the situation is, genetic relationships may not always be as straightforward and exclusive as we would like them to be. Not being a genetic linguist myself, I can only wonder now, at this initial stage of connecting my research on Kituba's development with the genetic linguistics enterprise, to what extent contact has been taken into account in the latter's agenda.

Lexicostatistics seems to have played a very important role, over morphosyntactic considerations, in the genetic classification of Bantu languages. The practice has, nonetheless, led most of the scholarship to infer morphosyntactic typological kinship from lexical-genetic kinship. As the research on Kituba's development shows that there may be more typological variation, even within individual languages, than may have been assumed in the literature, it is also conceivable that Proto-Bantu may have allowed not only grammatical but also lexical variation. It may help for geneticists to explain to the rest of us not only why lexicostatistics is weighted so heavily compared to other potential measures of kinship, but also why Proto-Bantu has been assumed to be so monolithic. Is it not conceivable to think of Proto-Bantu as a group of lexically and typologically related language varieties that were not necessarily identical?

The above comments and questions follow naturally from research on language contact. The reason for formulating them here and showing the implications of studying Kituba's development for Bantu genetic linguistics is precisely to highlight the not-so-marginal role that contact must have played during the spread of Bantu and its speciation into different subgroups. It was also to suggest the possibility of variation in the putative Proto-Bantu system, suggesting that pure innovations qua creations ex-nihilo were probably very limited during the speciation processes.

Notes

*. I am grateful to Yeno Matuka, Jerry Sadock, and Bill Samarin for comments on the conference version of this paper. I assume full responsibility for all the remaining shortcomings.

1. To the extent that Nigerian Pidgin English is used by some speakers as a vernacular, it is as much a creole as Krio in Sierra Leone and their counterparts in the New World.

2. The following abbreviations are used in this paper: AGR = subject agreement prefix, ANTER = anterior tense, CL# = class number, CONN = connective, DUR = durative, EMPH = emphatic, FUT = future, HABIT = habitual, NARR = narrative, NPERF = near perfect, RPERF = remote perfect, SG = singular, PL = plural, OBJ = object marker, NPAST = near past, RPAST = remote past, REFL = reflexive.

3. The Kiwumbu examples are from Mputubwel Makim (p.c.), to whom I am very grateful. The Kimanyanga examples are from Matuka (1990). He provides no tone diacritics, a shortcoming that does not reduce the usefulness of the corpus. The glosses and translations have occasionally been adjusted when I thought his analysis was inaccurate.

4. The graphic sequence [ɛa] is a diphthong in which [ɛ] is weaker than [a]. Rottland (1970) treats this [ɛ] as a glide. However, this putative diphthong combines with other glides, as in [mbwɛág] 'road'.

5. I argue in Mufwene (1989a, 1991) that in ethnographic settings which produce creoles free morphemes are generally selected over their bound alternatives because they are more salient, even if the bound ones do not violate Seuren and Wekker's (1986) principle of semantic transparency.

6. I have not yet been able to verify this feature in Kiwumbu (Kiteke).

7. According to Matuka (p.c. 1994), the prefix $m+$ in (7c) is syllabic and 'heavy' when meaning 'him/her'. Still I wonder how clear the distinction is to speakers of languages, such as Kiyansi, which do not have syllabic nasals at all or might have them only in word-initial positions.

8. According to Jerry Sadock (p.c. 1994), Yiddish right-adjoins the morpheme *aleyn* 'alone' to the reflexive form *zix* (from German *sich*) to mark semantically transparent reflexives. Kituba may thus not be unique in selecting this grammaticization path.

9. The Matuka texts suggest that Kimanyanga may make more subtle PAST (or ANTER) distinctions than I originally thought. This observation, which is consistent with the best known cases of creolization, does not, however, affect the morphosyntactic point made below.

10. Fehderau (1966) conjectures that Kituba's habitual suffix $+áka$ has come from Lingala. As discussed in Mufwene (1988), this analysis faces some problems, in part because it is not consistent with the fact that the two creoles developed around the same time. In fact, if Samarin (1982) is right, Kituba must

have developed a little bit earlier and may have influenced the development of both Lingala and Sango. There is also the problem of why Lingala's LL tone pattern would have changed to a HL pattern in Kituba. This is significant because Lingala words have generally maintained their tone pattern in Kituba, as shown in Mufwene (1989b).

11. Both Kimanyanga and Kiyansi nominalize verbs also by not marking them with any prefix at all, as in the Predicate-Cleft construction *diá ká+dià* 'He/She ate RPAST' in Kimanyanga or the complex deverbal noun phrase *diá kó adiá ndí la* 'the way he is eating' in Kiyansi (Mufwene 1987).

12. As explained in Mufwene (1988), *vánda* (from Kimanyanga *vwanda*) 'LOCA-TIVE be, sit' is used as the past allomorph of *kéle* 'be'.

13. The connective becomes *na* before personal pronouns. This variation, the only type I can remember, is still simple compared to the complex kind of variation in Kikongo, which is conditioned by the class of the head noun. The simplification in Kituba is of course consistent with the general loss of agreement discussed in Mufwene (1988, 1989b, to appear-a).

14. I derive this seemingly uncharitable interpretation of Bickerton's thesis from his association of his innovations with the creation of the relevant creoles putatively by children.

15. According to Hjelmslev, after two or more languages have competed with each other, the one that prevails does so often at the cost of incorporating several elements from its competitors. It wins only a Pyrrhic victory, so to speak, undergoing a fate similar to the lexifier of a creole language.

16. As often noted by Samarin, Manyanga also played an important role in the recruitment of porters for colonial explorations from the Atlantic coast during the Congo Free State.

17. In Zaire West Africans are often referred to indiscriminately as *Sénégalais*.

18. Ngalasso (1991) observes correctly that speakers of ethnic Kikongo varieties claim they speak separate languages, which is consistent with Fehderau's observation that they are not mutually intelligible. The term *Kikongo* is used either by outsiders to refer to the cluster of languages or occasionally by speakers of these varieties to contrast their group of languages with those of others, by analogy to the ethnic term *Mu+/Ba+kongo*. Otherwise they typically introduce themselves as *Mu+/Ba+ntandu, Mu+/Ba+yombe, Mu+/Ba+ladi*, etc.

19. I am also reminded here of Copper Island Aleut, which developed at the beginning of this century out of the contact of several Aleut dialects and Russian (Golovko and Vakhtin 1990). Being an endogenous contact-based language variety, with Aleut as the lexifier, most of its grammatical patterns and morphemes have been selected from Aleut. As in Kituba and other creoles, there is little evidence, if any, of Bickertonian innovations. A similar case is that of Nubi, which, according to Owens (1990), derives many of its grammatical features from (Sudanese) Arabic, its lexifier, and a few from the substrate (originally Chadic and Easteran Sudanic) languages.

20. Samarin (1989:34) reports that 'many of the Kru [and probably other West Africans from British territories] were probably speakers of pidgin English. Indeed, it was said that they all spoke a little English, and some even spoke some French (Valcke 1886:29).' Regarding the importance of English in the colonization of the Congo Free State, he reports: 'That English was an important lingua franca of the Belgian foreign work force is seen in the fact that Zanzibari were addressed in it at the military instruction camp at the equator ([*Le Congo Illustré*] 1892:186' (1989:34). Samarin also explains that indeed the Belgians recruited most of their non-indigenous African workers from British rather than French territories, thus clarifying the privileged position of English as a lingua franca between the Belgians and their interpreters at least during the Congo Free State stage.

21. To be sure, Rickford's article in Bright (1992) lists Kituba among creole languages but, because of its focus, it does not link it to Bantu languages.

22. Extent of restructuring is a debatable factor in deciding whether or not a contact-based variety is a creole, just as whether or not the term *creole* is used in discussions of language-based varieties changes little, if anything of significance, in investigating how they developed (Mufwene, to appear-b).

23. Orin Gensler has kindly brought to my attention Trubetzkoy's (1939) article which raises similar questions about positing a monolithic Proto-Indo-European (IE) as the ultimate root of IE languages. Note that Trubetzkoy thinks of this family not only as genetic but also as typological. From the latter point of view he claims that a language may become IE by acquiring IE features through contact, just as another may cease to be IE by losing IE features. Contact, with which he associates borrowing, plays an important role in his view of language classification.

24. Ecology in this model remains a necessary, though still elusive, notion. In this particular case, it subsumes the substrate languages, whose particular features are a relevant factor in determining those of the Bantu superstratum.

References

Bendor-Samuel, John, ed. 1989. *The Niger-Congo languages: A classification and description of Africa's largest language family.* Lanham: University Press of America.

Bickerton, Derek. 1981. *Roots of language.* Ann Arbor: Karoma.

Bickerton, Derek. 1984. The language bioprogram hypothesis. *Behavioral and Brain Sciences* 7.173-221.

Bickerton, Derek. 1988. Creole languages and the bioprogram. In *Linguistics: The Cambridge survey*, vol. 2: *Linguistic theory: Extensions and implications*, ed. by Frederick J. Newmeyer, 268-84. Cambridge: Cambridge University Press.

Bickerton, Derek. 1992. The creole key to the black box of language. In *Thirty years of linguistic evolution: Studies in honor of René Dirven on the occasion of his sixtieth birthday*, ed. by Martin Pütz, 97-108. John Benjamins: Amsterdam.

Boretzky, Norbert. 1993. The concept of rule, rule borrowing, and substrate influence in creole languages. In *Africanisms in Afro-American language varieties*, ed. by Salikoko S. Mufwene, 74-92. Athens: University of Georgia Press.

Bright, William, ed. 1992. *International encyclopedia of linguistics*. New York: Oxford University Press.

Chaudenson, Robert. 1979. *Les créoles français*. Paris: Fernand Nathan.

Fehderau, Harold. 1966. *The origin and development of Kituba*. Ph.D. dissertation, Cornell University.

Golovko, Eugeni V. and Nikolai B. Vakhtin. 1990. Aleut in contact: The CIA enigma. *Acta Linguistica Hafniensia* 22.97-125.

Guthrie, Malcolm. 1953. *The Bantu languages of Western Equatorial Africa*. Oxford: Oxford University Press.

Hagège, Claude. 1992. *The language builder: An essay on the human signature in linguistic morphogenesis*. Amsterdam: John Benjamins.

Heine, Bernd. 1970. *Status and use of African lingua francas*. Munich: Weltforum Verlag.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Hinnebusch, Thomas J. 1989. Bantu. In Bendor-Samuel 1989.

Hjelmslev, Louis. 1938. Etudes sur la notion de parenté linguistique. *Revue des Etudes Indo-Européennes* 1.271-286.

Matuka, Yeno Mansoni. 1990. *Three Koongo palavers: Wedding, bereavement and reconciliation*. Ms., Ball State University.

Mufwene, Salikoko S. 1986. Notes on durative constructions in Jamaican and Guyanese creoles. In *Varieties of English Around the World: Focus on the Caribbean*, ed. by Manfred Görlach and John Holm, pp. 167-82. Amsterdam: John Benjamins.

Mufwene, Salikoko S. 1987. An issue on predicate-clefting: Evidence from Atlantic creoles and African languages. In *Varia Creolica*, ed. by Philippe Maurer & Thomas Stolz, 71-89. Bochum: Brockmeyer.

Mufwene, Salikoko S. 1988. Formal evidence of pidginization/creolization in Kituba. *Journal of African Languages and Linguistics* 10.33-51.

Mufwene, Salikoko S. 1989a. Some explanations that strike me as incomplete. Column. *Journal of Pidgin and Creole Languages* 4.117-128.

Mufwene, Salikoko S. 1989b. La créolisation en bantou: Les cas du kituba, du lingala urbain, et du swahili du Shaba. *Etudes Créoles* 12.74-106.

Mufwene, Salikoko S. 1990a. Time reference in Kituba. In *Tense-modality-aspect systems in pidgins and creoles*, ed. by John Singler, 97-117. Amsterdam: John

Benjamins.

Mufwene, Salikoko S. 1990b. Creoles and Universal Grammar. In *Issues in creole linguistics*, ed. by Pieter Seuren and Salikoko Mufwene, 783-807. *Linguistics* 28.

Mufwene, Salikoko S. 1991. Pidgins, creoles, typology, and markedness. In *Development and structures of Creole languages: Essays in honor of Derek Bickerton*, ed. by Francis Byrne & Thom Huebner,123-43. Amsterdam: John Benjamins.

Mufwene, Salikoko S. 1992a. Africanisms in Gullah: A re-examination of the issues. In *Old English and new: Studies in language and linguistics in honor of Frederic G. Cassidy*, ed. by Joan Hall, Nick Doane, and Dick Ringler, 156-182. New York: Garland.

Mufwene, Salikoko S. 1992b. Why grammars are not monolithic. In *The joy of grammar: A festschrift in honor of James D. McCawley*, ed. by Diane Brentari, Gary Larson, and Lynn MacLeod, 225-250. Amsterdam: John Benjamins.

Mufwene, Salikoko s. 1993. Creole genesis: A population genetics perspective. Paper presented at the meeting of the Society for Pidgin and Creole Languages, University of Amsterdam. To appear in *Caribbean language: Issues old and new*, ed. by Pauline Christie. Mona, Jamaica: University of the West Indies Press.

Mufwene, Salikoko S. to appear-a. Kituba. In *Pidgin and creoles lexified by non-European languages* (?), ed. by Sarah Grey Thomason. Amsterdam: John Benjamins.

Mufwene, Salikoko S. to appear-b. Jargons, pidgins, creoles, and koinés: What are they? In *Pidgins and creoles: Structure and status*, ed. by Arthur Spears and Donald Winford. Amsterdam: John Benjamins.

Ngalasso, Mwatha Musanji. 1991. Tons ou accents? Analyse des schèmes tonals du kikongo parlé dans la région de Bandundu (Zaire). *Etudes Créoles* 14.147-162.

Nurse, Derek. 1985. *The Swahili: Reconstructing the history and language of an African society, 800-1500*. Philadelphia: University of Pennsylvania Press.

Ohly, Rajmund. 1982. *Swahili: The diagram of a crisis*. Vienna: Afro-Pub.

Owens, Jonathan. 1990. East African Nubi: Bioprogram vs. inheritance. *Diachronica* 7.217-250.

Rottland, Franz. 1970. *Die Verbformen des Yanzi*. Doctoral thesis, University of Leiden.

Samarin, William J. 1982. Colonization and pidginization on the Ubangi river. *Journal of African Languages and Linguistics* 4.1-42.

Samarin, William J. 1989. *The black man's burden: African colonial labor on the Congo and Ubangi Rivers, 1880-1900*. Boulder, Colorado: Westview Press.

Samarin, William. 1990. The origins of Kituba and Lingala. *Journal of African Languages and Linguistics* 12.47-77.

Schuchardt, Hugo. 1909. Die Lingua Franca. *Zeitschrift für Romanische Philologie* 33.441-61.

Seuren, Pieter and Herman Wekker. 1986. Semantic transparency as a factor in creole genesis. In *Universals versus substrata in creole genesis*, ed. by Pieter Muysken and Norval Smith, 57-70. Amsterdam: John Benjamins.

Thomason, Sarah G. and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Trubetzkoy, N. S. 1939. Gedanken über das Indogermanenproblem. *Acta Linguistica* 1.81-89.

Valcke, (Lt.). 1886. Description de la région des cataractes: de Vivi à Stanley-Pool. In *Conférence sur le Congo*. Brussels: Librairie Universelle de Veuve J. Rozez, 27-41.

Whitehead, John. 1899. *Grammar and dictionary of the Bobangi language, as spoken over a part of the Upper-Congo—West Central Africa*. London: Gregg International Publishers.

The Nature and Size of Linguistic Contact Possible in Three Centuries

Derek Nurse
Memorial University of Newfoundland

1. Introduction

Around AD 1600, the Orma, the southern tip of a huge Eastern Cushitic-speaking language /dialect chain with several million members stretching up into Ethiopia, moved aggressively into northern East Africa. They occupied the valley of the Tana River in Kenya, the country on either side, and spread south along the coast. Their occupation altered the linguistic landscape. For almost three centuries they surrounded and subjugated all the small communities living along the Tana, and subjected them to massive linguistic pressure (see Sasse 1979). These included the (Cushitic-speaking) Aweera (= Boni) and Dahalo, and the (Bantu-speaking) Pokomo and Ilwana (= Malakote): the Ilwana are one of the two targets of this discussion. Ca. AD 1875 communities speaking northern Somali arrived along the river and defeated the Orma in battle.The Orma retired to the west/south bank of the Tana and henceforth played a lesser role in the area. The changes in Ilwana resulting from Orma influence are fairly transparent and the state of Ilwana before contact with Orma can also be to some extent confirmed by comparison with its nearest relatives.

The other target of this discussion is the Daiso community who speak a Bantu language closely related to Kamba. Their early history is not quite clear, but in AD 1571 Daiso seem to have been at Malindi on the Kenya coast, twenty years later at Mombasa, and to have arrived shortly thereafter at their present location some fifty kilometers northwest of Tanga in northeastern Tanzania. Their language was first recorded there near the end of the 19th century. So if they are assumed to have split off from their Kamba brothers shortly before they were at Malindi, their language also had some three centuries to undergo the changes that now distinguish it from Kamba.

Here are two (Bantu) languages which have both undergone three centuries of disturbance, even the same three centuries. As their earlier state can be worked out or assumed, we can compare the kinds and size of contact-induced change possible over the same time period - a rather unique opportunity. As will be seen, the results and the circumstances leading to them are different.

2. Ilwana and Orma

2.1 Orma as a Language

Heine and Vossen (1980) contains a typological overview of Kenya's Cushitic languages, including Oromo, of which Orma is one dialect. Stroomer (1987) sets out a mass of detail on three Oromo dialects, including Orma. They form the basis for all typological and areal claims below.

In the NP of most Kenyan languages: N precedes possessive, adjective, numeral and demonstrative; the governing NP precedes the governed/genitive NP. Also the indirect object precedes the direct object. As a Kenyan Cushitic language Orma is SOV; AUX follows the main verb; postpositions predominate; verb derivational extensions are mostly preverbal except the causative, which is suffixal; the number of verbal extensions is relatively small (Orma has three: causative,

'passive', middle voice); there is a grammatical distinction masculine vs. feminine; and possessive, adjective, demonstrative, and verb agree with the noun in gender. As most Kenyan Cushitic languages, Orma has minimal gender or number marking on the noun. In the verb, Cushitic languages, including Orma, show past vs. non-past, which Heine, in later work, interprets as perfect vs. imperfect. Oromo's consonant inventory, as that of some other Cushitic languages, is large: depending on dialect, it contains at least twenty-five members, including an ejective series. As in many surrounding languages, there is an (inherited) contrast of implosive (partial series) vs. explosive stops. Consonants may geminate, and the status and behaviour of [h] and the glottal stop are distinctive. Oromo has five vowels with phonemic length. Vowels may devoice word-finally. Long vowels may shorten under certain conditions and suffixal -n(i) may delete under certain conditions. As all Kenyan Cushitic languages, Orma is tonal.

2.2 Orma as a Linguistic Community

During the period under investigation, the Orma were demographically superior, a fact obscured by today's situation. The most recent (1991) estimate puts the Ilwana (Rossbach, pc) at 15,000: assuming an average of four children per family, then adults would not exceed 5,000. Estimates of the size today of the four Kenyan Oromo dialect communities combined are 70,000 plus, somewhat larger than that of the Ilwana. Although these southern Oromo are part of a huge language chain going up into Ethiopia, that is hardly relevant since the northern relatives were never in the south. But the Oromo have undergone an almost catastrophic decline since the last quarter of the nineteenth century, as a result of military defeat by Somali, epidemics, rinderpest, and other factors. On the other hand, it is likely that the number of Ilwana has risen quite substantially over the same period. Thus the discrepancy in population size is likely to have been much larger a century or more ago and the Orma could well have outnumbered the Ilwana ten to one.

Orma were also militarily and economically superior to Ilwana, and doubtless felt themselves socially superior. They were cattle herders, with a similar set of assumptions about ownership of cattle and about neighboring peoples as the Maasai. They were effortlessly mobile, while the Ilwana, although partly dependent on hunting, were essentially sedentary farmers. Judging by the widespread adoption of kinship terminology by Ilwana from Orma, the Orma imposed themselves socially on the Ilwana.

