PROCEEDINGS OF THE THIRTY-FIRST ANNUAL MEETING OF THE

BERKELEY LINGUISTICS SOCIETY

February 18-20, 2005

SPECIAL SESSION

on

LANGUAGES OF WEST AFRICA

Edited by

Rebecca T. Cover and Yuni Kim

Berkeley Linguistics Society
Berkeley, CA, USA

TABLE OF CONTENTS

SPECIAL SESSION

A note regarding the contents of this volume

The following papers were also presented at the conference, though they do not appear in this volume:

Steven Moran, "Endo- and exocentric compounding in Western Sisaala"
Paul Newman, "New findings about Klingenheben's Law: It's more nuanced and more interesting than we thought"
Ekkehard Wolff, "Encoding topography and direction in the verbal systems of Central Chadic Lamang and Hdi"

Foreword

We are pleased to present the proceedings of the BLS 31 Special Session on Languages of West Africa, held at the University of California, Berkeley in February 2005. We would like to thank the contributors to this volume and all those who attended and participated in BLS 31, as well as the students who volunteered their time before and during the conference. We also acknowledge the generous support of the Center for African Studies at UC Berkeley, which was critical to the success of the Special Session. Special thanks are due to the UC Berkeley Graduate Assembly for their financial assistance in the publication of this volume.

Rebecca T. Cover and Yuni Kim

Volume editors

Interpreting Yorùbá Bare Nouns as Generic[*]

ỌLÁDIÍPỌ̀ AJÍBÓYÈ
University of British Columbia

0.   Introduction

There are two relevant facts about Yorùbá nouns that I discuss in this paper. First, they lack obligatory determiners (contra Bámgbóṣé 1967, 2001, Awóbùlúyì 1978)[1], and so they can be bare. Second, these bare nouns (BNs)[2] can be interpreted in at least one of these three ways: generic, indefinite, or definite.

(1)   a.   Mo   fẹ́ràn   **ajá**.   *generic*
              1sg.   like   dog
              'I like dogs.'

       b.   Mo   rí   **ajá**.   *indefinite*
              1sg.   see   dog
              'I saw a dog.'

---

[1] However, Manfredi (1992) claims that Yorùbá morphemes usually glossed as determiners, and (mis)translated as English definite articles, are both syntactically and semantically more like demonstrative adverbs or deictic nouns.

(i)   a.   ìwée   Ayọ̀   (náà)
              book   of   at that very

       b.   ìwée   Ayọ̀   (yẹn)
              book   of   that.one

       c.   ìwée   Ayọ̀   (yǐí)
              book   of   this.one (Manfredi 1992: 207)

In Ajíbóyè (2005), I give a full account of *náà* and *kan*, and I demonstrate that these elements are neither determiners nor adverbs.

[2] The bare nouns discussed in this paper focus mainly on count nouns partly because mass nouns are not clearly distinguished from count nouns.

    c.   **Ajá**   gbó   mi.   *definite* (in discourse context)
          dog   bark   me
          'The dog barked at me.'

This paper is concerned with the first interpretation. There are two broad generalizations about the availability of a generic construal of bare nouns in Yorùbá. First, a bare noun cannot be construed as generic when it occurs as the subject or object of an eventive verb or as the subject of a stative verb.

(2)   a.   Ajá   rí   mi.   *subject of eventive*
          dog   see   me
          ≠ 'Dogs see me.'

    b.   Mo   rí   ajá.   *object of eventive*
          1sg. see   dog
          ≠ 'I see dogs.'

    c.   Ajá   fẹ́ràn mi.   *subject of stative*
          dog   like   me
          ≠ 'Dogs like me.'

These are the same environments which obligatorily require the presence of the imperfective *máa-ń* to mark genericity. The second generalization is that a bare noun can be construed as generic when it occurs as the object of a stative verb (3a) or the subject of a stative verb, but only if the object is also a bare noun (3b).