2.3 Ilwana as a Language

As a Kenyan language, Ilwana shares the characteristics described in the first sentence of the second paragraph of 2.1 above. As a Kenyan Bantu language, Ilwana is SVO; AUX precedes the verb; prepositions predominate; and all verbal derivational extensions are suffixal. The number of extensions today appears to be five (causative, applicative, passive, stative, reciprocal), with the reversive at least having been widely replaced by individual lexical items, and there being no evidence for other extensions present in Ilwana's Bantu relatives. There is also one extension of Orma origin but it occurs only on verbs of Orma origin. Heine and Vossen point out that Cushitic languages have typically five or fewer extensions, whereas Bantu languages have five to eight. Bantu languages are characterized by noun class/gender systems, with upwards of a dozen members, marked by prefixes and arranged in typical singular/plural pairs. All the constituents of the NP, and also

the subject of the verb, agree with the head noun. Bantu languages tend to have past vs. present vs. future, with past and future often further subdivided. Tense is most often indicated at pre-stem position, and aspect, when present, at post-stem position. All Ilwana's relatives have a reduced set of tense distinctions, and both tense and aspect tend to cluster in pre-stem position.

Ilwana has inherited seven vowels and distinctive length. Word-final vowels may devoice, and long vowels may shorten optionally under certain conditions. Ilwana is a tone language but details are limited. Kenya's Bantu languages typically have between '15 and 24' consonant phonemes (Heine and Vossen: 14). Ilwana's relatives are at the high end of this scale because they have absorbed many consonants from outside. Ilwana itself has a huge consonant inventory, with at least thirty-two consonants, including (non-inherited) ejective and implosive series. Consonants occur geminated. [h] and the glottal stop are not inherited and their behaviour is distinctive and similar to that in Orma. There are three different -(V)ni suffixes, and all may delete, as in Orma.

Ilwana is a Sabaki Bantu language, so its nearest relatives include Swahili and Pokomo (Nurse and Hinnebusch 1993).

2.4   Ilwana as a Linguistic Community

As a group, the Ilwana have been very vulnerable. This is clearly true for the recent and remembered past, and also in the farther past, if the linguistic evidence is to be believed. They were sedentary farmers, who also hunted widely and exploited the Tana and its products. They lived in a thin line of hamlets strung out along both sides of the river. The average size of a hamlet today is eighty, including children.

Prior to AD 1600 their history is not clear. For much of the period from AD 1600 to 1875, they were massively influenced by the Orma. The Orma must have been an overwhelming presence, in bush, field, on paths, in villages and markets. Judging by the lexical evidence, and by the few historical accounts, the Ilwana restructured their society towards that of the Orma: terminology for kinship, economy, political institutions, even personal names, point to huge Orma influence. The Orma took Ilwana wives, not vice versa. For at least the latter part of the period, most if not all Ilwana were bilingual in Orma - the bilingualism did not work the other way.

From AD 1875 on, the Ilwana must have managed to achieve a precarious balance between Orma on the west bank, and Somali on the east bank, of the river. In the early 1990s the political events in Somalia overturned the balance again.

2.5   Ilwana Influenced by Orma

2.5.1   Lexicon

Currently, some 2200 verb stems and nonverbal lexemes are available. Of these only some 22.5% are inherited. Of the remaining 77.5%, some 20% come from Swahili, slightly over 15% from Orma, and some 43% are of unclear or unidentified origin. Swahili lexical sources are more complete and available than those for Orma. Thus the Swahili figure is likely more or less accurate while the Orma figure is probably an underestimate. When the source of all items is identified the Orma figure will probably rise, perhaps equaling the percentage of inherited

items in Ilwana. Orma vocabulary can be shown to cluster in certain areas, some pointed out above.

2.5.2 Sound System

While much in the Ilwana sound system is inherited, much is also innovated. Somewhat over half of the consonant inventory is inherited, as are the seven vowels and two lengths. On the other hand, nearly half the current consonant inventory has no inherited source, and many phonological features are of outside, mostly Orma origin: all four ejectives, geminates (or gemination ?), the presence and behaviour of [h] and [ʔ], the deletion of three suffixes of the shape -(*V*)*ni*, vowel shortening, final vowel devoicing, etc.

2.5.3 Nominal System

Some 1400 nouns were collected. Singular and plural forms are available for many as is much detail on NP constituents in general, concord, locatives, relatives, and pronouns.

Historically, the most curious feature of all this is the over 100 nouns, mostly in Classes 9-10, which can form plurals by what are apparently Cushitic suffixes (e.g. *si:ba* 'one lion' but *si:be:na* 'lions'). This feature does not appear to be Orma in origin.

In general, the noun class system appears to be little affected by contact with other languages, including Orma. The only ways in which Orma may have influenced Ilwana here is in a couple of new singular-plural groupings (1-10 and 9-2) and in some new possessive pronouns. An example of 1-10 is *mubokomo* 'Pokomo', pl. *bokomo* ; of 9-2- is *da:di* 'elder sister', pl. *wa-da:di* , and representative of the 'new' pronouns is -*emi* 'my, mine' = 'of me', from *a* 'connective' plus *imi* 'I, me'. While this may be 'natural' and common enough world wide, it is definitely not common in recent Eastern Bantu but is a prevalent pattern in Eastern Cushitic.

2.5.4 Verb

The verb has been significantly affected by contact with Orma. Most obviously, Ilwana's tense-aspect system has been reduced and restructured in the direction of Orma. Ilwana has fused past and perfect into one, based on the older perfect suffix: loss of a category, and use of inherited morphology to express a transferred category. There is a new future whose structure is anomalous but identical to that of Orma: inherited morphemes reused on a transferred pattern to express a transferred category. There is a progressive whose structure is also anomalous but exactly that of Orma: inherited morphemes restructured on a transferred pattern to express a category which may be transferred. Finally in two-word verbs, the Orma auxiliary 'be' has replaced the Bantu verb as the first word.

At post-stem position, the number of extensions is apparently reduced: Cushitic verbs are characterized by a smaller set of extensions than found in Bantu verbs. This needs more confirmation. An Orma extension (*at* 'middle voice') is used with many verbs, but not with native Ilwana words. Finally the Ilwana passive has extended its functions in the direction of Orma. Thus for example, the Ilwana verb -*nyota-wa* 'be thirsty', denominalized from *nyota* 'thirst', corresponds exactly

to an Orma example, as do several others. In Ilwana's relatives, such functions are not found.

2.5.6   Syntax

Heine and Vossen look at well-known order features in Kenya languages: constituents of S, NP and VP, etc. Although my investigation was not specifically aimed at such features, I had fairly abundant syntactic data and it is my impression that little change has occurred in this area.

3.   Daiso and its Neighbors

3.1   Historical Background

The background of the Daiso is quite complex. Once they formed part of the eastern section of the Kamba; even today that area is referred to as U-thaisu by western Kamba. When they left is not clear but a point in the early or mid sixteenth century is likely. When the community migrated it had to split in two before a raging river. One section crossed, went down to the coast, and turning south arrived at Malindi in AD 1571 and at Mombasa twenty years later. Renowned fighters, they were recruited by a Digo ruler to help him in a local war. When victorious, he pleaded poverty and gave them Digo and Chifunzi wives as reward. The ensuing children were brought up by the mothers to speak their language. Hence these Daiso lost their language and their offspring spoke Segeju, a variant of Digo.

The other section crossed the river later and, following an undetermined route, finished up some fifty-sixty kilometers northwest of Tanga, in the eastern part of the (Bantu) Shambala-speaking area, between Bwiti and Maramba. They and their offspring today continue to speak Daiso.

The 'Segeju' seen at Malindi were herding cattle: the Daiso still have herds today. Judging by the farming base of their Kamba brothers and by what they themselves do today, they were also subsistence peasants. Onto this inherited strand was grafted a coastal connection. Despite the language rift that had developed, the Daiso and the Segeju, now Digo, communities remained close. Not only was there - and still is - extensive intermarriage with Segeju/Digo, but a trade connection emerged that had the Daiso trading to the coast and to Kilimanjaro on their own account, and also acting as middlemen for Swahili caravans starting on the coast. In order to communicate with their Digo brothers, and to participate in the trade networks, they had to acquire Swahili (and Islam). Finally, they have been living in their present location for some four centuries, in daily contact with their Shambala (and maybe Bondei) farming neighbors.

3.2   Daiso as a Language Community

The previous apparent discursion was needed as a basis for showing how Daiso got to be as it is. The Daiso, as the Ilwana, community has been linguistically vulnerable, but for different reasons. As the Ilwana, they are relatively few in number: 8,000-10,000 was the estimate offered by several people (no official figure available) and may have been significantly fewer at the start of the twentieth century. The other reason has been their forced historical openness to outside influences. They have had to be bi- or multilingual for much or all of the last four

centuries. As a community they started off by speaking a form of Kamba. Then some (many? all?) acquired Digo to talk to their coastal brethren. Then they acquired Swahili in order to carry out their trading and all Daiso claim to be bilingual in Swahili today. Because of the coastal connection they became Muslims and Udaisu is today a Muslim island in a Christian sea. Because they were Muslim, some members of the community acquired Arabic. And because they lived in the Shambala area, some (many? all?) acquired some knowledge of Shambala, the dominant local language. I also found some evidence of knowledge of other languages (e.g. Zigula). There was no suggestion that Digo/Segeju, Swahili, or Shambala can speak Daiso, so, as for the Ilwana, bilingualism only works one way.

I had the impression that relations between Daiso and Digo/Swahili were cordial. On the other hand, this was less true for Daiso-Shambala relations, perhaps because in the past the Daiso were closely connected with coastal trade, which had had a proven slave component until the nineteenth century.

3.3 Daiso as a Language

As a Bantu language, Daiso would have had and has the following characteristics: SVO; all NP constituents follow the head noun; the governing NP precedes the governed/genitive NP; prepositions; the usual noun class system, with over a dozen classes, familiar singular-plural pairings and agreement with the noun across the NP and into the verb; Aux precedes verb; suffixal extensions.

As a member of the Central Kenya Bantu group, Daiso would have had: aspect marked predominantly in suffixal and tense in pre-stem positions: multiple (three) past and future (two or three) distinctions; seven vowels and distinctive length; distinctive tones; and a relatively small consonant inventory because Bantu spirantization had not occurred. As a member of this group, Daiso would also have had specific morphemes occurring at suffix and pre-stem positions, specific allomorphs, and, of course, certain specific lexical items.

How this was altered will be seen in 3.5 below. It should be noted that, although Daiso appears to have had four centuries to change, today's Daiso is apparently identical to that described by Baumann in 1891 - only three centuries after the community settled at Bwiti. Thus the essential changes - less some Swahili lexis perhaps - had already occurred then. They might, of course, have taken less than three centuries, just as the changes in Ilwana.

3.4 Contact Languages (Shambala, Swahili, Digo)

Digo and Swahili are very similar and closely related languages. Both are similar and related, though slightly less so, to Shambala. While they share many general Bantu features, all three also share many specific features. They share general Bantu features with the Central Kenya languages but differ in specifics.

They have five vowels but no distinctive length; Digo and Shambala (but not Swahili) are tone languages sharing certain details; they have larger consonant inventories than the Central Kenya languages, having undergone spirantization. They share particular verbal morphemes, specific allomorphs, and a certain set of lexical items. They make more use of verbal auxiliaries. They have a reduced set of verbal suffixes compared to Central Kenya languages, make more use of the pre-stem position for aspectual distinctions and have fewer past and future distinctions.

3.5   Daiso Influenced by its Neighbors

3.2, 3.3 and 3.4 are set out in such a way as to make certain outcomes seem likely. First, the language of a small insignificant community, separated from its nearest relatives, forced by circumstances to make bi- or multilingualism a way of life, living for centuries amidst communities speaking different languages, is likely to be reshaped. Second, given that Daiso and the contact languages are genetically related and typologically similar, most adaptations are likely to be minor rather than major.

While the quality of some of the Daiso data could be better, the following changes are clear.

3.5.1.   Sound System

At the level of sound Daiso has: neutralized its inherited length distinction in vowels but kept seven vowels (all the vowels of the class prefixes however appear to be the highest vowels in the system and thus likely taken from or modelled on one of the contact languages); enlarged its consonant inventory so that of twenty-seven consonant phonemes eight occur mainly or exclusively in non-inherited words; phonological processes and morphophonological processes have been frequently adapted so that the outcome is that of Daisu's current neighbors. The single most obvious source of these phonological changes is Shambala.

3.5.2.   Nominal System

Several aspects of the NP were also examined: the class system, class markers and allomorphs (esp. of Classes 5, 7, and 9-10), singular-plural pairings, concord, and NP constituents. Very little significant change was found. New are: a minor pairing (14-10), third person pronouns based on demonstratives, and allomorphy for classes 5 and 9-10. All these appear to come from Shambala. Animate concord has entered from either Swahili or Bondei.

3.5.3.   Lexicon

The over 1050 verb stems and other lexemes in the currently available lexicon were examined. Some 40% are inherited, 60% are not inherited. The latter were compared with corresponding items in Swahili, Shambala, Digo, Pare and other local languages. It should be borne in mind that Swahili and Shambala are better documented than the other languages. The two biggest sources, roughly of equal size, were Swahili and 'indeterminate': by the latter is meant either that the origin of an item could not be determined or that an item's shape and/or meaning was such that it could have come from more than one of the contact languages. Together, Swahili and 'indeterminate' contributed 82% of the non-inherited material (=49% of the total), Shambala 12% (7% of total), and Digo 4%. Many agricultural and domestic terms came from Shambala, and some kinship terms came from Digo, insofar as their origins could be determined. Swahili terms are ubiquitous.

How is this lexical picture to be explained? Initially, in their new location, the Daiso clung to the connection with their coastal brethren, the Segeju. From the beginning the Segeju had shifted to Digo. Intermarriage continued, even to today. So it is not surprising to find Digo kinship terms in Daiso. What then became important about the coastal connection was its link with trade, and the language of

the trade caravans from the coast to the Usambaras, Kilimanjaro, and elsewhere was Swahili. As central figures in this trade, the Daiso acquired Swahili and there was considerable bilingualism, probably from an early point. So the high level of Swahili lexical penetration is probably not a twentieth century phenomenon, but of longer standing. Finally, Shambala lexical penetration is easy to understand, given Daisu's location, but surprisingly light, given the phonological picture.

So far, there is a curious discrepancy. A majority of Daisu's lexicon is not inherited and the largest identifiable contributor is Swahili, as just explained. On the other hand, in terms of phonemic inventory, (morpho)phonological processes, and minor details of N and NP, Daiso has moved considerably towards Shambala. We have to think that three or four centuries of regular daily contact with Shambala produced the latter adaptation, with the underlying and unspoken acknowledgement that it was necessary or unavoidable, probably hardly noticed.

3.5.4. Verb

Finally, several features of the verb were examined in some detail: extensions, tense, aspect, pre- versus post-stem morphology and categories, and auxiliary use. A speaker of Kamba or Kikuyu, looking at Daiso extensions, would see few differences, not surprising, given that the contact languages and the Central Kenya languages are not greatly different in this area. But the same speaker would find Daiso's TA system quite unfamiliar. It is characterized by three main changes: inherited suffixes expressing aspect are reduced from five to two; while past reference is largely maintained, 'present' and future reference is changed and reduced; there is a general reduction in verb-morphological and -categorial complexity.

As in the sound and nominal systems, so here the largest single source of individual non-inherited morphemes is Shambala.

It is much harder to point to the source of the general structural and categorial changes. The end results do not correspond to those in any one contact language. What has emerged is a morphologically and categorially reduced system, looking less like a Central Kenya language, and more like coastal languages in general, but with no single clear source.

4. Ilwana and Daiso: Generalizations

Ilwana and Daiso were small communities surrounded by larger communities and open to their influence. In one case (Daiso) the community was also exposed to the influence of another, not much larger, not even adjacent, community (Swahili), but one that was important for commercial reasons. In one case (Ilwana) the larger community's threat was backed up by military force and in the case of the other, it was not. Both apparently became bi- (or in the case of Daiso possibly multi-)lingual: this polylingualism was unidirectional. In one case the source of outside influence was a single typologically and genetically different language: in the other case the contact languages were of the same family.

In three centuries, or possibly less, the following changes characterized both languages.

Both have undergone massive lexical change. Although the quantity of vocabulary available for both is not great (2200, 1100 items), 60% of this in one case, and over 75% in the other, has been transferred. In one (Ilwana), some 15% of these transfers come from the contact language (likely more if the sources of all

transfers could be identified); in the other case 25% come from a commercial contact language. Judging by what we know of other languages, these percentages would be likely to increase if more vocabulary were available. These heavy transfers occur throughout the vocabulary and both languages also show concentrations in specific semantic areas. One feature of both is that these heavy concentrations of transfers do not show up well on Swadesh-type lists: on the 100-word list Ilwana has only 25% not inherited, Daiso 15%.

In one case (Daiso) the major source language for the new vocabulary is different from that of the non-lexical material.

In both cases phonemic consonant inventories have enlarged, in one (Ilwana) massively so. One interpretation might be that, at least for Ilwana (which has ejectives and an implosive vs. explosive contrast), this involves the assimilation of many marked sounds. A better interpretation would be that both languages have simply moved taxonomically in the direction of the contact languages. In both cases the new consonants are those common to several surrounding languages.

In both cases a number of (morpho)phonological processes have been taken over from the contact language. This simplifies the task of a bilingual.

Tense-aspect systems have undergone considerable change towards the contact languages. In both cases there is restructuring, and reduction of verb-morphological and -categorial complexity. The reduction is only a function of the target languages. Both languages have restructured their categories towards those of the target languages. In some cases inherited morphemes and morphology are used; in other cases transferred morphemes are used; and in other cases inherited morphemes are used in structures calqued on those of the target languages. Where inherited morphemes or morphology are not useful, they are dropped or relegated to minor functions.

In both cases, noun class systems have remained largely intact, with only minor adjustments, mostly phonological. Similarly, in syntax, no major changes were found, although this was not the focus of my investigation.

In neither case was there any evidence of language shift into the minority language nor of pidginization. I therefore conclude that what is described above, and referred to neutrally as 'transferred', in fact results from borrowing.

References

Baumann, O. 1891. Proben der Kissegeju- und Kiparesprache. Usambara und seine Nachbargebiete, O. Baumann, ed., 340-346. Berlin.

Heine, Bernd, and Rainer Vossen. 1980. Language structures in Kenya. The Non-Bantu languages of Kenya, Bernd Heine, ed. 13-27. Berlin: Dietrich Reimer.

Nurse, D. 1982. Segeju and Daisu: a case study of evidence from oral tradition and comparative linguistics. History in Africa 9:175-208.

Nurse, D. n.d. Inheritance, contact, and change in Daiso. St. John's. Ms.

Nurse, D. n.d. Inheritance, contact, and change in Ilwana. St. John's. Ms.

Nurse, D, and Thomas. J. Hinnebusch. 1993. Swahili and Sabaki, a linguistic history. Berkeley: University of California Press.

Rossbach, Hans-Joseph. 1991. 1990-1991 Annual report on Bible translation and literacy in the Malakote language. Nairobi: Bible Translation and Literacy.

Rossbach, Hans-Joseph. 1992. 1992 Annual report on Bible translation and literacy in the Malakote language. Nairobi: Bible Translation and Literacy.

Rossbach, Hans-Joseph. 1993. Preliminary working dictionary Malakote-English-German-Swahili. Nairobi. Ms.

Sasse, Hans-Jürgen. 1979. Entlehnung morphophonemischer Regeln im Boni. Sprache und Geschichte in Afrika 1:93-108.

Stroomer, Hendrikus, J. 1987. A comparative study of three southern Oromo dialects in Kenya. Leiden University dissertation..

Thomason, Sarah G. and Terrence Kaufman. 1987. Language contact, creolization, and genetic linguistics. Berkeley: University of California Press.