(3)   a.   Mo fẹ́ràn ajá.   *object of stative*
          1sg. like   dog
          'I like dogs.'

    b.   Ajá fẹ́ràn egungun.   *subject and object of stative*
          dog   like   bone
          'Dogs like bones.'

These are the same environments which optionally permit the presence of imperfective *máa-ń* to mark genericity.

    I demonstrate that generic construal of bare nouns in Yorùbá can be obtained in one of two ways: lexical conditioning (cf. Carslon and Pelletier 1995, Chierchia 1995) or grammatical conditioning (Déchaine 1993, Delfitto 2002, Longobardi 2000). I propose an operator-based analysis to account for the genericity of these bare nouns (cf. Chierchia 1995).

1. Lexically Conditioned Genericity

Lexically conditioned genericity is when the interpretation of a bare noun as generic is dependent on the predicate type. One characteristic feature of a lexically conditioned generic reading is that it does not require any special marking. As the examples in (3) and (4) show, genericity is encoded in stative verbs in Yorùbá.

(4) a. Bọ́sẹ̀   kórira ológbò.
         PN   hate   cat
         'Bose hates cats.'

     b. Jẹ́nrọ́lá   bẹ̀rù   ejò.
         PN   afraid snake
         'Jenrola is afraid of snakes.'

The facts above are schematized in the structures in (5).

(5) a.
```
        _____
    DP          _____
            V            BN
         stative      ✓generic
```

     b.
```
        _____
    BN          _____
  *generic   V          (DP)
          stative
```

     c.
```
        _____
    BN          _____
  ✓generic   V           BN
          stative     ✓generic
```

Several questions arise at this point:

- Why is there a split between stative and eventive verbs?
- Why are BN objects of statives obligatorily construed as generic?
- Why do subject BNs need object BNs that are generic before they can be generic?

To answer the first question, following Chierchia (1995:219), I propose that stative verbs are inherently generic. However, as observed in Chierchia (1995: 202), the idea that certain verbs "are somehow inherently generic cannot be straightforwardly implemented in strict lexicalist terms." According to Chierchia, this is so because such verbs "cannot stand on their own and need to be operated

on by GEN" (1995:202). In light of this, I propose that stative verbs in Yorùbá require the presence of a generic operator (GEN) that binds any BN in its scope. I assume that this GEN is generated first in VP-adjoined position (Chierchia 1995: 213). It is this GEN that licenses bare nouns as generic, subject to certain other conditions that are discussed below. In pursuance of this proposal, I assume a υP-shell structure (6).
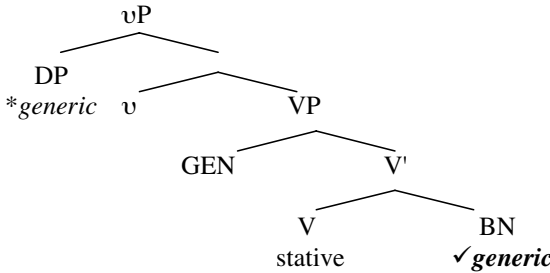
(6) statives: transitives = υP
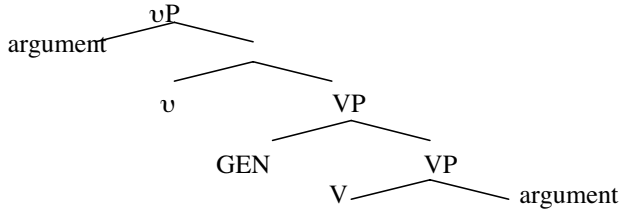
(cf. Chierchia 1995: 213)

Thus, GEN is located in the lower part of the υP-shell, i.e. in Spec,VP. The fact that it is Spec,VP that hosts GEN explains why there is subject-object asymmetry in terms of how subject and object bare nouns can be interpreted as generic.

(7) a.

b. Mo fẹ́ràn **ajá**.
'I like dogs.'

As shown in §1.1, despite the fact that both subject and object nouns are bare, the subject bare noun continues to be non-generic even with the presence of GEN. My claim as to why this is so is that the subject bare noun is not in a position where GEN can scope over it.