The Origin of Leftward Tone Shift in Masasi Chiyao[1]

David Odden
OSU

1. Introduction

The purpose of this paper is to trace the historical development of a leftward H tone shifting process which affects the dialects of the Bantu language Chiyao spoken in Masasi District in Tanzania. It is argued that tone shift arises in these dialects due to contact with Makonde, a related Bantu language.

The organization of the paper is as follows. Section 2 outlines the basic principles of tone assignment common to all dialects of Chiyao, using data from the tonally conservative Tunduru dialect. Section 3 demonstrates the first stage of the historical change with data from the Mangaka dialect: tone shift in this dialect is restricted to leftward spreading of a final H. Section 4 considers tone shift in the Ndanda dialect, which completely prohibits final H's; section 5 argues that this tone shift is due to contact-induced language change, and originated with the prohibition against final H's in neighboring Makonde.

2.   Tonal Basics

This section gives an overview of tone in Chiyao, using data from the Tunduru dialect of Tanzania. More detailed information on tone can be found in Hyman and Ngunga 1994 for Mozambican Chiyao, Odden 1994 for Tanzanian Chiyao, and Mtenje 1993 for Malawian Chiyao. As is typical of languages in Guthrie's Zone P, there are no lexical tone classes in verbs. In verbs, surface H tones are assigned on the basis of morphological properties. Thus in the infinitive, H is assigned to the first root mora; in the recent perfective H is assigned to the second stem mora; in the remote perfective H goes on the final vowel. Interacting with these initial tone mapping rules are various general tone rules.

2.1.   Doubling

One of these rules is Doubling, which spreads any H from its underlying mora to the mora immediately following. This can be seen in (1) with examples of the future tense. In that tense, H is assigned to the first root mora.[2]

(1)   chíná[chi-téléche   'I will cook it'
       chítwá[a-pílíkane   'we will listen to them'
       chítú[chi-súúmisye   'we will sell it'

Doubling does not spread H to a prepausal syllable; however spreading is allowed to a word-final syllable in phrase-medial position.

(2) cháá[ji-líme   'he will cultivate them'
     chíná[si-táve   'I will build them'
     cháá[ji-límé miguunda   'he will cultivate the fields'
     chíná[si-távé pe   'I will merely build them'

Doubling does spread final H from a word-final syllable to a following word, as shown in (3).

(3) chitúúmbili chigwiíle   'the monkey fell'
     chivigá chígwiíle   'the pot fell'
     chíváaváláánjile mandaanda   'he will count eggs for them'
     cháambé mándaanda   'he will give me eggs'

A formulation of Doubling is given below.

(4) **Doubling**   H
                        ⌈ ˋ ˎ
                        V   V

The prohibition against spreading to the prepausal syllable is due to a postlexical constraint against prepausal H's, possibly involving extraprosodicity.

In the negative past tense, H is assigned to the second stem mora. If the stem has the shape CVCV, that H will appear on the final syllable, and in the citation form this results in a prepausal H. In the Tunduru dialect of Chiyao, this situation is tolerated. Morphosyntactically conditioned assignment of H in verbs is lexical rather than phrasal, so assignment of H within the word proceeds without consideration for the phrasal consequence of creating a prepausal H. The data in (5) illustrate application of Doubling within the word in this tense, along with the suspension of Doubling in CVCVCV stems. These data further illustrate application of Doubling between words.

(5) nganiim[bilíkána   'I didn't hear'
     nganiin[deléka   'I didn't cook'
     nganitu[limá   'we didn't cultivate'
     nganitu[limá míguunda   'we didn't cultivate fields'

The data in (6) involve the remote perfective where H is assigned to the stem final vowel. This H spreads to the following word.

(6)   naa[sósilé   'I dug'
       naa[sósilé lísiimbo   'I dug a hole'

2.2.   Retraction, Absorption

Data from the remote perfective also shows that if a final H tone is preceded by a long syllable, that H shifts to the preceding syllable, resulting in a rising tone.

(7)   naá[liíle   'I ate'
       naa[téleéche   'I cooked'
       naa[súúmiísye   'I sold'

Phrase-medially the H only appears on the final syllable, and does not spread to the following word.

(8)   naa[téleeché mandaanda   'I cooked eggs'
       naa[súúmiisyé mandaanda   'I sold eggs'

These alternations are accounted for as follows. The rule Retraction shifts final H to the preceding long syllable.

(9)   **Retraction**
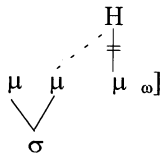
Following this, Doubling applies in phrase-medial position. This would create a rising tone followed by a H tone, which is not a possible sequence in Chiyao. Thus the following Absorption rule applies.

(10)   **Absorption**
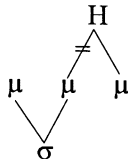
The derivation in (11) illustrates the interaction between Retraction, Doubling and Absorption in prepausal and phrase-medial contexts.

(11)

	H H	H H	**Underlying**
	naateleeche	naateleeche mandaanda	

H H   H   H   **Underlying**

naateleeche   naateleeche mandaanda

H H   H   H   **Retraction**

naateleeche   naateleeche mandaanda

H   H   H   H   **Doubling**

naateleeche   naateleeche mandaanda

H H   H   H   **Absorption**

naateleeche   naateleeche mandaanda

2.3. Lexical Tone in Nouns

Nouns, in contrast to verbs, have lexical tone. The canonical noun stem of the form CVCV may have one of four tone patterns, corresponding to the free specification of any mora for tone.

(12)
njété	'salt'	njété jígwiíle	'the salt fell'
chivigá	'pot'	chivigá chígwiíle	'the pot fell'
chisúvi	'leopard'	chisúvi chígwiile	'the leopard fell'
chijuni	'bird'	chijuni chígwiíle[3]	'the bird fell'

This gives the basic paradigm of tonal behaviour for nouns in the relevant tone classes: HH and LH nouns spread their final H to the following word, and HL nouns spread their H within the word to the word final syllable.

3. Leftward Tone Spreading in Mangaka Chiyao

Whereas the Tunduru dialect maintains a four-way tone contrast in bimoraic stems, the Chiyao dialects of Mangaka and Ndanda in Masasi District have reduced this to a three- or two-way contrast. In the Mangaka dialect, HH and LH neutralise to HH in the citation form, and in the Ndanda dialect HH, LH and HL neutralize to HL.

(13)

	Tunduru	Mangaka	Ndanda	
HH	njété	njété	njéte	'salt'
	mipíní	mipíní	mipíni	'handles'
	malóvé	malóvé	malóve	'words'
LH	chivigá	chivígá	chivíga	'pot'
	liijaní	liijání	liijáni	'baboon'
	manyasí	manyasí	manyási	'grass'
HL	chisúvi	chisúvi	chisúvi	'leopards'
	ngúku	ngúku	ngúku	'chicken'
	machíle	machíle	machíle	'root'
LL	ngoji	ngoji	ngoji	'rope'
	majela	majela	majela	'hoes'
	chijuni	chijuni	chijuni	'bird'

It is also appropriate to note that final HH in the Mangaka dialect is not phoneti-
cally comparable to HH in the Tunduru dialect: the pitch of final HH in Mangaka
is noticeably downstepped.

In the Mangaka dialect, nouns with the historical patterns HH and LH
both act like original HH nouns — phrase medially the final H spreads to the
following word. Nouns with the tone pattern HL act just as they do in the Tun-
duru dialect: the H spreads to the final vowel in phrase medial position.

(14)

Tunduru pattern	Mangaka citation form	Mangaka phrase medial	
HH	njété	njété jáangu	'my salt'
	lusúló	lusúló lwáangu	'my river'
	liijóká	liijóká líwiíle	'the snake died'
LH	manyásí	manyásí gáangu	'my grass'
	chivígá	chivígá cháangu	'my pot'
	mapúkú	mapúkú gáangu	'my rats'
HL	ngúku	ngúkú jaangu	'my chicken'
	mbéju	mbéjú jaangu	'my seed'
	ukána	ukáná ulíingwá	'how much beer'

The neutralization of the HH and LH patterns is due to (15).

(15) **Leftward Backspread**

$$\begin{array}{ccc} & & H \\ & \nearrow & | \\ \mu & \mu & \omega] \end{array}$$

Since tone is not assigned by productive rules in nouns as it is in verbs, it is impossible to motivate the backspreading rule with CVCV noun stems, since there are no alternations. In such nouns it is just as simple to assume that nouns which formerly had the tone pattern LH now have the tone pattern HH. However Leftward Backspreading can be motivated in monosyllabic noun stems with H; in such nouns, the noun class prefix which would otherwise be toneless has a H tone. Note that this H appears both prepausally and utterance-medially.

(16) ḿ-bwá ḿ-bwá jwáangu   'my dog'
      mí-twé mí-twé jéetu   'our heads'

Thus the backspreading rule of the Mangaka dialect applies at the word level, not the phrase level — compare this to the more common variety of backspreading found in Kikerewe which spreads H to the left only if it is prepausal, thus giving alternations such as *ndalá* → *ndálá* 'leopard' ~ *ndalá yáanga* 'my leopard'.

In the realm of verbs, Leftward Backspread is easy to motivate given the paradigmatic nature of tone in verbs. Such data also shows that there is a restriction on (15), namely that it is optional if the antepenultimate syllable has a H tone — in which case applying the rule would create adjacent H tones in violation of the OCP. In that case, the two H's are separated by a downstep: this is due to the fact that all final H's in this dialect have a lower pitch.

(17) aka[téléche   'go cook'   aka[gáláusye 'go change'
      aka[líme   'go cultivate'   aká[lyé   'go eat'
      mna[téléche   'don't cook'   mna[líme   'don't cultivate'
      mná[lyé   'don't eat'
      cháá[tu-límíle 'he'll cultivate for us' cháá'[tú-pé 'he will give us'
                                         ~cháátupé

Backspreading applies in the negative future, where H is assigned to the second stem mora. If the stem is bimoraic, H would be assigned to the final vowel, and therefore the H spreads leftward by (15).

(18) ngaam[bilíkánila   'I won't listen'
      ngaan[deléka   'I won't cook'
      ngaan[dímá   ← /ngaandimá/   'I won't cultivate'

In the remote perfective, H is assigned to the final vowel. This H is retracted to a long penult; otherwise the H spreads (optionally, if the penult is preceded by a H tone) to the penult.

(19) náá[ng'weéle   'I drank'   naa[téleéche 'I cooked'
      naa[líˈmílé   'I cultivated'   ná[á-puutílé 'I beat him'
      ~naalímílé

Phrase-medially, this final H is subject to both Leftward Backspreading and Doubling.

(20) naa[sóóˈsílé lísiimbo   'I dug a hole'
      náá[liilé mandaanda   'I ate eggs'

    The account of this alternation is analogous to that for the Tunduru dialect; the added complication found in the Mangaka dialect is that word final H spreads (not shifts) leftward.

(21)   H   H   H H   **Underlying**
          | | | |
      naasoosile lisiimbo   naaliile mandaanda

          H   H   H H   **Retraction**
          | | | |
      naasoosile lisiimbo   naaliile mandaanda

          H   H   H H   **Doubling**[4]
          |\ |\ | |\
      naasoosile lisiimbo   naaliile mandaanda

          H   H   H   H   **Absorption**
          | \ | |
      naasoosile lisiimbo   naaliile mandaanda

          H   H   H H   **Leftward Backspread**
          | ╱\ | | |
      naasoosile lisiimbo   naaliile mandaanda

In summary, owing to paradigmatic alternations in verbs, the tonal system of the Mangaka dialect can be described as being that of the non-Masasi dialects of Chiyao, with the addition of a rule spreading word final H to the left. The system of contrasts in nouns, on the other hand, has undergone restructuring so that nouns in the original LH pattern have been reanalyzed as having the HH pattern.

4.   Tone Shift in Ndanda Chiyao

The Ndanda dialect of Chiyao has taken a further step in neutralizing tone contrasts: HH, HL and LH are all represented in this dialect as HL in the citation form, alternating with HH phrase-medially, due to Doubling.

(22)

Tunduru pattern	Ndanda citation form	Ndanda phrase-medial	
HH	njéte	njété jaangu	'my salt'
	lusúlo	lusúló jaangu	'my river'
	lilóve	lilóvé lyaangu	'my word'
LH	lijáni	lijání lyaangu	'my baboon'
	usílwa	usílwá waangu	'my callous'
	chivíga	chivígá chaangu	'my pot'
HL	litúnu	litúnú ali	'that hyena'
	njípi	njípí jinnúmíle	'a louse bit me'
	chisúvi	chisúví chaangu	'my leopard'

This neutralization of patterns can be understood as the combined effect of Leftward Backspreading as found in the Mangaka dialect (which neutralizes LH and HH to HH) plus prepausal delinking (which neutralizes HH and HL to HL): indeed, prepausal H tones are impossible in the Ndanda dialect.

(23)   **Final Delinking**

$$
\begin{array}{c}
\mathrm{H} \\
\bcancel{|} \\
\mu \quad]]
\end{array}
$$

As can be seen in (22), irrespective of tone class all nouns which have a H now behave like original HL nouns.

Monomoraic stems with H present a different picture: they have the citation form HL, and phrase-medially HH with spreading of the final H to the next word. Phrase medial examples show that Final Delinking only affects phrase-final H's, but Leftward Backspread applies to a word final H even when phrase-medial.

(24)

/n-twé/	ń-twe	ńtwé wáangu	'my head'
/lii-ú/	líi-u	lííú lyáangu	'my ash'
/m-bwá/	ḿ-bwa	ḿbwá jwáangu	'my dog'

In other words, the pattern LH found with disyllabic noun stems is restructured as HL, but the same pattern with monosyllabic stems is not since the pattern is not limited to a single morpheme, precluding lexical reanalysis.

Verbs present robust evidence in the form of paradigmatic alternations for this analysis of leftward shifting of final H. Consider infinitives: in the infinitive, H is assigned to the initial root mora, but if the root is monosyllabic, the H surfaces on the infinitive prefix.

(25)   kú[wa   'to die'   kú[ng'wa   'to drink'

Similar evidence can be marshalled from other tenses.

(26)   tuka[téléche   'let's go cook'   tuka[súúme   'let's go buy'
        tuká[lye   'let's go eat'   nká[ng'we   'go drink'
        nká[ng'wé máveele 'go drink milk'

In the remote perfective, H is assigned to the final vowel. In most dialects H remains if the preceding vowel is short (it will spread to the preceding vowel in the Mangaka dialect by Leftward Backspreading). But in the Ndanda dialect, all final H's shift to the penult, so the question arises whether there is evidence that this dialect retains the original pattern assigning H to the final. It is conceivable that this verbal tone pattern has been reanalyzed as one assigning H to the penult, and indeed such a reanalysis is found in nouns which originally had the tone pattern LH. Nevertheless, there is evidence that even in this dialect H is assigned to the final syllable, and shifts to the left either by Retraction or by the combination of Leftward Backspreading and Final Delinking.

If the penult is short, a H appears on the penult in the citation form, and on the penult, final, and following word phrase medially.

(27)   nááválaasíle   'I counted'
        náásúú'míle   'I bought'
        náásúú'mílé lígóombo 'I bought a banana'

This follows from assigning H to the ultima, applying (15) in all contexts, and either delinking the final H in prepausal position, or else applying Doubling phrase-medially.

When the penult is long, that syllable bears the H in the citation form; phrase medially, the final syllable has the H and that H does not spread to the following word.

| (28) | náátéleéche | 'I cooked' |
| | náátéleeché mbataáta | 'they cooked a potato' |

This pattern is exactly like that found in the other dialects, reflecting the interaction of Retraction, Doubling and Absorption. Thus paradigmatic alternations support the retention of the original final-H pattern in verbs. But in nouns, all historical final H's are reanalysed as being lexically on the penult (except in the case of monosyllabic nouns, where such reanalysis is impossible).

5.   The Source of Leftward Tone Shifting.

It is apparent that the Masasi dialects of Chiyao are undergoing a historical change which eliminates final H tones. The initial seeds of this change are found in the Mangaka dialect, with its Leftward Backspreading rule and phonetic lowering of final H. The most fully developed version of this change appears in the Ndanda dialect which systematically moves H off of the final vowel.

The question is why this change has affected these dialects. Certain sociolinguistic facts of the area help to explain the origin of this tone shift. Tunduru is located in the most northwestern region of Chiyao speaking territory; Mangaka is located to the east near Masasi town, and Ndanda is spoken at the extreme east edge of Chiyao speaking territory, right at the base of the Makonde Plateau and Newala District. Thus Ndanda is in most intimate contact with speakers of Makonde, and Tunduru is quite insulated from Makonde speakers, with Mangaka being situated between the two but still in close proximity to Makonde speakers. In this area, bilingualism in Yao and Makonde is common.[5]

A fundamental fact about tone in all Makonde dialects (see Odden 1990a,b for discussion of Makonde tone) is that H can never be assigned to the final syllable of a word. Consider the alternations in (33).

(29)	nindaloóla	'I will see'	nindaalya	'I will eat'
	nindataleéka	'I will cook'	nindaloodya	'I will show'
	nindakalamoóla	'I will cough'	nindavingiliidya	'I will chase'
	kungúluúma	'to bite me'	kungúloodya	'to show me'
	kungútoteéla	'to sew for me'	kungútoteeya	'to make me sew'

H is assigned to the penultimate mora in the future tense. Some verb stems end in a vowel, cf. the examples in the righthand column. In such forms, H would be assigned to a prevocalic vowel, thus /nindaloodía/. A regular rule of glide formation desyllabifies that vowel, and therefore it can no longer bear tone. One would expect H to transfer to the final syllable, but that would result in a word final H which is prohibited. Therefore, the final H tone is deleted.

As shown in Odden (1990a,b) these and a number of related facts can be explained by positing that the final syllable in Makonde is extraprosodic, and thus not a fit tone-bearer. The similarity between Makonde and the western dialects of Chiyao lies in the abstract principle that final H's are disallowed. The languages differ in how this principle is enforced, in that Chiyao allows final H's at the lexical level and takes steps to remove them postlexically, whereas Makonde systematically eschews final H's even at the lexical level.

Notes

1. Data for this paper was gathered at the University of Dar es Salaam in 1989, with the support of a Fulbright Research grant. I would like to thank the University of Dar es Salaam and the Tanzanian Commission for Science and Technology for assistance in conducting that research; I would also like to thank my Yao language consultants, Rashid Akwilombe, Emmanuel Hamisi, Gabriel Njunju, Mohammed Zuberi. I have benefited from discussion of Yao with Larry Hyman, Chuck Kisseberth, Al Mtenje and Armindo Ngunga, none of whom can be held responsible for errors in this paper.
2. The verb stem, whose left edge is marked with '[', contains an optional object prefix, which is separated from the root by a hyphen. Thus the stem in the first example is *chiteleche* and the root is *telech*.
3. H is assigned to the verb *chigwiile* by a phrasal rule inserting H after a toneless word.
4. Doubling does not spread H from the prefix *-a-* to the stem *-liile* due to a constraint against spreading to a H-toned syllable.
5. Indeed, my Mangaka Chiyao consultant is bilingual in Makonde and Chiyao, and controls the tone system of the Chimahuta dialect of Makonde.

References

Hyman, L. and A. Ngunga. 1994. On the non-universality of tonal association 'conventions': Evidence from Ciyao. to appear in *Phonology*.

Mtenje, A. 1993. Verb structure and tone in Chiyao. *Topics in African linguistics*, S. Mufwene and L. Moshi (eds.), 179-190. Philadelphia: Benjamins.

Odden, D. 1990a. Tone in the Makonde dialects: Chimaraba. *Studies in African linguistics* 21: 61-105.

Odden, D. 1990b. Tone in the Makonde dialects: Chimahuta. *Studies in African linguistics* 21: 149-187.

Odden, D. 1994. Principles of tone in Tanzanian Yao. Ms, OSU.