1.1. Consequence 1: Subject-Object Asymmetry

This section accounts for why at LF (before covert movement of the generic operator takes place to a position where it can scope over the subject), even when both subject and object nouns are bare, only the object can be construed as

generic.

(8)   a.   υP

          BN
          generic   υ   VP

                         GEN   V'
                                      V   BN
                                             ✓*generic*
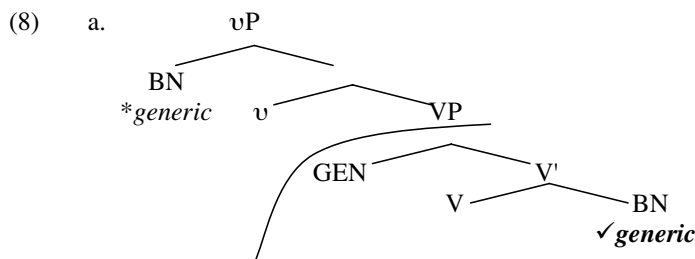
     Given the structure in (8), it is assumed that GEN scopes over only the object bare nouns, since it is introduced within the VP and not outside. This then explains why only the object bare noun can be construed as generic in (7). I give a summary of my discussion up till this point in (9).

(9)   Absence of scope over subject of statives
          a.   If bare N is OBJ, then GEN
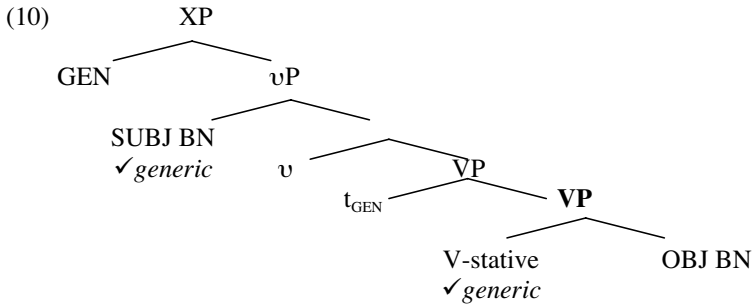   b.   If bare N is SUBJ, then *GEN

     There still remains the issue of why the interpretation of subject bare nouns as generic has to rely on the type of nominal expression that occupies the object position.

1.2.   Consequence 2: Object Dependency with Statives in Relation to Scope

What we have seen so far seems to suggest that the interpretations of bare noun subjects versus objects as being generic are independent of each other. However, we must also account for cases where both subject and object bare nouns are construed as generic, and particularly for the fact that the interpretation of bare noun subjects as generic is dependent on the interpretation of bare noun objects as generic. I propose that an account of this dependency can be carried out when we understand the scope relation between the generic operator and these bare nouns.

     First I define the scope of an operator, along the lines of Szabolcsi (2000:607), as the domain within which it has the ability to affect the interpretation of other expressions.

     I propose that the abstract lexical GEN operator undergoes Q-raising from Spec,VP to a position where it can scope over both the subject and the object BN, causing both bare nouns to be within the domain of the operator (cf. Diesing 1992, Szabolcsi 2000).

(10)

```
                    XP
          ┌──────────┴──────┐
        GEN               υP
                    ┌───────┴──────┐
                SUBJ BN            │
                ✓generic    υ         VP
                           ┌────────┴──────┐
                        t_GEN              VP
                                    ┌───────┴──────┐
                                V-stative       OBJ BN
                                ✓generic
```

However, movement of GEN must satisfy one condition: the object must be bare. When the object is not bare, there is no genericity of subject bare nouns.

With this I conclude the discussion of genericity that is lexically conditioned and turn to grammatically conditioned genericity.

2. Grammatically Conditioned Genericity

This section looks at the occurrence of bare nouns in eventive predicates. The first observation is that when bare nouns appear as subjects and/or objects of eventive predicates, they cannot be construed as generic. This I show in (11) and (12).

(11) a. **Ajá**   rí   mi.   *subject of transitive eventive*
          dog   see   me
          ≠'Dogs see me.'

      b. **Ajá**   hàn.   *subject of intransitive eventive*
          dog   snore
          ≠'Dogs snore.'

(12) a. Mo   rí   **ajá**.   *object of transitive eventive*
          1sg. see   dog
          ≠ 'I see dogs'

      b. **Ajá**   jẹ   **egungun**.   *subject & object of transitive eventive*
          dog   eat bone
          ≠ 'Dogs eat bones.'

In order to obtain a generic reading, the imperfective marker *máa-ń*[3] must be introduced.

---

[3] Barczak (2004) treats *máa-ń* as grammatical particles that jointly mark a sentence as habitual. In this analysis, I treat the two as imperfective.

(13) a.   Ajá **máa-n** rí   mi.
           dog **IMP**   see me
           'Dogs see me.'

      b.   Ajá **máa-n** hàn.
           dog **IMP**   snore
           'Dogs snore.'

(14) a.   Mo   **máa-n** rí   ajá.
           1sg. **IMP**   see dog
           'I see dogs.'

      b.   Ajá **máa-n** jẹ   egungun.
           dog **IMP**   eat bone
           'Dogs eat bones.'

Based on these data, I argue that a subject or object bare noun of an eventive verb, or a subject bare noun of a stative verb, is generic if imperfective *máa-ń* is present. It follows that grammatically conditioned genericity is that which involves interpreting bare nouns as generic by the presence of an imperfective marker. Consequently, I assume that the imperfective morphemes *máa-ń* carry a quantificational feature [+Q] that forces the presence of GEN in its local environment, along the lines of Chierchia (1995:202).
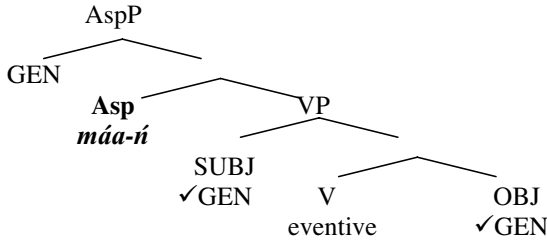
In what follows, I develop an aspectual theory of genericity along the lines of Kamp and Reyle (1993) to address grammatically conditioned genericity.

2.1.   Theory of Aspectually Conditioned Genericity

The theory of aspect and generic interpretation of bare nouns that I propose draws largely from the work of Chierchia (1995) and Kamp and Reyle (1993:569). Kamp and Reyle treat "progressive" and "perfect" as aspectual operators (Asp OP). The claim is that these operators transform the meaning of the underlying non-progressive or non-perfect verb, verb phrase or sentence into that of its progressive or perfect counterpart, respectively.

Adapting their aspect-model analysis, I propose that imperfective *máa-ń* is associated with the presence of the υP-external aspectual marker. This makes available the generic construal of bare nouns in either subject or object position.

(15)

```
                    AspP
              ┌──────┴──────┐
          GEN              
                     ┌───────┴───────┐
                    Asp             VP
                  máa-ń        ┌─────┴─────┐
                            SUBJ          
                           ✓GEN      ┌─────┴─────┐
                                     V          OBJ
                                  eventive      ✓GEN
```

The question that immediately arises is this: how do we motivate the interpretation of bare nouns as generic with the presence of the imperfective? To answer this question, I further propose that this imperfective is associated with a GEN, which is introduced in Spec,AspP (Déchaine 1993, Barczak 2004). What I am assuming is that whenever an eventive verb is accompanied by the IMP, a bare noun either in subject or object position can be construed as generic. Observe that GEN, which is introduced in Spec,AspP, is able to scope over both the subject and object bare nouns. I extend this analysis to the cases involving stative verbs discussed in §1.

2.2. Extending Aspectually Conditioned Genericity to Stative Verbs

Recall that in §1.1, we saw that in lexically conditioned genericity, a bare noun subject could not be construed as generic even with the presence of a null GEN. What I intend to do is to extend the grammatically conditioned approach that accounts for all cases involving eventive predicates to those that involve stative predicates. In doing so, I begin with those cases that cannot be construed as generic at all, i.e. subject bare nouns of statives.