The Historical Development of Secondary Articulation in Gurage*

Sharon Rose
McGill University

1. Introduction

Gurage (South Ethio-Semitic) languages are characterized by widespread palatalization and labialization processes, often with no obvious trigger. These secondary articulations often mark morphological classes either alone or in combination with other affixes. In addition, only a subset of palatalizable or labializable consonants are affected in different contexts and across the languages. Building on proposals by Hetzron (1971, 1975, 1977) and Leslau (1967), I address how secondary articulation arose in Gurage, and propose three separate stages corresponding to the types of consonantal targets. I also examine the interaction between palatalization and labialization in certain verb forms such as the impersonal, and conclude that there is no interaction. In section 2 I introduce the Gurage languages and discuss the process of palatalization and labialization. In section 3 I discuss previous analyses of the impersonal form of the verb and in section 4 I propose a modified analysis, which represents the first stage of palatalization. In sections 5 and 6 I discuss the second and third stages of palatalization and in section 7 I show synchronic phonological representations.

2. Classification of Gurage and Some Preliminary Data

I will follow Hetzron's (1977) classification of the Gurage languages (shown in 1).

(1)   PROTO-ETHIOPIC

                    South-Ethiopic   North-Ethiopic

                                                        Tigre
                                            Ge'ez Tigrinya

    Outer South-Ethiopic   Transversal South-Ethiopic

Gafat   Central   Eastern
  .(Northern Gurage)
                                      Amharic   Argobba   Harari
Soddo   Goggot Muher
                        Western Gurage   East Gurage

                    Masqan   Selti (etc.) Zway

                    Central   Peripheral
                Western Gurage   Western Gurage

        Eža   CHAHA Gumer Gura   Gyeto INOR Endegeñ Ener
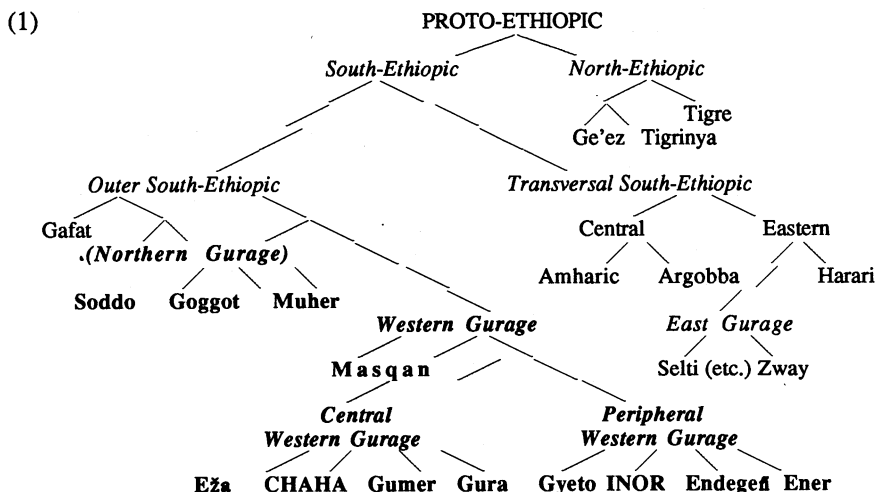
He argues that East Gurage (Selti, Zway, Wollane, etc.) are more closely related to Amharic and Harari than the Western and Northern Gurage languages, and as a result the term "Gurage" more aptly characterizes a geographical region rather than a linguistic family. Due to this split, and the fact that information on East Gurage is

somewhat scant, I will restrict my focus to the Northern Gurage and Western Gurage languages.

Chaha will be the representative language of the Central Western Gurage (CWG) group, and data on this language holds for the other dialects, and also for Gyeto, which patterns with the CWG dialects as far as palatalization and labialization are concerned. Inor (often referred to by its Amharic name Ennemor) will represent Peripheral Western Gurage (PWG), although differences with Endegeñ will be pointed out where applicable. Soddo, Goggot, Muher and Masqan all have crucial differences as regards secondary articulation, so I will treat them separately. The main focus of the discussion will be on verbal contexts, as the patterns are more readily identifiable, although see Rose (1992b) for some discussion of noun forms in Chaha.

2.1. Palatalization and Labialization

Palatalization turns alveolar obstruents into palato-alveolar in all Gurage languages. In Western Gurage and Muher, velars and coronal sonorants may also be palatalized (T=alveolar ejective, C=palato-alveolar affricate ejective, q=velar ejective):

(2)   __All Gurage__   __Western Gurage and Muher__

$t \rightarrow \check{c}$ $k \rightarrow k^y$

$d \rightarrow \check{j}$ $x \rightarrow x^y$

$T \rightarrow C$ $q \rightarrow q^y$

$s \rightarrow \check{s}$ $g \rightarrow g^y$

$z \rightarrow \check{z}$ $r \rightarrow y$ (CWG & PWG)

     $n \rightarrow \check{n}$ (some dialects of PWG)

Palatalization of coronals when triggered by a suffix is restricted to the final consonant of the stem, but palatalization of velars when triggered by a suffix affects the rightmost velar, and occurs in fewer contexts.

Labialization targets labials and velars in Western and Northern Gurage, and when triggered by a suffix affects the rightmost labializable consonant:

(3) $p \rightarrow p^w$ $k \rightarrow k^w$

     $b \rightarrow b^w$ $x \rightarrow x^w$

     $(\beta \rightarrow w)$ $q \rightarrow q^w$

     $f \rightarrow f^w$ $g \rightarrow g^w$

     $m \rightarrow m^w$ $\text{?} \rightarrow \text{?}^w$ (PWG only)

2.2. 2nd Singular Feminine Subject[1]

Palatalization is the sole indicator of the 2nd singular feminine subject marker in all of Western Gurage and in Muher. The final coronal (including /r/, except in Masqan and Muher (Hetzron 1977)) or velar is palatalized.

(4)   Chaha

	2sg masc	2sg fem	
a.	kɨft	kɨfč	'open!'
b.	nɨkɨs	nɨkɨš	'bite!'
c.	sɨβɨr	sɨβɨy -> sɨβi	'break!'
d.	dɨrg	dɨrgʸ	'hit!'

If a velar consonant occupies the first or second radical position, it may be palatalized, providing all the consonants to its right are labial (labials are never palatalized):

		2sg masc	2sg fem	
(5)	a.	nɨxəβ	nɨxʸəβ	'find!'
	b.	qɨfɨf	qʸɨfɨf	'cut the edges!'

However, if a coronal is in the first or second radical position, it cannot be palatalized, and palatalization is instead realized on the second vocalic position (between C2 and C3) through fronting the vowel: ɨ or no overt vowel becomes i, ə becomes e or a becomes ɛ. Furthermore, palatalization of a velar in stem-initial position is blocked by an alveolar in medial position (6d):

		2sg masc	2sg fem		
(6)	a.	nɨzɨβ	nɨziβ	*nɨžiβ	'be flexible!'
	b.	sɨrəf	sɨref	*sɨyəf	'be scared!'
	c.	Tɨməm	Tɨmem	*Cɨməm	'be contrary!'
	d.	kɨtɨf	kɨtif	*kɨčif *kʸɨtif	'chop meat!'

There are other complications associated with the realization of the 2sf subject marker, but these will not be dealt with here (see Rose 1992b, 1993 for more details).

In Soddo, there is a suffix /-i/ marking a 2sfem. subject, and this triggers palatalization of the immediately preceding coronal obstruent. In addition, the first vowel of the stem is fronted. In Goggot, the /-i/ is not present, but the same pattern occurs (data from Goldenberg 1969 - /-u/ and /-n/ are main verb markers):

(7)   Soddo

	2s masc	2s fem	
a.	tɨsəbr-u	tɨsebrɨ-n	'you break'
b.	tɨləbs-u	tɨlebši-n	'you dress'
c.	tɨsnəbbɨt-u	tɨsnebbɨči-n	'you pass time during the week'

In contrast with Western Gurage, Soddo and Goggot do not have palatalized velars, and in this manner, the forms in (7) are similar to Amharic, whose 2sfem. suffix /-i/ triggers palatalization of the final coronal. However, Soddo and Goggot do not allow palatalization of the sonorants /l/ and /n/ in this form, whereas Amharic does.

2.3. Plural Forms (Peripheral Western Gurage)

Peripheral Western Gurage dialects are unique in having secondary articulation in

their plural verb forms (see Bahru Lilaga 1988, Habte Mariam Marcos 1974 and Hetzron & Habte Mariam Marcos 1966 for discussion of Inor). The 3rd masculine plural (past and non-past) and the 2nd masculine plural (non-past) are marked by labialization of the rightmost velar or labial and palatalization of the final coronal obstruent. The 3rd feminine plural (all forms) and the 2nd feminine plural (non-past) display palatalization of the final coronal obstruent only[2]:

(8) Inor

		3masc.pl.	**3fem.pl.**	
a.	√kfd	kəfᵂəj-u-m	kəfəj-a-m	'they opened'
b.	√nks	nəkᵂəš-u-m	nəkəš-a-m	'they bit'
c.	√drg	dənəgᵂ-u-m	dənəg-a-m	'they hit'
d.	√sbr	səpᵂə-m	səpər-a-m	'they broke'

2.4. Impersonal

The plural forms in PWG are similar to the Impersonal form in Western Gurage in general. The impersonal verb forms (perfective, imperfective and jussive) are used when the subject is unspecified, although they may also be used for stylistic reasons when there is an explicit subject (see Leslau 1967 for more details). In Western Gurage (including Masqan), the impersonal is characterized by palatalization of the final coronal obstruent and labialization of the rightmost velar or labial. The impersonal almost always appears with a heavy object suffix (for the light/heavy distinction see Polotsky 1951). If there is no object, the 3smasc. object marker /-i/ will be used, even for intransitive verbs:

(9) Chaha

		Perfective	
a.	√kft	kəfᵂəč-i-m	'one opened'
b.	√rks	nəkᵂəš-i-m	'one bit'
c.	√drg	dənəgᵂ-i-m	'one hit'
d.	√gtr	gᵂətər-i-m	'one put to bed'

In Northern Gurage, however, different patterns obtain. In Goggot, only the final segment undergoes palatalization or labialization, according to Leslau (1979). But Hetzron (1968, 1971 1977:81) claims that this language only has palatalization.

(10) Goggot (Leslau 1979)

a.	√sykt	šəkkəč-i-m	'one made'
b.	√Tbs	Təbbəš-i-m	'one roasted'
c.	√drg	dərrəgᵂ-i-m	'one hit'
d.	√lygm	leggəmᵂ-i-m	'one mounted a horse'
e.	√qbr	qəbbər-i-m	'one buried/planted'
f.	√kdn	kəddən-i-m	'one thatched a house'

Hetzron (1971) reports that only palatalization of the final coronal obstruent marks the Impersonal in Muher. (Leslau (1981) claims that Muher also has labialization.)

(11)   Muher (Hetzron 1971)

a.   √qbT   qǝbbǝC-i-m   'one missed (it)'
b.   √aks   ɛkkǝš-i-m   'one waited'
c.   √agd   y-agǰ-i-t   'one ties'
d.   √sbr   yɨ-sǝbr-i-t   'one breaks'
e.   √srq   yɨ-sǝrq-i-t   'one steals'

In Soddo, no palatalization or labialization occurs:

(12)   Soddo

a.   √grf   gǝrrǝf-u-t   'one freed'
b.   √sbr   sǝbbǝr-u-t   'one broke'
c.   √lbs   lǝbbǝs-u-t   'one dressed'

3. Previous Analyses of the Impersonal

Various proposals have been advanced to account for the origin of secondary articulation found in the impersonal form of the verb in Gurage. Most researchers agree on one point: labialization is due to the Proto-Ethiopic 3rd person plural marker /*-u:/. The loss of a length contrast between peripheral vowels in Ethio-Semitic rendered it unstable. This view is accepted by Polotsky (1938), Leslau (1967) and Hetzron (1968, 1971, 1977). Nevertheless, the origin of palatalization remains an unresolved issue.

Polotsky (1938) and Leslau (1967) propose that palatalization is due to the 3smasc. object marker /-i/. This marker is found on all impersonal forms which do not take another object marker. It may even appear on intransitive verbs: ex. fenǝq^w-i-m 'one burped' (Chaha). Polotsky and Leslau argue that once palatalization was triggered by the object marker, it was extended throughout the paradigm and became a property of the impersonal even before other object markers. A potential problem with an analysis in which /-i/ is the trigger is that palatalization in Gurage always involves loss of the trigger. It is what Bhat (1978) labels 'absorbing' type palatalization.

Hetzron (1971) takes a different view. He proposes that once labialization was triggered by the /u:/ suffix, this suffix underwent a dissimilation to /i/, which triggered palatalization:

(13)   *kǝ(f)fǝtu: --> kǝ(f)f^{w}ǝtu --> kǝ(f)f^{w}ǝti --> kǝ(f)f^{w}ǝč

Hetzron backs up his claim with other examples of dissimilation of /u/ to /i/, such as the 1s. perfect subject marker /ku/ plus the main verb marker /u/ yielding /ki/ in Soddo, and /kwi/ in Goggot. However, all his examples involve cases of strictly adjacent segments, whereas the impersonal form often involves labialized consonants several segments away from the suffix /-u/. A more serious objection to this analysis is that palatalization is intrinsically dependent on prior labialization, yet other palatalized forms (Inor 3plfem., Muher impersonal) have no labialization. In addition, labialization in Gurage is always absorbing: /Cu/ results in the sequence [C^{w}i].

Finally, Elmedlaoui (1992) compares the impersonal form to the middle

passive CuCiC of Classical Arabic. He suggests that the short /u/ and /i/ of the passive were set afloat and are responsible for the secondary articulation we find in Gurage. While intriguing, this analysis does not coincide with other properties of palatalization and labialization in these languages. These processes are suffixal; they affect the rightmost velar or labial, or the final coronal. In addition, palatalization of coronals may only occur preceding the trigger, as in Type B verbs (see Rose 1992a,b for more details). There are adjectival/nominal forms in Gurage which resemble the passive form, however. In these cases, we do not find the typical suffixal secondary articulation, but rather labialized and palatalized (coronal) segments *within* the stem. The forms in (14) illustrate that those consonants affected are not necessarily the final coronal or the rightmost velar or labial. In (14a,b), palatalization of a coronal occurs in C2, a form ruled out by suffixal palatalization (see forms in (6)). In (14c,d), the initial or medial labial is labialized, despite the presence of velars in final position:

(14) <u>Chaha</u>

a.	√xdr	x^{w}ɨjɨr	<	CuCiC	'clothes'
b.	√fTm	f^{w}ɨCɨm	<	CuCiC	'cruel (person)'
c.	√nfg	nif^{w}ɨg	<	CuCC	'stingy (person)'
d.	√fg	f^{w}ɨgɨg	<	CuCC	'last night of the full moon'

4. Impersonal and Plural Forms

4.1. Impersonal

The analysis I propose to account for the impersonal verb essentially follows that of Leslau (1967) and Polotsky (1938). The impersonal was formed from a 3rd person plural marker /-u/ and the 3smasc. object marker /-i/. However, I will offer an account to solve the problem of the 'absorbtion of palatalization'. Gurage languages in general do not tolerate vowel-vowel sequences. If both vowels are high, a glide is always inserted between them: ex. /abi-u/ --> [abiyu] 'It is Abi'. The sequence /u-i/ requires glide insertion in Gurage: /uyi/. The /-u/ triggered labialization and became the exponent of the impersonal:

(15)   *kəfəd-u-i -> kəfəduyi -> kəf^{w}ədyi

The /y/ palatalized the preceding coronal obstruent: (Note Polotsky (1938) proposes an original /*-yi/, being aware of the absorbtion property.)

(16)   kəf^{w}ədyi --> kəf^{w}əji

It is worth noting that Soddo, the only Northern or Western Gurage language with no palatalization in the impersonal, is also the only one of the group to have /-u/ and not /-i/ as the 3smasc. object marker.

4.2. Plural Forms

Labialization in the plural forms of PWG can also be attributed to the Proto-Ethiopic 3rd plural subject marker /*-u:/. in PWG, where it labialized the rightmost labializable consonant. As compensation, one might propose that the 3rd person

plural object marker was adopted as a subject marker: /-yu/. The /y/ would then be responsible for the palatalization:

(17)   *kəfəd-u: --> kəfədu --> kəfʷəd --> kəfʷəd-yu --> kəfʷəj-u

In the feminine forms, which only have palatalization, the feminine plural object marker /-ya/ would have been adopted by analogy with the masculine form, despite the presence of a subject marker /a/. This marker also triggered palatalization:

(18)   *kəfəd-a: -> kəfəd-a-ya -> kəfəd-ɛa -> kəfəja

While in PWG, the original subject markers /*-u:/ (and /*-a:/) were either reduced to floating features or shortened, in the other Gurage languages, they came to be supported by other vowels and consonants:

(19) | | **PWG** | **CWG** | **Masqan/N. Gurage** | **Gura/Gyeto** |
|---|---|---|---|---|
| 3mp | -u(a) | -əw -> o | -mu / mʷ | -əw -> o |
| 3fp | -a(a) | -əma | -(ə)ma | -əβa |

4.3.   Interaction of Labialization and Palatalization

In this section, I will address the question of whether labialization and palatalization interact in those forms which have both. McCarthy (1983), Lieber (1988) and Elmedlaoui (1992) argue that labialization and palatalization potentially target the same consonant (velars) and since labialization occurs, it must be ordered first, and the features representing palatalization are not realized. The following illustration is taken from McCarthy (1983):

(20)   [+round] [+high, -back]
                                      /
          d   n   g
          | | |
          C v C v C   --> dənəgʷ

However, recall that the Muher impersonal has no labialization, and yet velars are still not palatalized (21c), this despite the fact that velars may be palatalized in other forms in the language, such as the second person singular feminine (21f):

(21)   **Impersonal**   **2s.fem.**
a.   yagj-i   'one ties'   d.   tagij   'she ties'
b.   yɨsəbr-i   'one breaks'   e.   tɨsəbir   'she breaks'
c.   yɨsərq-i   'one steals'   f.   tɨsərqʸ   'she steals'

Thus, the first problem with their proposal is that it is based on the assumption that velars can be palatalized in the impersonal, since they can be in the 2sfem. But, the 2sfem. also allows palatalization of /r/ and vowels, which the impersonal does not.
      Another problem is that labialized consonants routinely succumb to palatalization in other contexts, but not vice versa. Thus the 3smasc. light object marker which consists of labialization (ʷ) and /-n/ in CWG, will have no effect on palatalized consonants (22a-b), but labialized consonants do become palatalized in

the 2sfem forms (22c-d) (palatalization is represented as /- ^{i}/):

(22)   Chaha
a.	/yɨ-gʸəkʸɨr - ʷ..n/	-->	[yɨgʸəkʸɨnn]	'he arranges it'
b.	/yɨ-məkʸɨr - ʷ..n/	-->	[yɨmʷəkʸɨnn]	'he burns it'
c.	/kʷərkʷɨm - ⁱ/	-->	[kʷɨrkʸɨm]	'hit on the head! (f)'
d.	/qʷɨm - ⁱ/	-->	[qʸɨm]	'stop! (f)'

I conclude that palatalization in the impersonal and in the plural forms of PWG only targets final coronal obstruents, reflecting an early stage in which only coronal obstruents could undergo palatalization across Western Gurage and Muher (and Goggot if Leslau (1979) is correct). This represents <u>Stage I</u> in the development of secondary articulation.

5. Type B Verbs

Type B is a group of verbs which have gemination of the medial radical in geminating languages (i.e. Eža, Amharic). In Gurage, they are also characterized by either palatalization of the initial or medial radical or by a front vowel in the first vocalic position of the stem in the perfective and imperfective forms. Rose (1992a,b) and Degif Petros (1992) propose that Type B verbs are quadriliterals, the second segment being /i/ or /y/, triggering palatalization.[3] In Western Gurage, initial velars and coronal obstruents are palatalized (23). This is also true in Muher, but Goggot has only palatalized coronals:

(23)   Chaha

		Perfective	**Imperfective**	
a.	√dygm	jəkəm-ə-m	yɨ-jəkɨm	'strike'
b.	√sygt	šəkət-ə-m	yɨ-šəkt	'repair'
c.	√kys	kʸəsəs-ə-m	yɨ-kʸəsɨs	'be near'
d.	√gytm	gʸətəm-ə-m	yɨ-gʸətɨm	'sell on credit'

When the initial consonant is labial or a coronal sonorant and the penultimate stem consonant is velar, the latter is palatalized (except in Endegeñ where only initial velars may be palatalized):

(24)

		Perfective	**Imperfective**	
a.	√mygr	məkʸər-ə-m	yɨ-məkʸɨr	'burn'
b.	√rykm	nəkʸəm-ə-m	yɨ-rəkʸɨm	'mount a horse'
c.	√byqr	bəqʸər-ə-m	yɨ-bəqʸɨr	'brew'

Finally, when the first consonant is not palatalizable, that is, labial or coronal sonorant, and the second consonant is coronal, the vowel /e/ appears following the initial consonant (except in Eža, where there is /ə/):

(25)		Perfective	Imperfective	
a.	√bysr	besər-ə-m	yɨ-besɨr	'observe'
b.	√myTs	meTəs-ə-m	yɨ-meTɨs	'break a rope'
c.	√fyrq	fenəq-ə-m	yɨ-fenɨq	'burp'

Soddo has no palatalization, but an /i/ vowel in the first vocalic position (Leslau, 1968b, 1979; Hetzron, 1977):

(26)	Perfective	Imperfective	
	zibbər-ə-m	yɨ-zibbɨr	'return'
	tikkəm-ə-m	yɨ-tikkɨr	'take a mouthful'
	gibbər-ə-m	yɨ-gibbɨr	'pay taxes'
	miTTər-ə-m	yɨ-miTTɨr	'choose'

Type B verbs represent <u>Stage II</u> of the development of palatalization in Gurage, during which palatalization was extended to velars in Western Gurage and Muher.

6. 2nd Singular Feminine

Finally, let us reconsider the 2nd singular feminine subject forms. The palatalization in these forms is clearly due to an /-i/ suffix, which is found in Amharic optionally, and in Ge'ez and Tigrinya. In Western Gurage, not only are coronal and velar obstruents palatalized in these forms (27a-d), but coronal sonorants and vowels, too (27e-h, 28):

(27) <u>Chaha</u>

		2sg masc	2sg fem	
	a.	kɨft	kɨfč	'open!'
	b.	dɨrg	dɨrgʸ	'hit!'
	c.	nɨxəβ	nɨxʸəβ	'find!'
	d.	qɨfɨf	qʸifɨf	'cut the edges!'
	e.	siβɨr	siβiy -> siβi	'break!'
	f.	sɨrəf	sɨref	'be scared!'
	g.	Tɨməm	Tɨmem	'be contrary!'
	h.	kɨtɨf	kɨtif	'chop meat!'

(28) <u>Inor</u>

	a.	tən	təň	'come!'
	b.	tɨ-mʷən	tɨ-mʷəň	'you are jealous'

The 2nd singular feminine represents <u>Stage III</u>, the final stage of palatalization development. At this stage, the palatalization of coronal sonorants was introduced in Western Gurage, and the palatalization of coronal obstruents was introduced in Soddo.

7. Synchronic Representation of Palatalization and Labialization

In this section, I will briefly discuss the representation of palatalization and labialization from a synchronic point of view. I will propose that the three stages correspond to two different representations of floating features or segments.

7.1. Features [labial] and [posterior]

Labialization may be represented as a floating feature or node [labial], which attaches to preceding consonants:

(29)

```
                        [labial]
                       /
        d    n    g
        |    |    |
        C v  C v  C        -->  dənəg^W
```

If no labializable consonant is found, it is not realized and remains floating:

(30) a. nəTər-i-m 'one melted'
      b. q^yəCə-i-m 'one measured'

Palatalization in the impersonal and PWG plural forms is also a feature: [posterior], which attaches to segments bearing a Coronal node or feature. It targets final coronal obstruents only. If none are found, it is not realized.

(31)

```
            [labial]    [posterior]
               |       /
        k    f    t
        |    |    |
        C v  C v  C        -->  kəf^Wəč
```

7.2. Floating Segment /i/

In contrast, palatalization in Type B and in the 2nd singular feminine, is a full floating segment /i/ or /y/. As such, it must be realized, and will be realized on the vowel as a last resort.

(32)

```
                                      i
    k    f    t << i             k    t  |  f
    |    |    |                  |    |  |  |
    C v  C v  C   --> kifč       C v  C v  C   --> kitif
```

This approach also explains why palatalizing velars will overpower labialized velars, and not vice versa. Realizing (or licensing) the full segment /i/ is more important than realizing a single feature such as [labial].

8. Conclusion

In this paper I have proposed three stages in the development of palatalization in Gurage: I. coronal obstruents (Impersonal, PWG Plurals); II. velars (Type B) and III. coronal sonorants (2nd singular feminine) in Western Gurage and coronal obstruents in Soddo. These results are summarized in the following table:

(33)

	Impersonal	Plurals	Type B	2s.fem.
CWG	Cor Obs		Cor Obs, Vel	Cor Obs, Vel, Cor Son
PWG	Cor Obs	Cor Obs	Cor Obs, Vel	Cor Obs, Vel, Cor Son
Muher	Cor Obs		Cor Obs, Vel	Cor Obs, Vel
Goggot	Cor Obs		Cor Obs	Cor Obs
Soddo				Cor Obs

In addition, I have proposed a new analysis of the impersonal form of the verb which does not rely on the interaction of labialization and palatalization, but rather reflects the stages of palatalization development.

Notes

1. Palatalization in the converb or pseudo-gerundive operates in parallel fashion to the 2nd singular feminine forms (Leslau 1969, Hetzron 1975). However, it is unlikely to result from the vowel /i/ in the Ge'ez converb səbir-, as Hetzron (1975) suggests, since palatalization of coronals occurs preceding the trigger in the language (see section 3), and the feminine forms are clearly suffixal. It is more likely to have originated from a suffix /-i/ whose origin remains obscure at present. Note also that I have not discussed all cases of verbal secondary articulation, for lack of space. Labialization occurs with the 3smasc. light object marker in Western Gurage, as discussed in section 4.3. Palatalization and labialization parallel to the impersonal and plural forms are found in the PWG infinitive.
2. All central vowels adjacent to labialized consonants in Inor are realized as rounded vowels. Thus kəfʷəj-u-m is phonetically kofoj-u-m (Prunet 1991 and Paradis & Prunet (to appear)). In the form səpʷə-m, the /r/ is deleted.
3. Grover Hudson objected to this portrayal of Type B verbs as quadriliterals on the grounds that these verbs parallel Akkadian and other non-Ethiopic South Semitic geminating triliteral forms. However, there are several reasons to believe that they are quadriliterals in Gurage, although their origin may have been a geminating triliteral with a vowel /e/: (1) The pervasive palatalization is akin to weak triliterals such as səčə-m 'to drink' (Chaha) from the root √sty (Tigrinya: Sətəyə). (2) The close parallel with labialized forms, such as bʷənəsəm 'to feel alone' which have almost identical conjugation patterns (Rose 1992a,b, Degif Petros 1992 who labels these Type D). (3) The large number of regular quadriliteral forms in Ethio-Semitic not found in languages like Arabic suggests that triliterals were augmented in different ways: with a coronal sonorant in regular forms (approximately 75% of regular quadriliterals in Chaha have /r/ in C2), with /y/ in Type B, with /w/ in Type D, and perhaps with /a/ in Type C.

References

Bahru Lilaga. 1988. The morphophonemics of nouns and verbs in Enor. Addis Ababa: Addis Abba University MA Thesis.

Bhat, D. 1978. A general study of palatalization. Universals of language Vol. 2: Phonology, ed. by J. Greenberg, C. Ferguson and E. Moravcsik, 47-92. Stanford, CA: Stanford University Press.

Degif Petros. 1992. La dérivation verbale en chaha. Montreal, Quebec: Université du Québec à Montréal MA Thesis.

Elmedlaoui, Mohamed. 1992. Aspects des représentations phonologiques dans certaines langues chamito-semitiques. Rabat, Morocco: Université Mohamed V Thèse de doctorat d'Etat.

Goldenberg, Gideon. 1969. Kistaniñña. Studies in a Northern Gurage language of Christians. Orientalia Suecana 17. 61-102.

Habte Mariam Marcos. 1974. Palatalization in Ennemor. Proceedings of the IV Congresso Internazionale di studi Etiopici, 251-265. Roma: Accademia Nazionale dei Lincei.

Hetzron, Robert. 1968. Main verb-markers in Northern Gurage. Africa 38. 156-72.

Hetzron, Robert. 1971. Internal labialization in the tt-group of Outer South-Ethiopic. Journal of the American Oriental Society 91.2. 192-207.

Hetzron, Robert. 1975. The t-converb in Western Gurage (The role of analogy in historical morphology). Afroasiatic Linguistics 2.2. 39-50.

Hetzron, Robert. 1977. The Gunnän-Gurage languages. Naples: Istituto Orientale di Napoli.

Hetzron, Robert and Habte Mariam Marcos. 1966. Des traits pertinents superposés en ennemor. Journal of Ethiopian Studies 4. 17-30.

Leslau, Wolf. 1967. The impersonal in Chaha. To Honor Roman Jakobson. Essays on the occasion of his seventieth birthday, 1150-1162. The Hague: Mouton.

Leslau, Wolf. 1969. The pseudo-gerundive in Chaha. Rassegna di studi etiopici 23. 27-42.

Leslau, Wolf. 1968a. Toward a classification of the Gurage dialects. Journal of Semitic Studies 14.96-109.

Leslau, Wolf. 1968b. Ethiopians speak. Studies in cultural background. Vol. III (Soddo). University of California Publications. Near Eastern Studies 11.

Leslau, Wolf. 1979. Etymological dictionary of Gurage (Ethiopic). Wiesbaden: Otto Harrassowitz.

Leslau, Wolf. 1981. Ethiopians speak. Studies in cultural background. Vol. IV (Muher). Äthiopistische Forschungen 11. Wiesbaden: Franz Steiner Verlag.

Lieber, Rochelle. 1988. Configurational and nonconfigurational morphology. Morphology and modularity, ed. by M. Everaert, A. Evers, R. Huybregts and M. Trommelen, 187-215. Dordrecht: Foris.

McCarthy, John. 1983. Consonantal morphology in the Chaha verb. Proceedings of WCCFL 2. 176-188.

Paradis, Carole & Jean-François Prunet (to appear) A reanalysis of velar transparency cases. The Linguistic Review.

Polotsky, Hans. 1938. Etudes de grammaire gouragué. Bulletin de la Société de linguistique de Paris 34.2. 137-175.

Polotsky, Hans. 1951. Notes on Gurage grammar. Jerusalem: The Israel Oriental Society.

Prunet, Jean-François. 1991. A note on dialectal variation in Ennemor. Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday, ed. by Alan Kaye, 1233-1242. Wiesbaden: Otto Harrassowitz.

Rose, Sharon. 1992a. La palatalisation du Type B en chaha. Paper presented at Journées d'Etudes chamito-sémitiques, Institut catholique de Paris, France.

Rose, Sharon. 1992b. De la palatalisation en chaha. Montreal, Quebec: Université du Québec à Montréal MA Thesis.

Rose, Sharon. 1993. Palatalization, underspecification, and plane conflation in Chaha. Proceedings of WCCFL 12.

The Dynamics of Morphotactic Change in Sango

William J. Samarin
University of Toronto, Emeritus

1. The Argument

Social factors, no less than biological ones, play a role in determining the nature of language at one level or another.

While this verity is not presented here for the first time, it remains to be documented in a comprehensive manner. Moreover, while social factors have been invoked in works of a sociolinguistic nature to *explain*, as writers have been wont to say, the distribution of linguistic variables, there are still far too few attempts to really explain why certain languages are what they are in what might be called socio-psychological terms.[1] And whereas it is sometimes possible to arrive at satisfying *post hoc* explanations, only infrequently—if not rarely—do we have the privilege of being witnesses to substantive linguistic change. This work is the report of one such witness about one change that has taken place in Sango, the lingua franca of the Central African Republic.

I propose, in the following order of probability, that Sango, has, as *langage pratiqué*, acquired number agreement in about thirty years, that this innovation was made familiar to the inhabitants of Bangui by radio broadcasters and was adopted by other speakers because it was stylish or fashionable, and that it is more likely based on French than on any other language. Now in the pool of linguistic patterns in Bangui, the nation's capital, and other urban centers, it is being adopted by the young, a practice apparently reinforced by some subconscious notion of what constitutes a noun phrase.

If this proposition can be argued convincingly, one would have contributed to understanding something about the dynamics of language change through language contact. Some light will also have been shed on the alleged role that so-called creolization has in the evolution of a pidgin to what is considered a creole.

2. The Language

Creolization is relevant to the present discussion, because the language has a pidgin in its ancestry. This pidgin arose in the context of colonization, when the Ubangi River basin was invaded by representatives of the Congo Free State in 1887, followed very quickly by military personnel acting on behalf of the French government and those interested in trade, not only French and Belgian but also Dutch. As Belgian and French nationals competed with each other for the occupation of the territory, they penetrated to the headwaters of the Ubangi River. This was the home of the people whose language, in a pidginized form, came to be used as a medium of communication by the foreigners.

There, where the Mbomu and Uele rivers join to make the Ubangi, are found a large population speaking mutually intelligible dialects of a language having no single name: these are Dendi, Yakoma, Ngbandi (or Mongwandi), and Sango (the latter not to be confused with the pidgin that emerged). Because of their riverine habits, and being involved even before Europeans arrived in regional trade, they played an important role in colonization, providing, for example, canoes and canoers.

The Ngbandi, as they might be called collectively, speak a Ubangian language, to which family most of the other languages of the Central African Republic belong. Only a few other non-Ngbandi languages were found along the banks of the Ubangi at that time. Starting from Bangui, going eastward, they were Ngbaka (also known as Ngbaka-Ma'bo), Gbanziri, and a few Banda dialects in between these. Typologically the same, these languages, with the possible exception of Banda (for which see below), could not have contributed number agreement to the pidgin, because their grammars are without it.[2]

However, some of the other languages brought by the expatriate African workers and militia of the Europeans do have number agreement. These were Bantu languages spoken in the lower Congo area and along the Congo (now Zaire) River as well as Swahili. But Bantu languages were not the only ones that were spoken by the multilingual work force, because African personnel came also from what have become Senegal, Liberia, and Nigeria. It was these expatriate Africans, not the Europeans, who played the most important role in creating Sango.[3]

3. The Evidence

There was no evidence when I began to learn Sango in June of 1952—nor during the eight following years as I used the language, along with Gbeya, which I began to learn in 1954—that number agreement occurred in anybody's speech.[4] Nor was there any evidence in 1962, when I travelled throughout the country to make tape-recordings of Sango in preparation for writing a grammar. There was none, that is to say, except for what I found in a few radio broadcasts. That was noted in my grammar (1967a:136), but its significance was not perceived.[5]

In 1988, however, its significance could not be missed for two reasons. First, it had become common in the speech of children and young people. Second, the plural prefix was found to occur three times and, in one instance, example (5), four times in a noun phrase, evidence that the rule was becoming stronger. In other words, the prefix (highlighted in boldface in this chapter), rather than moving to the left, so to speak, was distributed to each word that could carry it: not only the head noun but all its preposed attributives. In the following examples (1-4) only those in the left-hand column would have been used in the pidgin stage:[6]

(1) *zò* 'person' *ázò* 'people'
(2) *mbéní zò* 'someone' *ámbéní ázò* 'some people'
(3) *kótá zò* (big person) 'adult' *ákótá ázò* 'adults'
(4) *mbéní kótá zò* 'an adult' *ámbéní ákótá ázò* 'some adults'

The following phrase, used by an eight-year-old girl (a capital letter indicates gender; the figure, age), is special, because it has four prefixes, because it is the only example of its kind, and because the noun phrase itself is different in structure:[7]

(5) *ágbá*	*tí*	*ákété*	*ámérengé*	*tí*	*ákóndò* (F8)
bundles of		small	children	of	chickens
'a lot of little chicks'					

However, a few others are similar to it:

(6)· *áíta*   *tí*   *áwáli*   *nà*   *ákóli* (M12)
     siblings   of   females and   males
     'sisters and brothers'

(7) *íta*   *tí*   *mbi*   *tí*   *kóli*   *nà*   *tí*   *áwáli* (F12)
     sibling of me   of   male   and   of   females
     'my brother and sisters' (?)

(8) *à-kè*   *sárà*   *òsìì*   *áviì*   *tí*   *ádèboòʃ*(M23)
     subj.prfx.-be do   also   lives   of   debaucheries
     '(the girls) also lead lives of debauchery'

In these last four examples (5-8) we see that speakers seem inclined to use the prefix to mark a whole phrase as plural. Years ago (6) and (7), for example, would have been ungrammatical, because the nouns *wáli* and *kóli* are used in constructions such as these to mark gender in a generic sense. For this reason (7) is ambiguous. Was the speaker trying to say 'my brothers and sisters'? That could have been done by saying *áíta* ... 'siblings of ...'. Or was the speaker indeed trying to say 'my brother and [my] sisters'? In that case, one could have said *íta tí mbi tí kóli nà áíta tí wáli* (sibling of me of male and siblings of female). It is not clear what is happening in (8), from a twenty-three-year-old male, where the French words *vie* and *débauche* are being used, unless each noun was intended as plural. (Native speakers of English have been observed using the plural in ways that reveal some similarity to what is happening in Sango.)[8] Most of their peers, however, remain behind, using phrases constructed like (4) or, in my opinion, haphazardly like (9):

(9) *mbéní ákótá ázò*
     'some adults'

It would appear that they are applying a rule that they do not yet fully understand.

The last statement is supported by the fact that variation occurs between one and the same utterance in the same discourse of a single speaker with no difference in meaning.[9]

Nonetheless, there are enough examples from a sizeable corpus to conclude that something is indeed going on in Sango and has been since at least the beginning of the sixth decade of this century. From a total corpus of transcribed texts of about 150,000 words a sample of about 89,000 words was examined for the use of this prefix. This corpus comes from 254 texts and as many different speakers of both genders, ranging from three to twenty years in age. In this corpus were found eighty-three examples of phrases in which two or more instances of the prefix occurred. They come from speakers of both genders, one of whom was three years old, the others six to sixteen in age, each age being represented.[10] (Only from ages eleven to sixteen do we find noun phrases with three words pluralized.) Another corpus of about 10,000 words from radio broadcasts is discussed below.

This number is impressive partly because about a decade earlier (in the fifties or sixties) there were no such phrases at all and also because number is not an obligatory category in Sango or, it would appear, in any of the other Ubangian languages except some dialects of Banda. In the Gbaya dialect of Bossangoa called

Gbeya by its speakers), for example, it is used parsimoniously and mostly with animate nouns (Samarin 1967b). This seems to have been the case with Sango too. Although inanimate nouns are now frequently marked for plurality in Sango, possibly under the influence of French, its use is neither general nor consistent.[11]

Among the many features that distinguish twenty Banda languages in the Central African Republic are the following: (1) some mark only animate nouns for plural, others both animate and inanimate; (2) in noun phrases some prefix the noun, others the attributive, and in five languages both the attributive and noun. The latter are found in two groups: the Central Group, which includes Gbaga-sud (Southern Gbaga) and Hai (found near Bocaranga in the northwest corner of the country), and the Peripheral Group, which includes Gbaga-nord (Northern Gbaga), Wojo, and Ngbugu. Whereas three options occur for Hai with respect to agreement, it is obligatory in Wojo and Ngbugu.[12] We return to these languages after another explanation is proposed.

4. The Explanation

Having established (a) that number agreement was not in the language that provided Sango with its lexical base, (b) that it cannot be attributed to any of the substrate Ubangian languages other than a Banda one, and (c) that it could have been introduced by speakers of Bantu languages but was not, we are left with trying to explain where it did come from. Who was responsible? And what was their language?

The most reasonable explanation is that Sango-French bilinguals incorporated number agreement in their Sango on the pattern of French. If this is true, the calque would have arisen in response to written French, not spoken, because the phonetic realization of the plural suffix occurs only in *liaison*. Although French borrowings in Sango have not yet been studied with respect to their oral or written sources, we can observe that *wànyoòn* (final *n* indicating nasalization) is a common pronunciation of the French word *oignon* 'onion;' in the pidgin stage it was *zònyoòn* from *des oignons*.