(16) a. **Ajá** fẹ́ràn mi.   *subject of transitive stative*
          dog   like   me
          ≠ 'Dogs like me.'

      b. **Ajá** tóbi.   *subject of intransitive stative*
          dog   be.big
          ≠ 'Dogs are big.'

The prediction is that with the presence of IMP, we should be able to have generic construal of these bare nouns.

(17) a. Ajá **máa-ń**   fẹ́ràn mi.
          dog   **IMP**   like   me
          'Dogs like me.'

b. Ajá   **máa-ń**   tóbi.
    dog   **IMP**   be.big
    'Dogs are big.'

This prediction is borne out, since the bare noun in each of the examples in (17) is now interpretable as generic.

    However, in grammatically conditioned genericity, the presence of the imperfective is optional, as expected. This is shown in (18).

(18) a. Mo   (**máa-ń**)   fẹ́ràn   **ajá**.   *object of stative*
          1sg. **IMP**   like   dog
          'I like dogs.'

     b. **Ajá**   (**máa-n**)   fẹ́ràn   **egungun**   *subject of stative*
          dog   **IMP**   like   bone
          'Dogs like bones.'

I extend the analysis of Yorùbá data to English in the following section.

3. Extending Our Analysis to English

First, I show that bare nouns in English (i.e. nouns without overt determiners), as in Yorùbá, can be construed as generic with stative predicates.

(19) a. Dogs like bones.   *stative*
          GENx GENy [dog(x), bone(y), like(x,y)]

     b. **Ajá**   fẹ́ràn   **egungun**.   *stative*
          dog   like   bone
          'Dogs like bones.'

However, with eventives, Yorùbá contrasts with English in how bare nouns are construed as generic. While English bare nouns can be construed as generic, (20a) Yorùbá bare nouns cannot (20b).

(20) a. Dogs eat bones.   *generic*
          GENx GENy [dog(x), bone(y), like(x,y)]

     b. **Ajá**   jẹ   **egungun**.   *generic*
          dog   eat   bone
          ≠ 'Dogs eat bones.'

So English does not show the kind of distinction that Yorùbá shows with respect to predicate type and generic reading. Thus, generic construal of bare plural noun (subject and object) is possible with both stative and eventive verbs in English but

not in Yorùbá. The generalization that emerges is the following: in English generic construal is always available, whereas in Yorùbá, it is sometimes available. As I argue below, this distinction is only at the surface level.

First, recall that despite what is shown above, Yorùbá has a way of obtaining a generic reading in eventive constructions. This is what I discussed under aspectually conditioned genericity. Note that I also revisited cases involving stative verbs using the same approach. My concern here is how to generalize the resolution of genericity in Yorùbá to English.

I propose that the English example in (20a) be interpreted as imperfective as shown in (21). If this proposal is correct, then Yorùbá and English converge.

(21)   IMPERFECTIVE DOGS eat BONES

The fact that the two languages converge at some abstract level suggests that IMP is the source of GEN in both languages. Extending the analysis of Yorùbá bare nouns to English, I propose that the imperfective always introduces GEN, and that while Yorùbá has an overt imperfective, English has a covert imperfective. Note, however, that in English, based on the examples shown above, there is no overt imperfective marker present. The question that arises is how then do we harmonize the facts of Yorùbá, which shows overt presence of an imperfective marker, and the facts of English, which shows its absence. I seek to address this issue through the use of the default aspect theory.

4.   The Default Aspect Theory

In order to understand how English and Yorùbá converge on the generic construal of bare nouns, I propose a theory of aspect for the two languages. I claim that in Yorùbá, the default aspect/tense is the perfective/past, whereas in English, the default aspect is the imperfective/present.

I give the example in (22) to illustrate the case of Yorùbá. Observe that the sentence can only be interpreted as simple past (22a) or present perfective (22b), and neither as progressive (22c) nor as imperfective (22d).