One might suppose that it would take an impressive number of bilinguals to introduce a new rule, but it is clear that a small number of bilingual people, for reasons we do not yet fully understand, had more importance than their numbers would have justified. Indeed, their influence has already been demonstrated for Sango: in the adoption of the word for money (Samarin 1989c), of the copula (Samarin 1986), and probably also of the negative (Samarin 1984/1985). Moreover, a small number of bilinguals inclined to marking plurality more than once would have great influence if they played important roles in their society. In the broadcasters for Radio Bangui we have exactly such a number of speakers. It is reasonable to argue that they are the ones who introduced number agreement to the people of Bangui.[13] The argument is supported by the positioning of the adverb *ngá* 'also' (see below). As for numbers of Sango-French bilinguals at independence, there could not have been very many. I certainly encountered very few in my early years in the country. My own impressions are supported by figures on the number of students in schools where French was the medium of instruction. At the end of the 1961-1962 school year, there were only fifty in the final year of secondary school (called *première*), including Europeans. Moreover, only six Central Africans (or those who were not considered Europeans) in the whole country took the *baccalauréat* examination in 1962.[14]

If the number of Sango-French bilinguals was small, the number of those who were inclined to use number agreement was presumably even smaller. That can be taken as axiomatic on the grounds that no innovation is started by all speakers of a language at the same time. The fact can also be demonstrated. First, the nation's first president, Barthélémy Boganda, was a well-educated person, one whom in the colonial period the French would have considered an *évolué*. Once a priest, a school teacher, then a territorial deputy to the French parliament and leader of the MESAN (Mouvement d'Evolution de l'Afrique Noire) political party, in his radio address in March 1959 urging the inhabitants of Bangui to vote he did not once use number agreement even though he had three opportunities to do so.[15]

By contrast, out of the seven persons who broadcast in Sango at that time, one of them a Central African priest (*abbé*), all but one used number agreement in my sample of about 10,000 words, in which there were fifteen instances of number agreement. For example:

(10) *ámbéní wáli   òta*
some   woman   three
'three women'

(11) *ákétékété ámìkroòb*
'tiny microbes'

(12) *ákótá ázò   tí sésè*
big   people of earth
'important people of the country'

Three of these speakers varied between patterns, sometimes in the same phrase. The one broadcaster who did not use plural agreement used the older construction eight times; another case is problematic. This speaker, still broadcasting today, remains 'conservative' or old-fashioned in his use of Sango—at least in my opinion.

However, we must return to those Banda languages that are characterized by number agreement. Although some might consider a substratal explanation rather than the one proposed here, it does not appear to have much in its favor. First, if speakers of these languages introduced number agreement into Sango, it would have to be demonstrated that they did this in their own regions. No such evidence is available. Second, it would have to be demonstrated, assuming that in Sango they used plural agreement as in their ethnic languages, that they were in the position to influence other speakers. Here also no evidence is available. What is very attractive is the role that Ngbugu might have played, because it is located immediately north of Mobaye, the center for the speakers of vernacular Sango, one of the dialects that make up the cluster that served as the basis of the pidgin. There is no evidence that the Pidgin Sango of this area was ever any different with respect to marking than any other form of the lingua franca. One could, of course, hypothesize that it was a native speaker of one of these few Banda dialects whose tendency to use plural agreement on radio broadcasts, and possibly elsewhere, was adopted by his colleagues. If such was the case, we are still left with radio broadcasters as being responsible for disseminating the innovation to the general public.

Having found reasons for believing that number agreement was introduced by Sango-French bilinguals thirty-some years ago and that it was introduced to the

population of Bangui over the air, in the case of some people perhaps only reinforcing what they were already doing, the argument could be left there. It might be too bold to attempt an explanation as to what motivated the innovators, apart from bilingual competence. However, a study of the radio archives would probably demonstrate that they had created a certain 'radio style' of Sango. Today certainly, there is a 'radio Sango' that is, in my opinion, quite different from the vernacular (that is, what one hears in ordinary usage).[16]

It is not, however, too bold to suggest that people are influenced by the variety of Sango they hear on the radio, because the broadcasters make up part of the society's élite. In support of this argument one would like various kinds of information about radio broadcasting in the country. This, unfortunately, is not available. Only the following can be given at this time. (1) In 1974 85 percent of broadcasting time was in French, not Sango. (2) In 1975 of 500 persons interviewed in the *quartier* of Boy-Arabe 50.6 percent said that they listened to the radio frequently (*souvent*); these appear to have been younger persons, because the 28.2 percent who said that they listened to the radio rarely or never, were persons of age 45 or more; 45.5 percent said that they preferred listening to broadcasts in Sango to those in French (Déchamps-Wenezoui 1981:111).[17] (3) In a sociolinguistic study of 113 students in Bangui in 1987 (Gerbault, ms) it was found that the second highest score for positive attitudes with respect to the Sango of various categories of persons (e.g., journalists, missionaries, teachers, nurses) was found with radio broadcasters (*journalistes*): i.e., 78.76 percent of the subjects. (The slightly higher percentage was for judges: i.e., 79.64 percent.)[18] (4) Two persons of high status whom I have worked with have used a radio style, one a male catechist of St. Paul's Cathedral. After I had recorded something from him, I remarked on the similarity of his manner of speaking with the radio's. The explanation, he told me, was that he had a friend who was a broadcaster. The other was a forty-year-old man with a B.A. from an American university and a certificate in translating from a French university whom I employed to translate extemporaneously a sermon from French into Sango. He had never been a broadcaster.

In assessing the influence of radio Sango one will have to take into consideration the radio's history. In the present study we are attempting to go back as far as possible.[19] Before 1983 there was, as far as I know, no attempt to use the radio to effect changes in Sango. In that year, however, and again in 1985 the Central African linguist, Marcel Diki-Kidiri, had workshops in Bangui to train broadcasters to 'improve' their Sango. He reported that they had succeeded both in vocabulary and in syntax (personal communication, Cologne, September 1992). It is not known what syntactic changes were recommended; in vocabulary there were proposed many neologisms to replace French loans, like *tonda* (tones unknown) for *commencer* 'to begin', and *pàkàrà* for *monsieur*.[20] (On the introduction of neologisms see also Samarin 1980.) In the last ten years these changes have been deliberately and systematically taught on the radio. It is no wonder, then, that Central Africans might be convinced that, as Diki-Kidiri reported (1992), 'la radio-télévision centrafricaine est considerée aujourd'hui comme l'un des lieux où le sango est le meilleur et le plus moderne.'[21]

5. Supporting Evidence

To demonstrate further the influence of the speech of radio broadcasters one can

bring evidence about what has happened to the distribution of *ngá* 'also'. Whereas it used to be found at the end of the clause, except for other uses not relevant here, it has been found immediately following the verb in speech heard or recorded in Bangui. It was at first assumed that the new position was semantically motivated: some sentences could, it seemed to me, be translated with 'even', as in (14). For example:

(13) *mbi báà   lò   ngá*   *apè*
       I   see   he/she   also   not
     'I also didn't see him/her'

(14) *mbi báà ngá lò apè*
     'I didn't even see her'

But there were too many problematic sentences. Besides, there were sentences like (15) in which *ngá* is repeated, as if the speaker were using the adverb haphazardly:

(15) *lò mu   ngá wáli   ngá nà ngú   só à-língbì*
       he take   also woman   also at water   this subj.prfx. be-able
     'he marries the girl at the proper age'

In order to see if the positioning and the repetition of *ngá* might be due to the influence of broadcasters, I checked the broadcast recordings. The hypothesis was confirmed: the same kind of correlation was found as with number agreement. Out of seventeen examples of sentences in which the adverb could have preceded a complement, it did appear there fifteen times. Moreover, one female broadcaster is responsible for eight of these, all of them of the marked pattern. In other words, at the time when tape-recordings of extemporaneous speech in Sango from throughout the country revealed one pattern, radio broadcasts revealed an additional one.

French competence might here also be the causative or disposing factor. If so, it would be reinforced by a substratal or adstratal one, because Ngbandi allows a post-verb position (Lekens 1958[2]:651). Its equivalent in Sango would be as follows, where final *n* represents a nasalized vowel:

(16) *lò   kè   kèn   ngá mbi*
       he   is   refuse   also me
     'he's refusing me also'

It is not surprising therefore that the speaker who produced (14) is also Yakoma, a twenty-four-year-old male. Indeed, the 'new' position is preferred by him.[22] The ethnicity of the female broadcaster mentioned in the preceding paragraph is not known; only that she was from Cameroun.

6. Conclusion

This study appears to be one of the few on the actual process of creolization: that is, change in progress (Mühlhäusler 1986:205-206). There have been, of course, many such studies for 'natural' languages in the last thirty years, falsifying this statement about change in African languages: 'changes will be almost imperceptible within one generation' (Welmers 1970:2). Let us note the following.

• From this study of the history of the use of the 'plural' prefix in Sango we learn that a few bilingual speakers can contribute to language change when they are considered models of language use. Fashion or style plays a role in language change, since all change assumes that some people are adopting an innovation by others for one reason or another. What the present study contributes, however, is evidence of morphotactic and syntactic change motivated by wanting to be fashionable.[23]

• We must recognize that number agreement is not necessarily an established rule in the grammar of Sango.[24] It seems safe, however, to assume that its use will become more and more prominent and that it will spread to rural areas (as can already be demonstrated). For some time, however, it will probably remain a spoken phenomenon. Indeed, it may for a long time be one of the many features that will distinguish spoken and written Sango.[25]

• This study ought to lead linguists to be cautious about claiming that when a pidgin becomes a creole it becomes more 'complicated,' somehow meeting the 'needs' of the people for whom it is now the first language. The emergence of number agreement in Sango is not—or not necessarily—an instance of the linguistic consequences of creolization (especially not when this is teleologically or functionally conceived). The following generalization, therefore, does not wholly apply to Sango: 'As the usefulness of a pidgin grows and as its functions extend the lexicon increases and the syntactic properties are refined' (Todd 1974:58; see also Valdman 1977:155-157, and the word of caution in Mühlhäusler 1986:204). The change discussed in this study of Sango is just another case of what can happen when languages are in contact.

• Finally, if radio broadcasters have indeed introduced changes in Sango, we have evidence that has been lacking on this modality of change. Up until now it has been denied that the media could have any effect. Thus: 'It appears that relatively few speakers are directly influenced by the speech patterns heard on radio and television' (Labov 1967:74fn).

Acknowledgements

Notes

1. Without attempting to be anywhere near exhaustive in citing works on this topic, one might include Dozier's explanation for the differences between the Tewa and Yaqui response to Spanish relations (1956) and the role of ideology in shift to Tok Pisin in Papua New Guinea (Kulick 1992).
2. This statement is endorsed by the linguist F. K. Erhard Voeltz, who spent

several years in the Central African Republic. Having 'worked quite a bit on the noun phrases of Ubangian languages,' he found that 'generally, especially if a discussion [presumably with language consultants] regarding the data is allowed, only one sign for the plural is accepted in nearly all of the languages, including Sango.' However, he reports that whereas in Gbayi (Kpatiri) one pluralizes the noun phrase only once, there are instances of other patterns: e.g., the title of a fable is *á-ngbo nà déli* 'snake and toad', not 'snakes and toads' or snakes and toad' (personal communication, 25 February 1994, used with permission). Gbayi is a language related to the Ngbandi dialect-cluster, spoken along the Kotto River at and south of the village of Mingala (Boyeldieu and Diki-Kidiri 1982).

3. The first possible reference to Sango may be the one made by a missionary in 1896: he wrote about a *volapuck* that was spoken by the workers. The first published linguistic data appeared in 1906-1907 (Gaudefroy-Demombynes, a book I have not yet succeeded in finding). Grammatical notes accompanied a glossary in 1911 (Calloc'h), followed by one in 1950 (Tisserant). There is no single work on the history of Sango. My research on this topic has resulted, however, in the following works: 1982a, 1982b, 1984, 1984/1985, 1989a, 1989b. On the basis of very little information it has been suggested that Dendi was a nonpidginized lingua franca before Europeans arrived. I have argued against this position in Samarin 1985, supported by Eric de Dampierre (personal communications), who has written about this area (1967). There may have been other nascent jargons before Sango was adopted along the length of the upper Ubangi. One possibly based on Banda is mentioned in Samarin 1985.

4. By this time, of course, there was a considerable amount of literature in Sango that was prepared by Catholic and Protestant missionaries. The former had established a mission just outside the military post of Bangui in 1894; the latter did not enter the country until the 1920s. Among Protestants, the first to use Sango were those serving with Baptist Mid-Missions (an organization supported by cooperating churches of the General Association of Regular Baptists). Their stations were distributed from the middle of the territory and eastward, and they used Sango exclusively. To the west were found missionaries with the Mission Evangélique de l'Oubangui-Chari (M.E.F., of the Grace Brethren Church in the United States) and the Swedish Baptists. Until just after the Second World War these used indigenous vernaculars in their work. The New Testament had been published in the 1930s; the manuscript of the Old Testament was nearly finished in the late 1950s and was subsequently published.

5. Copies of that grammar, produced by photo-offset, were sent to Africanists and libraries of universities where there were African studies programs in 1963.

6. The orthography being used here, like the orthography endorsed by Marcel Diki-Kidiri, does not distinguish between upper and lower mid and back vowels. In any case, there is so much variation between speakers that in the absence of a sociolinguistic study of the matter writing one or the other phonemically would be arbitrary. Diki-Kidiri's comments on variation (1982), although more thorough than my own of 1967a, are based, like my own, on personal experience, not careful research. See Bouquiaux et al. 1978 for another orthography. There are three register tones: high, mid, and low, indicated as in Bouquiaux et al., but not as endorsed by Diki-Kidiri. The examples are written in their uncontracted forms. Both assimilation and contraction characterize the phonological change that is found in urban and urban-influenced Sango (see Samarin in press, Walker and Samarin, in press): e.g., *ámbén ázò* (the two words pronounced without hiatus, *n* representing here the consonant) 'some people' (see [2] above).

7. One wonders if the child is not using a superlative here, when she shoots all her cannons, so to speak. In the following sentence she correctly notes that there are ten chicks with the hen (not hens). The context of the recording is having her comment on pictures in a French-language book for children. She was describing a picture 5.5" x 4" (13.97 cm x 10.16 cm). This would be a large number for most if not all Central African children. I certainly never saw a brood of ten, rarely over five, and usually just two or three. Hens have a difficult time bringing chicks into the world. To add another touch to this girl's portrait, we might note that she also says *kétá kétá méyengé tí ákóndò*, in which there is no plural on the word for 'child' and assimilation produces something resembling affixation in the word for 'small' (unless, of course, one prefers *két ákét áméyengé*). However, F. K. Erhard Voeltz (see fn. 2) reports that with noun phrases like (5) he generally got from his consultants more than one plural marking, the case depending on the meaning: e.g., 'child of chicken', 'children of chicken', 'child of chicken (pl.)', 'children of chicken (pl.)', adding that 'All four cases were generally possible in the Ubangi languages of the CAR with four different sort of plural/singular markings.' While respecting my colleague's competence as a linguist and expertise in Ubangian languages, I must demur by saying that I have found that educated native speakers of Ubangian languages, which his consultants for the most part were, I believe, allow Sango to influence their native languages.

8. Here are a few examples collected only recently, offered without analysis and interpretation, italics added: 'They all have terrific *senses* of humor' (Bill Maher, 'Larry King Live,' January 26, 1994); '*these types* of *players*' (head coach, Syracuse University, New York, television interview, September 18, 1993); 'And after the claymores, the *sounds* of the *enemy's* dying—the cries' (Wentz and Jurus 1992:252); 'Someone put them [junk-mail flyers] in everybody's *boxes*' [one mailbox per house] (WJS); 'Most *creoles* do not even have *pidgins* in their *ancestries*' (WJS); 'Have you got our *various kinds* of *waters*?' (with reference to drinking water and water for painting, said probably in exaggeration or irony, July 20, 1993, WJS); 'The construction trucks left all *kinds* of oil' (Ruth Samarin); 'Where's the popcorn?'— 'It's up there with the chips *and stuff like that*' (where in Sango the plural prefix on the preceding noun can have this meaning).

9. During a period of language change, of course, speakers can be pushed into a corner by an aggressive foreigner into thinking up an explanation. Such seems to have been the case when Peace Corps volunteers were informed in Bangui that with the placement of *á-* one emphasizes ('*insiste*') either the adjective (*big* houses) or the noun (big *houses*). The Central African assistant's examples suggest that only one modifier occurs with a noun.

10. The three-year-old girl used these pluralized phrases: *ámbéní áfàmiì* 'some relatives', *ágbá tí áíta tí i* (bundles of siblings of 3p.prn.) 'a lot of our siblings (or, friends)'. For the latter, compare in English *a lot of* vs. *lots of*.

11. A more comprehensive study of the prefix is now under way. See also Samarin 1994.

12. I am grateful to France Cloarec-Heiss for providing me with this information, based on her forthcoming work on Banda dialects. See also Cloarec-Heiss 1986.

13. Although a Lingala radio style in Kinshasa is well recognized (*Lingala ya ba-je-le-connais*), it is considered amusing and is used only jocularly; moreover, it has not influenced the spoken language (Salikoko Mufwene, personal communication, 21 January 1994, used with permission).

14. The information was provided on July 12, 1962 by a Msr. Dartois, attached to the Ministry of Education. It can be assumed that at independence there were

secondary schools only in Bangui. In 1962 only about 450 in the whole country, including Europeans, passed the *examen d'entrer à sixième*, thereby qualifying for entrance into secondary school. In any case, in the 1950s neither the Brethren nor the Mid-Missions missionaries had French schools. At that time in the District (now Sous-Préfecture) of Bossangoa, with about 80,000 inhabitants, there was one 'official' school and one at the Catholic mission in the town of Bossangoa and a primary school at a cotton experimental station about thirty kms away.

15. This was just seventy years after the French had established their post at Bangui.

16. Most striking is the intonational patterns. What is curious is that the most salient one of dragging pitch upward at the end of some kind of a unit is one that I have never heard in France. But other features could be cited. Radio Sango deserves a study of its own. With respect to people's attitudes, I might mention that while riding in taxi-busses, in which the radio is usually played, I have imitated and commented on something that is characteristic of radio Sango in an attempt to elicit criticism or approval. The response has always been silence. I explain this by the unwillingness of the others to criticize their compatriots. One might suppose that they might have been willing to express an opinion in private; even in this setting I have not had success. One should not conclude from these observations that Central Africans do not have opinions; only that one must be skilful and patient in eliciting them.

17. The author's name has elsewhere appeared as Wenezoui-Dechamps.

18. Although one can learn something by interviewing subjects—otherwise, I would not be using this method myself—one must be critical of the methodology and the results. I find, for example, this appreciation of the speech of judges fanciful. Whereas all of the 113 persons can be expected to have heard the radio many times in their lives, how many would have had occasion to hear a judge speak? I would say, agreeing with Gerbault, that the answers of the subjects depended on their expectations: whatever it was that they considered 'good Sango,' it was spoken by persons who would have a high competence in French. I hold this view in spite of the fact that in this study Gerbault found that negative opinions were expressed concerning *Fransango*, i.e., speech characterized by an admixture of French and Sango. For example, when three kinds of language are compared—French, *Fransango*, and Sango—with respect to amount of education (*degré d'instruction*) the percentages are 44.35 percent, 38.53 percent, and 32.51 percent respectively; for residence in Bangui, 38.91 percent, 45.75 percent, and 42.27 percent respectively; for success, 33.44 percent, 28.62 percent, 34.51 percent respectively.

19. It has been reported, for example, that in the early history of Radio Rural the most popular broadcasters were Lucien Dambalet and Passy Wilibyro, possibly only because they dealt with topics of greatest interest to listeners (Nzapayeké 1987:124). One of these is the telling of Central African traditional tales.

20. In Ngbandi *pàkàrà* is used to refer to someone whose name one does not know or wants to avoid using for one reason or another (Lekens 1958).

21. Informed by occasional visits to the Central African Republic, Diki-Kidiri however has not, as far as I know, made a controlled study of people's attitudes towards the Sango spoken by all or different radio broadcasters. One study undertaken in 1986 by Radio Bangui in the rural areas found that most villagers had difficulty in following the Sango spoken on the radio ('la plupart de ceux-ci [villagers in group discussions] n'ont pas manqué d'insister sur les difficultés qu'ils éprouvent à suivre le sango parlé à la radio' [Nzapayéké 1987:142]).

22. The recording, obtained by my assistant Lamine Ndocko, is an animated conversation on the topic of marriage between this young man and another. However, the speaker also used the adverb in this position once.

23. One would like to say much more about style and linguistic fashions in Sango, but the topic has not yet been sufficiently studied. Here are a couple of examples that illustrate the way linguistic usage can change over time: (1) the virtual disappearance of the connective *sí* as a topicalizer (Samarin 1967a:104ff) in the speech of Bangui's youth, possibly linked to the next change; (2) an increase in the use of the topicalizer *làá* (now simply *lá* or *la*) to the point where its function seems weakened (Samarin 1967a:104ff); (3) the use of *gbá* 'bundle' for 'a lot', replacing *míngì*. The following mean 'a lot of people' or 'many people': *ázò míngì* (old), *gbá tí ázò* (new).

24. Since this study focusses on the origin of number agreement in Sango, not its distribution, and since a thorough study of its use by a sample of the population has not yet been pursued, it seems premature to talk about a so-called variable rule in the language (*langue*). However, in interviewing 150 persons in Bangui and ninety-eight in and around Nzoro, a rural village in the northwest corner of the CAR, about 550 km from Bangui, it was found in 1992 that 59.33 percent of the former and only 9.18 percent of the latter preferred the sentence with agreement in the sentence *ámbéní (á)màma àhíngà tí bàtà ámérengé tí álà nzoní* 'some mothers know how to care for their children well'. The optional prefix is parenthesized. Only two of the subjects were female (aged eleven and twenty-six); the ages of the males ranged from twelve to forty-five. None of the rural subjects had ever been to Bangui. Some of the villagers had radios, of course.

25. Two pieces of literature published by L'Assemblée Spirituelle Nationale des Bahá'ís de la République Centrafricaine were examined for instances of number agreement: *A Tene ti Nzapa* (1989) and *Vìngö lö tî Nzapä* (1990). They are of approximately equal size, having about 9,500 words in them. Only the first is characterized by number agreement, but it has only one instance (p. 47): *a sengue azo aga a kota azo* (pl. ordinary people subj.pref-come pl. big people) 'ordinary people become important people'. It is interesting that the word 'people', *azo*, is written solid, but that the prefix with the words *sengue* and *kota* is separated. One assumes the operation of some folklinguistic justification. See note 9. (The diacritics in the second publication indicate tone.)

References

BOUQUIAUX, LUC, et al. 1978. Dictionnaire sango-français et lexique français-sango. Paris: Société d'Etudes Linguistiques et Anthropologiques de France [SELAF].

BOYELDIEU, PASCAL and MARCEL DIKI-KIDIRI. 1982. Le domaine ngbandi. Bibliothèque de la SELAF, no. 93. Paris: SELAF [Société d'Etudes Linguistiques et Anthropologiques de France].

CLOAREC-HEISS, FRANCE. 1986. Dynamique et équilibre d'une syntaxe: Le banda-linda du Centrafrique. Paris, Cambridge: SELAF-CUP.

DÉCHAMPS-WENEZOUI, MARTINE. 1981. Le français, le sango et les autres langues centrafricaines: Enquête sociolinguistique au quartier Boy-Arabe (Bangui, Centrafrique. Paris: Société d'Etudes Linguistiques et Anthropologiques de France (SELAF).

CALLOC'H (RÉV. PÈRE) J. 1911. Vocabulaire français-sango et sango-français, langue commerciale de l'Oubangui-Chari, précédé d'un abrégé grammatical. Paris: Librairie Paul Geuthner.

DAMPIERRE, ERIC DE. 1967. Un ancien royaume bandia du Haut-Oubangui. Recherches en Sciences Humaines 24. Paris: Librairie Plon.

DIKI-KIDIRI, MARCEL. 1982. Variations libres entre phonèmes en sango langue nationale de R.C.A.: Étude et mesures. In Le domaine ngbandi, by Pascal Boyeldieu and Marcel Diki-Kidiri, 83-103. Bibliothèque de la SELAF, n. 93. Paris: SELAF [Société d'Etudes Linguistiques et Anthropologiques de France].

———. 1992. Sango: Its status and uses in the Central African Republic. Presented at the International Colloquium on Sango at the University of Cologne, Germany, 3-4 September. [To appear in the Proceedings, ed. by Helma Pasch.]

DOZIER, EDWARD P. 1956. Two examples of linguistic acculturation: The Yaqui of Sonora and Arizona and the Tewa of New Mexico. Language 32: 146-57. Reprinted in Language in culture and society, ed. by Dell Hymes, 509-517. Cambridge: Cambridge University Press, 1964.

GERBAULT, JEANNINE. n.d. Modes d'appropriation langagière en République Centrafricaine. Université de Bangui. Ms. 36 pp.

GAUDEFROY-DEMOMBYNES, M. 1906-1907. Documents sur les langues de l'Oubangui-Chari. Actes du XIVè Congrès International des Orientalistes. Paris: Leroux.

KULICK, DON. 1992. Language shift and cultural reproduction: Socialization, self, and syncretism in a Papua New Guinean village. Studies in the Social and Cultural Foundations of Language 14. Cambridge: Cambridge University Press.

LABOV, WILLIAM. 1967. The effect of social mobility on linguistic behavior. In Explorations in sociolinguistics, ed. by Stanley Lieberson, 58-75. Bloomington, IN: Indiana University. (International Journal of American Linguistics, Vol. 33, Nu. 4, Part II, Pub. 44.)

LEKENS, BENJAMIN. 1958. Ngbandi-idioticon, 2. Ngbandi en Frans-Nederlands. Tervuren: Annales du Musée Royal du Congo Belge.

MÜHLHÄUSLER, PETER. 1986. Pidgin and creole linguistics. Oxford: Basil Blackwell.

NZAPAYÉKÉ, ANDRÉ. 1987. Etude de l'auditoire de la Radio Rurale Centrafricaine: analyse des conditions sociales et techniques de l'écoute de la radio en milieu rural. Bangui: Projet Radio Rurale. G.T.Z. Coopération Centrafricano-Allemand.

SAMARIN, W. J. 1967a. A grammar of Sango. The Hague: Mouton and Co.

———. 1967b. The Gbeya language: Grammar, texts, and vocabularies. Berkeley and Los Angeles: University of California Press.

———. 1980. Review of Bouquiaux, et al. 1978. Journal of African Languages and Linguistics 2.159-64.

———. 1982a. Colonization and pidginization on the Ubangi River. Journal of African Languages and Linguistics 4.1-42.

———. 1982b. Goals, roles, and language skills in colonizing central equatorial Africa. Anthropological Linguistics 24.410-22.

———. 1984. The linguistic world of field colonialism. Language in Society 13.435-453.

————. 1984/1985. Communication by Ubangian water and word. Sprache und Geschichte in Afrika [University of Cologne] 6.309-73.

————. 1985. Plurigenesis in pidginization. Journal of Historical Linguistics and Philology 2(1/2).100-19.

————. 1986. The source of Sango's 'be'. Journal of Pidgin and Creole Languages 1(2).205-223.

————. 1989a. The black man's burden: African colonial labor on the Congo and Ubangi Rivers, 1880-1900. Boulder, Colorado: Westview Press.

————. 1989b. Language in the colonization of central Africa, 1880-1890. Canadian Journal of African Studies 23(2).232-49.

————. 1989c. The colonial heritage of the Central African Republic: A linguistic perspective. The International Journal of African Historical Studies 22(4).697-711.

————. 1994. Plurality and deference in Sango. Paper given at the 25th Annual Conference on African Linguistics. Brunswick, NJ: Rutgers University, 25-27 March.

————. in press. The creolization of pidgin morphophonology. In Pidgins and Creoles: Structure and status, ed. by Arthur K. Spears and Donald Winford. Amsterdam: John Benjamins.

TISSERANT, [Le Rév. Père]. 1950. Sango, langue véhiculaire de l'Oubangui-Chari. Issy-Les-Moulineaux: Les Presses Missionnaires.

TODD, LORETO. 1974. Pidgins and creoles. Language and society series. London & Boston: Routledge and Kegan Paul.

VALDMAN, ALBERT. 1977. Creolization: Elaboration in the development of creole French dialects. In Pidgin and Creole linguistics, ed. by A. Valdman, 155-89. Bloomington, IN: Indiana University Press.

WALKER, JAMES A. and W. J. SAMARIN. In press. The phonology of Sango. In Phonologies of Asia and Africa, ed. by Alan S. Kaye. Wiesbaden: Otto Harrassowitz.

WELMERS, WILLIAM E. 1970. Language change and language relationships in Africa. Language Sciences 12.1-8.

Diphthongization in the Refal Dialect of Lama

Charles H. Ulrich
Simon Fraser University

1. Introduction

Lama, a Gur language spoken in Togo, has two attested dialects: Refal (Aritiba 1987) and Kantè (Ourso 1989a).[1] The most striking difference between the two Lama dialects is the (diachronic) diphthongization of mid vowels in Refal. While the Common Lama mid vowels *e, *ẹ, *o,* and *ọ* have been retained in Kantè, they have undergone a startling array of diphthongizations in Refal. Refal has neither front mid vowels nor advanced mid vowels in roots, though the retracted mid back vowel *ọ is retained in certain environments.

This paper draws on my own reconstruction of Common Lama (Ulrich 1993), which is based on more than six hundred pairs of cognate noun and verb roots. This corpus was compiled by a comparison of the two aforementioned dissertations, with additional Kantè forms from Downing (1986), Kenstowicz (1989), Kenstowicz et al. (1988), Ourso (1988a, 1988b, 1989b), Yu (1991), and my own field notes.

2. Common Lama Reconstruction

Ulrich (1993) reconstructs the following consonant system for Common Lama:[2]

(1)   *p   *t   (*d)   *c   *k   *kp
             *f   *s   *h
             *m   *n   *ñ
             *w   *l   *r   *y

Consonants have undergone only a few changes in the two dialects. In Kantè, *k has been palatalized to *c* before mid front vowels.[3] In Refal, *w has become *v* before front and low vowels. And, synchronically, *s* is palatalized to ʃ before *y* (Aritiba 1987).

Ulrich (1993) reconstructs the following vowel system for Common Lama, where the underdot indicates retraction of the tongue root:

(2)   *i   *ɨ   *u
             *ị   *ụ
             *e   *o
             *ẹ   *ə̣   *ọ
                   *ạ

Kantè retains this vowel system virtually intact:

(3)   i   ə   u
        ị   ə̣   ụ
        e   o
        ẹ   ọ
              ạ

The only change is that the two non-low central vowels are aligned: both sound mid (at least to anglophone ears), but are phonologically high (Ourso 1988b, Ulrich in preparation).

Refal, on the other hand, has lost three of the Common Lama mid vowels and developed a retracted high central vowel:

(4)   i   ɨ   u
        ị   ɨ̣   ụ

              ə̣   ọ
              ạ

Within roots, Refal also has nine glide + vowel diphthongs (5), three vowel + vowel diphthongs[4] (6), and one glide + vowel + vowel triphthong (7):

(5)   yị   yɨ   wɨ
              yə̣   yọ   wə̣   wọ
              yạ   wạ

(6)   ịa   ịo   ɨị

(7)   wɨụ

As might be expected, it is the missing mid vowels that were the source of most of these diphthongs and triphthongs.[5] The correspondences are summarized as follows:

(8)   *Common Lama*   *Kantè*   *Refal*
        *e   e   yị, yə̣, ɨị, ị
        *ẹ   ẹ   yạ, ịa, ị
        *o   o   yọ, ịo, wɨ, ɨ, wɨụ
        *ọ   ọ   wọ, ọ, wạ, ạ

3. Vowel Changes in Refal

After consonants other than *y*, the advanced mid front vowel *e* has diphthongized to *yị* in root-final position (9) and to *yə̣* in non-root-final (always preconsonantal) position (10):

(9) a. K. lé-eu 'to weave' ~ R. lyí-ú 'tresser'
     b. K. kpê-en 'silos' ~ R. kpyí-n 'le grenier' (plural)
     c. K. rê-m 'sleep' ~ R. ryî-m 'le sommeil'

(10) a. K. tèm 'finish!' ~ R. tyə̂m 'finir'
     b. K. lèk-ú 'arrow shaft' ~ R. lyə̀k-ú 'le manche de flèche'
     c. K. sén-də̂ 'bean' ~ R. ʃyə́n-də̂ 'le haricot'
     d. K. tèel-à 'windows' ~ R. tyə̀l-à 'la fenêtre' (plural)

Example (10d) shows that diphthongization affected long *ee as well as short *e.[6]
     Note that it is root structure, not syllable structure, that conditions the different diphthongs. For instance, *e has become yə in the root tyə̂l-, regardless of whether the root-final consonant is syllabified as an onset (11a) or a coda (11b):

(11) a. K. tél-ṷ 'baobab tree' ~ R. tyə́l-ṷ 'le baobab'
     b. R. tyə̀l-m-o 'petite quantité de la poudre du fruit de baobab'

Compare also (9c), in which m is the class 10 suffix, with (10a), in which m is part of the root.
     After y, *e has become i̧i̧ root-finally (12) and i̧ non-root-finally (13):

(12)   K. yé-eu 'to quarrel' ~ R. yí̧í̧-ú '(se) quereller'

(13) a. K. yép-ə̂ 'to let' ~ R. yí̧p-ə̂ 'laisser'
     b. K. yélə́m 'blind person' ~ R. yí̧lə́m 'l'aveugle'
     c. K. yèn-də́ 'hippopotamus' ~ R. yi̧n-də 'l'hippopotame'

The retracted mid front vowel *ȩ has diphthongized to yạ when final in a high-toned verb root (14), and when short and non-root-final (15):

(14) a. K. lȩ́-ȩṷ 'to let go' ~ R. lyạ́-ú 'lâcher'
     b. K. sȩ́-ȩṷ 'to plant' ~ R. ʃyạ́-ú 'planter (arbre)'
     c. K. tȩ́-ȩṷ 'to throw' ~ R. tyạ́-ú 'lancer'

(15) a. K. tȩ́l-ṷ 'to escape' ~ R. tyạ̀l-ṷ 's'évader...'
     b. K. pȩ́l-ṷ 'to cut' ~ R. pyạ̀l-ṷ 'couper'
     c. K. nàmkpȩ̀t-ə̀r 'ground squirrel' ~ R. nànkpyạ̀t-ə̀r 'l'écureuil'

Between h and a consonant, *ȩ has become i̧ạ:

(16) a. K. hȩ̀k-ṷ́ 'middle' ~ R. hì̧ạk-ú 'le milieu'
     b. K. hȩ́r 'cut (belly) open!' ~ R. hì̧ạr-ṷ 'ouvrir le ventre (opérer)'
     c. K. hȩ́tə́ 'make an incision!' ~ R. hì̧ạt-ṷ 'faire des incisions'

*ȩ has been raised to i̧ elsewhere, i.e. when long (17a,b), when final in a noun root (17c,d,e), and when final in a low-toned verb root (17f,g):

(17) a. K. cèel-ụ̀ 'to turn over' ~ R. k̀ị̀l-ụ̀ '(se) tourner'
 b. K. wéẹr-ụ́ 'placenta' ~ R. vír-ụ́ 'le placenta'[7]
 c. K. àkpê 'mouse' ~ R. àkpî 'toute petite souris...'
 d. K. hé-ẹr 'yam' ~ R. hí-ír 'l'igname'
 e. K. sè-ẹr´ 'field mouse' ~ R. sì-ìr 'la souris (sauvage)'
 f. K. sé-eụ 'to greet' ~ R. sí-ụ́ 'saluer'
 g. K. hè 'pick (fruit by hand)!' ~ R. hí-ụ́ 'cueillir'

This development neutralized the contrast between Common Lama *ẹ and *i in these environments.
 Note that it is the underlying tone of the verb root that conditions diphthongization. A verb will have a single reflex of *ẹ throughout its conjugation, in spite of the fact that tones are neutralized in all forms but the imperative:

(18) *ẹ̀   *ẹ́
   hí-ụ́   lyáụ́   tyá-ụ́   infinitive
   hì   lyá   imperative
   hî   lyâ   progressive
   hì-ạ́   tyà-ạ́   perfective
   (17g)   (14a)   (14c)

 The advanced mid vowel *o has diphthongized to yọ after coronals (19) and to ịọ after kp (20):

(19) a. K. só-ou 'to lift up' ~ R. ʃyọ́-ụ́ 'basculer'
 b. K. àló-or 'transvestite' ~ R. ạ́lyọ́r 'personne de sexe masculin qui adopte
                                    les comportements de femmes'
 c. K. tòotǫ̀ 'estimate the weight!' ~ R. tyọ̀t-ụ̀ 'soupeser'

(20) a. K. kpòsǫ̀ 'bark!' ~ R. kpị̀ọ̀s-ụ̀ 'aboyer'
 b. K. kpóntá 'leopard' ~ R. kpị́ọ́ntạ́ 'le lion'
 c. R. kpị̀ọ̀l-ụ̀ 'assembler'

Example (19c) shows that diphthongization affected long *oo as well as short *o. After h, k, and p, *o has become wịụ root-finally (21) and wị non-root-finally (22):

(21) a. K. hó-ou 'to become liquid' ~ R. hwị́ụ́-ụ́ 'être trop liquide'
 b. K. kó-ou 'to cry' ~ R. kwị́ụ́-ụ́ 'crier'
 c. R. hwị́ụ̀-r 'le bourgeon'

(22) a. K. kóm-ụ̀ 'kapok tree' ~ R. kwị́m-ụ̀ 'le kapokier'
 b. K. hôm 'to wake up' ~ R. hwị̂m 'se réveiller'
 c. K. ạpor-ạ 'dance...' ~ R. ạ̀pwị̀r-à 'chants et danses...'

Between *w* and a consonant, *o* has become *i̧*:

(23) a. K. wól 'daughter-in-law' ~ R. wi̧l 'la bru'
     b. K. wòs-ú 'to wake someone up' ~ R. wi̧s-ū̧ 'réveiller'
     c. K. wònk-ò́ 'donkey' ~ R. wi̧nk-ò́ 'l'âne'

The retracted mid back vowel *o̧ has diphthongized only when short and only after *h* and *k*, to *wo̧ root-finally (24) and to *wa̧ non-root-finally (25):

(24) a. K. kò̧ 'sister' ~ R. kwò̧ 'personne de sexe feminin plus âgée'
     b. K. kò̧ 'accept!' ~ R. kwó̧-ú 'être content, satisfait'
     c. K. hô̧-o̧r 'heart' ~ R. hwo̧-o̧r 'le coeur (affectif)'

(25) a. K. à̧kò̧m 'visitor' ~ R. à̧kwa̧m 'l'étranger'
     b. K. hô̧m 'to pull' ~ R. hwa̧m 'tirer (vers soi)'
     c. K. kò̧tȩ̀ 'knock!' ~ R. kwa̧t-ū̧ 'frapper...'

After *w*, we find similar reflexes, but without a second *w*, namely *o̧ root-finally (26) and *a̧ non-root-finally (27):

(26) a. K. wó̧-o̧u 'to spread (a mat)' ~ R. wó̧-ú 'étaler (une natte)'
     b. K. à̧wô̧-r 'place' ~ R. à̧wô̧-r 'la place'

(27) a. K. wó̧k-ú 'fool' ~ R. vá̧k-ú 'l'idiot'
     b. K. wó̧p-ȩ̀ 'to roast' ~ R. wá̧p-ȩ̀ 'faire cuire au feu de braises'
     c. K. wo̧s-u̧ 'to boil' ~ R. và̧s-ū̧ 'bouillir'

Common Lama *o̧ is retained elsewhere, i.e. when long (28a,b) or when preceded by a consonant other than *h*, *k*, or *w* (28c,d,e):

(28) a. K. wò̧o̧r 'praise!' ~ R. wò̧ò̧r-ū̧ 'féliciter'
     b. K. kò̧o̧s-ù̧́ 'flute' ~ R. kò̧s-ù̧́ 'instrument à vent...'
     c. K. yó̧r-ô̧ 'Indian millet' ~ R. yó̧r-ô̧ 'variété de fonio...'
     d. K. mò̧s-ȩ̂m 'thought' ~ R. mò̧s-ȩ̂m 'le fait de penser'
     e. K. à̧cà̧mkpô̧-o̧r 'bachelor' ~ R. à̧cȩ̀nkpó̧-o̧r 'le célibataire'

4. Generalizations

While the Refal developments may seem bewildering at first, we can make a number of generalizations. First, Common Lama front and back mid vowels tended to diphthongize in Refal. All front mid vowels changed, usually to diphthongs. All advanced mid vowels changed to diphthongs (or triphthongs). And some retracted back mid vowels changed to diphthongs.

Second, all Refal reflexes of Common Lama mid vowels are retracted. Note, however, that no two mid vowels were neutralized in any environment, although *e/*i̧ and *o/*a̧ were neutralized in certain environments. Reflexes of advanced mid vowels tend to be higher than reflexes of retracted mid vowels: ya̧, wi̧ vs. ya̧,

w*ạ*. We see, thus, an interaction between the dimensions of tongue height and tongue root advancement: earlier non-high vowels are realized as retracted vowels, and earlier retracted vowels are realized as lower (often low) vowels.

Third, triphthongs developed from advanced mid vowels after certain consonants; **e* became †*yị̈* after *y*, **o* became *wị̈u* after *h, k, p, w*. (I use a dagger to mark unattested forms intermediate between Common Lama and contemporary Refal.) Fourth, the vocalic portions of diphthongs (and triphthongs) tended to be bleached (and sometimes lowered) preconsonantally: *yị* > *yạ*, †*yị̈* > †*yị̈*, *wị̈u* > *wị̈*, *wọ* > *wạ*. (And perhaps also **e* > †*yẹ* > *yạ*.) Fifth, glides within diphthongs were absorbed into homorganic glides in onsets: †*yị̈* > *ị̈*, †*yị* > *ị*, *wị* > *ị*, *wọ* > *ọ*, *wạ* > *ạ*.

Sixth, all conditioning factors were internal to the root. The only exceptions to this generalization involve vowel-final roots that are invariably followed by the same suffix. While most nouns have both a singular and a plural marked by different noun class suffixes, some nouns have only one form. In most such cases, the initial consonant of the class suffix suffices to condition the preconsonantal (i.e. non-final) reflex of a Common Lama mid vowel:[8]

(29) a. K. lê-m 'water' ~ R. lyậ-m 'l'eau'
     b. K. lệ-n 'intelligence' ~ R. lyậ-n 'l'intelligence'
     c. R. kwậ-tậ 'le plafond'

5. Diphthongization Cross-Linguistically

Donegan (1978) makes the following generalizations about diphthongization processes. First, diphthongization applies preferentially to long vowels as opposed to short vowels.[9] Second, diphthongization yields falling diphthongs (that is, diphthongs in which the syllabic element precedes the non-syllabic element), even from short vowels. Third, falling diphthongs (which are typically heavy, like long vowels) may become rising diphthongs (which are typically light, like short vowels) due to timing requirements (e.g. in closed syllables). Fourth, diphthongization is typically context-free.

However, the diphthongizations that have taken place in Refal contradict every one of these generalizations. In Refal, diphthongization applies preferentially to short vowels. Advanced mid vowels diphthongized regardless of length, but retracted mid vowels diphthongized only if short. Long retracted mid vowels were raised (17a,b) or retained (28a,b).

In every attested case in Refal, diphthongization has yielded rising diphthongs (that is, diphthongs in which the non-syllabic element precedes the syllabic element), even from long vowels. Donegan suggests that apparent cases of diphthongization yielding rising diphthongs involve an intermediate step of falling diphthongs, which subsequently undergo a shift of syllabicity. But considering the multiplicity of diphthongs in Refal, the absence of clear-cut cases of falling diphthongs casts some doubt on such a suggestion. Moreover, note that rising diphthongs result regardless of timing in Refal: we find (light) rising diphthongs even in open syllables.

Finally, diphthongization is context-sensitive, at least for retracted vowels. *ẹ diphthongized only if root-final in a high-toned verb or if non-root-final; elsewhere it was raised to į. *ọ diphthongized only if preceded by h, k, or w; elsewhere it was retained.

The Refal developments so consistently run counter to Donegan's generalizations that they raise the question: Are there two types of diphthongization, one deriving falling diphthongs from long vowels, and another deriving rising diphthongs from short vowels?

6. Alternative Interpretations

Two alternative interpretations of the synchronic Refal data immediately suggest themselves. What I have treated as a sequence of an onset consonant plus a diphthong (30a) could be analyzed as an onset cluster plus a vowel (30b) or as a palatalized or labialized consonant plus a vowel (30c):

(30) a. $[k]_O [wọ]_R$ $[p]_O [ya]_R$
      b. $[kw]_O [ọ]_R$ $[py]_O [a]_R$
      c. $[k^w]_O [ọ]_R$ $[p^y]_O [a]_R$

I will argue briefly against the interpretations in (30b) and (30c).[10]

Aritiba (1987) treats the sequences in question as containing onset clusters, as in (30b). However, if they are clusters, then they are the only tautosyllabic clusters in the language. Clusters of obstruent plus glide are the most highly preferred type of onset cluster cross-linguistically (Vennemann 1988), so their existence as the only clusters might not disturb us. However, sequences of liquid plus glide make very poor onsets, and these are amply attested in Refal, as in (29a,b) above. Note also that Aritiba does not discuss Kantè, and thus he may not have taken into consideration the vowel correspondences that suggest treating the sequences in question as simple onsets plus diphthongs.

A more likely interpretation might be that the sequences in question involve consonants with a secondary palatal or labial articulation, as in (30c). The occurrence of the putative w primarily after h and k—two eminently labializable consonants—could be seen as supporting such an analysis. Note, however, that Aritiba considers and rejects this possibility, always treating the glide as a separate segment. He indicates palatalization of *s by using the symbol ∫ in addition to writing a following y, which suggests that y is not simply a diacritic for palatalization. A palatalization analysis would also entail the existence of a four-way contrast $t/t^y/c/c^y$:

(31) a. K. tệm-ụ̂ 'gourd plant' ~ R. tệm-ụ̂ 'plante rampante...''
      d. K. tèm 'finish!' ~ R. tyệm 'finir'
      c. K. cệpậ 'mighty person' ~ R. cệpậ 'le puissant'
      d. K. cém-ụ̀´ 'hen' ~ R. cyệm-ụ̀´ 'la poule'

These putative contrasts—while not inconceivable—are at least a bit suspect.

Moreover, palatalization would have taken place in an odd array of contexts. Consider the forms in (32):

(32)	CL	Kantè		Refal	
a.	*sì-	sí-iu	'to wear'	sí-ú	'porter un vêtement'
b.	*sì̧-	sí̧-i̧u̧	'to enter'	sí̧-ú̧	'entrer'
c.	*sé-	sé-eu	'to run'	ʃyí-ú	'courir'
d.	*sȩ̀-	sȩ́-ȩu̧	'to greet'	sí̧-ú̧	'saluer'
e.	*sȩ́-	sȩ́-ȩu̧	'to plant'	ʃyá̧-ú̧	'planter (arbre)'
f.	*sà̧-	sá̧-a̧u̧	'to cook'	sá̧-ú̧	'faire cuire'
g.	*só-	só-ou	'to lift up'	ʃyó̧-ú	'basculer'

The palatalization of a consonant could not be predicted from the quality of the following vowel in contemporary Refal, since the putative palatalized and non-palatalized consonants contrast before *i* (32b,c,d) and before *a̧* (32e,f). Looking at the Common Lama vowels, palatalization would be predictable, but not according to the usual pattern. Palatalization would never have been triggered by following high front vowels (32a,b), typical triggers of palatalization. On the other hand, palatalization would have been triggered always by **e* (32c), and sometimes by **ȩ* (32e, but not 32d).[11] And the palatalization of a consonant before a mid back rounded vowel (32g) is baffling by any account.

Finally, note that there is one case of diphthongization in word-initial position. In Common Lama, words could not begin with non-low vowels. But in one word, an initial **w* has been lost before a diphthong (33a):

(33) a. K. wȩ̀dȩ̀ 'throw away!' ~ R. yà̧d-ũ̧ 'jeter'
      b. K. wȩ́-ȩu̧ 'to give way' ~ R. vyá̧-ú̧ 'céder la place à'

The sequence of *v* plus diphthong in (33b) apparently represents the regular development. Thus, we cannot say that the outcome of an attempt to palatalize *w* is *y*. Rather, we must say that, in one word (33a), the segment *w/v* has been deleted before the following segment *y*.

Thus, both the onset cluster and the secondary articulation interpretations run into problems. The diphthong interpretation appears preferable on several counts. First, it most simply reflects the correspondences between the two Lama dialects: diphthongs in Refal correspond to monophthongs in Kantè. Second, it avoids positing universally disfavored onset clusters such as *ly* and *ry*. Third, it reflects the fact that *y* may follow any consonant, but may precede a limited set of vowels; this is what is expected if *y* forms a constituent with the following vowel and not with the preceding consonant. Fourth, it conforms with Aritiba's judgment that postconsonantal glides are separate segments, not secondary articulations on the preceding consonant. Fifth, it avoids positing palatalization in an ill-motivated set of contexts.

The mismatch between diphthongization in Refal and Donegan's generalizations calls for further investigation of rising diphthongs in other languages, including those in other West African languages. If rising diphthongs in

other languages turn out to be derived unambiguously from short vowels, then diphthongization will have to be recognized as comprising two complementary types: one applying preferentially to long vowels and yielding falling diphthongs, and the other applying preferentially to short vowels and yielding rising diphthongs.

Notes

1. My study of Lama began in a field methods class at the University of Illinois in 1987. I am grateful to the participants in that class, and especially to Méterwa Ourso, who served as consultant. I also thank Larry Hyman, David Odden, John Ohala, Nikẹ Ọla, Edwin Pulleyblank, Nathalie Schapansky, Pat Shaw, and others at BLS and UBC for their comments.
2. To facilitate comparison, I have modified the transcriptions of Ourso and Aritiba. Following Ourso and Ulrich (1990), I consistently represent tongue root retraction with an underdot, regardless of whether it is contrastive on that vowel in that dialect. Following Ourso (1989a), I represent the palatal nasal as *ñ* and the retroflex stop—derived from underlying *r* (Ourso and Ulrich 1990)—as *d*. In all other cases, I have retained the original symbols. The following is a chart of the orthographic symbols used in the three works.

Aritiba (1987)	i	ɪ	—	—	—	a	o	ɔ	u	ɷ	—	ə	ɨ	ɪ	ʈ	ɲ	v̀	v̀´
Ourso (1989a)	i	i	e	e	a	a	o	o	u	u	ə	3	—	—	d	ñ	v	v̀
Ulrich	i	i̩	e	ẹ	a	ạ	o	ọ	u	ụ	ə	ə̣	ɨ	ɨ̣	d	ñ	v̀	v̀´

While Ourso (1989a) leaves low tones unmarked (except when followed by a floating high tone, which is then unmarked), I have attempted to indicate tone explicitly on every syllable. Where I omit tones, it is because they are absent or illegible in my sources. I have otherwise followed the practice of the original author, marking tone on the first vowel of each syllable for Kantè (after Ourso 1989a) and on every vowel for Refal (after Aritiba 1987).
3. Palatalization of *k* before high front vowels had apparently already taken place in Common Lama (Ulrich 1993).
4. Like the glide + vowel sequences in (5), the tautomorphemic vowel + vowel sequences in (6) correspond to single vowels in Kantè. Moreover, the two vowels always bear the same tone. I therefore treat both types of sequence as diphthongs. Note that the distinction between *yạ* and *iạ* poses problems for the now-customary treatment of glides simply as non-syllabic vowels. Edwin Pulleyblank (personal communication) informs me that a similar distinction is found in Vietnamese.
5. The diphthong *yɨ* is attested in only six roots, none of which have attested Kantè cognates. It may turn out to be a reflex of **e*, but there is insufficient evidence to say.

6. Diphthongs are usually short, even when derived from long vowels. However, there are a few examples of long rising diphthongs:

(i) a. R. tyọ̀ọ̀l-ụ̂ 'tenir un vase dans la paume'
     b. K. cá̧r 'trade!' ~ R. cyà̧à̧r-ụ̂ 'faire du commerce'

Unfortunately, all such forms either lack Kantè cognates, like (ia), or exhibit irregular correspondences, like (ib).
7. The lack of diphthongization in Refal indicates that the long vowel in Kantè is conservative, though the vowel has apparently been shortened in Refal. Cf. also (28b) below.
8. In a few cases, the root-final reflex is found:

(i) a. K. ré-en 'strength' ~ R. ryî-n 'la force'
     b. R. ʃyî-n 'la plaie'

9. Hayes (1990) promotes this generalization to an absolute principle, and bases a theory of phonological representations upon it.
10. For further discussion, see Ulrich (to appear).
11. Bhat (1978) claims that there are languages in which palatalization is triggered by mid front vowels but not high front vowels—but only when the consonant to be palatalized is velar. In Lama, the original place of articulation of the putatively palatalized consonant is immaterial. Ohala (to appear), on the other hand, argues that palatalization in some languages is triggered by mid front vowels but not high front vowels regardless of the original place of articulation of the consonant. Note, however, that mid front vowels would not always trigger palatalization in Refal (32d), so the picture would still not be neat.

References

Aritiba, Adji S. 1987. *Le lamba de Défalé (langue gurunsi du Togo), phonologie et morphologie*. Thèse de doctorat de 3ème cycle, université de Grenoble III.
Bhat, D. N. S. 1978. A general study of palatalization. *Universals of human language*, ed. by Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik, 47–92. Stanford, CA: Stanford University Press.
Donegan, Patricia J. 1978. *On the natural phonology of vowels*. Columbus, OH: Ohio State University Ph.D. dissertation.
Downing, Laura J. 1986. Lamba field methods report. Ms.
Hayes, Bruce. 1990. Diphthongisation and coindexing. *Phonology* 7:1.31–71.
Kenstowicz, Michael. 1989. Compensatory lengthening: from West Africa to East Slavic. Paper presented at the University of Connecticut.
Kenstowicz, Michael, Emmanuel Nikiema, and Méterwa Ourso. 1988. Tonal polarity in two Gur languages. *Studies in the Linguistic Sciences* 18:1.77–103.

Ohala, John J. To appear. Hierarchies of environments for sound variation; plus implications for 'neutral' vowels in vowel harmony. *Acta Linguistica Hafniensia* 27.403–414.

Ourso, Méterwa A. 1988a. Nominal morphology in Lama: further evidence for Lexical Phonology. Ms.

Ourso, Méterwa A. 1988b. Root control, underspecification, and ATR harmony. *Studies in the Linguistic Sciences* 18:2.111–127.

Ourso, Méterwa A. 1989a. *Lama phonology and morphology*. Urbana, IL: University of Illinois Ph.D. dissertation.

Ourso, Méterwa A. 1989b. Phonological processes in the noun class system of Lama. *Studies in African Linguistics* 20:2.151–177.

Ourso, Méterwa A., and Charles H. Ulrich. 1990. Sonorant-strengthening in Lama. *Studies in the Linguistic Sciences* 20:1.135–147.

Ulrich, Charles H. 1993. A preliminary comparison of two Lama dialects. Ms.

Ulrich, Charles H. To appear. Diphthongization of short vowels in Lama and elsewhere. *Proceedings of the 1994 annual meeting of the Canadian Linguistic Association, Toronto Working Papers in Linguistics.*

Ulrich, Charles H. In preparation. Labial dissimilation in Lama. Ms.

Vennemann, Theo. 1988. *Preference laws for syllable structure*. Berlin: Mouton de Gruyter.

Yu, Ella O. 1991. *Theoretical aspects of noun classes in Lama*. Urbana, IL: University of Illinois Ph.D. dissertation.

BERKELEY LINGUISTICS SOCIETY

2337 Dwinelle Hall, University of California, Berkeley, CA 94720 USA
Telephone: (510) 642-5808, (510) 642-2757; e-mail: bls@garnet.berkeley.edu

The Berkeley Linguistics Society is a graduate student organization at the University of California, Berkeley. The Society holds its annual meeting during Presidents' Day weekend in the middle of February. The meeting consists of three parts: a General Session on topics of varied linguistic interest, a Parasession on a particular linguistic topic, and, beginning in 1990, a Special Session on general linguistic topics of a certain geographic area. Around the middle of September, a Call-for-Papers is sent to those on our mailing list, and abstracts of papers are due around the middle of November. After reviewing and selecting the abstracts, the Society sends out the conference program in early January. The Proceedings of the conference are published yearly by the Society and become available in the fall. Order forms for proceedings are mailed with the Call-for-Papers and can be obtained from the above address. All orders require prepayment.

Book Prices*

	Parasession & Special Session (S) Topics	Individuals	Institutions
BLS-1 (1975)	....	$10.00	$15.00
BLS-2 (1976)	....	$10.00	$15.00
BLS-3 (1977)	....	$10.00	$15.00
BLS-4 (1978)	....	$10.00	$15.00
BLS-5 (1979)	....	$10.00	$15.00
BLS-6 (1980)	....	$10.00	$15.00
BLS-7 (1981)	....	$10.00	$15.00
BLS-8 (1982)	....	$14.00	$20.00
BLS-9 (1983)	....	$10.00	$15.00
BLS-10 (1984)	Subordination..	$14.00	$20.00
BLS-11 (1985)	Poetics, Metrics, and Prosody........................	$14.00	$20.00
BLS-12 (1986)	Semantic Typology.......................................	$14.00	$20.00
BLS-13 (1987)	Grammar and Cognition................................	$14.00	$20.00
BLS-14 (1988)	Grammaticalization.......................................	$14.00	$20.00
BLS-15 (1989)	Theoretical Issues in Language Reconstruction..	$14.00	$20.00
BLS-16 (1990)	Legacy of Grice..	$16.00	$20.00
BLS-16S (1990)	General Topics in American Indian Linguistics..	$8.00	$12.00
BLS-17 (1991)	The Grammar of Event Structure....................	$16.00	$20.00
BLS-17S (1991)	African Language Structures...........................	$8.00	$12.00
BLS-18 (1992)	Place of Morphology in a Grammar................	$16.00	$20.00
BLS-18S (1992)	Typology of Tone Languages........................	$8.00	$12.00
BLS-19 (1993)	Semantic Typology and Universals.................	$16.00	$20.00
BLS-19S (1993)	Syntactic Issues in Native American Languages.	$8.00	$12.00
BLS-20 (1994)	Anniversary Session.....................................	$16.00	$20.00
BLS-20S (1994)	Historical Issues in African Linguistics............	$8.00	$12.00
Author Index to BLS 1-16S...		$2.00 (includes shipping)	

California Sales Tax

8.25% for residents of Alameda, Contra Costa, Los Angeles, San Diego, San Francisco, San Mateo, Santa Clara, and Santa Cruz counties

7.75% for residents of Fresno, Imperial, Inyo, Madera, Monterey, Orange, Riverside, Sacramento, San Benito, San Bernardino, San Joaquin, and Santa Barbara counties

7.50% for residents of Sonoma county

7.25% for all other California residents

Shipping and Handling

	Within USA	Outside USA
BLS 1,2,6,7,9,11,12,13,14,15,16S,17S...................................	$2.00	$3.00
BLS 3,4,5,8,10,16,17..	$2.50	$4.00

ORDERS MUST BE ACCOMPANIED BY PAYMENT
Make checks and money orders payable to the Berkeley Linguistics Society
*All prices subject to change without notice. For current prices see most recent BLS volume.