(22)   Jímọ̀ jẹ iṣu.
        a.   Jimo ate yam   PAST
        b.   Jimo has eaten yam   (Present) PERFECTIVE
        c.   *Jimo is eating yam   *PROG
        d.   *Jimo eats yams   *IMP

I also illustrate the English imperfective/present with the examples in (23). As in Yorùbá, there are only two interpretations to (23): either as habitual and imperfective (23a) or as a historical present[4] (23b).

---

[4] This is in the context of story-telling when a series of successive events is encoded.

(23)   Jimmy eats yam.
     a.   habitual and imperfective
     b.   historical present

    The generalizations that emerge from the two languages are the following: the unmarked verb form in English is the imperfective, whereas the unmarked verb form in Yorùbá is the perfective. With this generalization, I argue that generic construal is closely linked to the imperfective. Generic construal is determined at least in these two languages through the use of the imperfective.

(24)   a.   ENG default as (IMP) is equal to GEN
      b.   YOR overt IMP is equal to GEN

To sum up, it has been established that in Yorùbá the imperfective is overtly realized through the use of *máa-ń,* whereas in English, the imperfective is covertly present.

(25)   a.   Jímò̩ [$_{IMP}$ máa-ń] je̩ is̩u   *Yorùbá genericity*
      b.   John [$_{IMP}$ $\varnothing$] eats yam   *English genericity*

5.   Conclusion

I have shown that a generic construal of bare nouns can be obtained in one of two ways: via lexical conditioning with stative verbs and via grammatical conditioning with eventive verbs through the use of the imperfective *máa-ń*. I have been able to reduce the two analyses to one by claiming that a bare noun can be interpreted as generic using the grammatically conditioned approach. This analysis has also contributed to our knowledge of genericity in Yorùbá and English. With the use of IMP, bare nouns in stative and eventive predicates can be construed as generic. This IMP can be overt, as in the case of Yorùbá, or covert, as in the case of English.

References

Ajíbóyè, O. 2001. The internal structure of Yorùbá DP. Paper presented at the 32nd Annual Conference of African Linguistics, Berkeley, March 22-25, 2001.

--. 2005. Topics on Yorùbá nominal expressions. Ph.D. diss., University of British Columbia.

Baltin, Mark, and Chris Collins. 2000. *Handbook of Contemporary Syntactic Theory*. Malden, MA: Blackwell.

Bámgbós̩é, Ayò̩. 1967. *A Short Yorùbá Grammar*. London: Cambridge University Press.

--. 2000. Coreferentiality and focus in textual cohesion: Evidence from Yorùbá. *Journal of West African Languages* 28(2):59-70.

Barczak, Leszek. 2004. Syntax and semantics of the Yorùbá auxiliary *máa*. 34[th]Annual Conference of African Linguistics, Boston, April 1-4, 2004.

Carlson, Gregory N., and Francis Jeffry Pelletier, eds. 1995. *The Generic Book*. Chicago: University of Chicago Press.

Chierchia, Gennaro. 1995. Individual level predicates as inherent generics. In G.N. Carlson and F.J. Pelletier, eds., 176-223.

Déchaine, Rose-Marie. 1993. Predicates across categories: Towards a category-neutral syntax. Ph.D. diss., University of Massachusetts, Amherst.

Déchaine R.-M., and V. Manfredi. 1998. SVO ergativity and abstract ergativity. *Recherches linguistiques de Vincennes* 27:71-94.

Delfitto, Dennis. 2002. *Genericity in Language: Issues of Syntax, Logical Form and Interpretation*. Alessandria: Edizioni dell'Orso.

Diesing, Molly. 1992. *Indefinites*. Linguistic Inquiry Monograph 20. Cambridge, MA: MIT Press.

Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Languages, Formal Logic and Discourse Representation Theory, Part* 2. Amsterdam: Kluwer Academic.

Longobardi, Giuseppe. 2000. The structure of DPs: Principles, parameters and problems. In M. Baltin and C. Collins, eds., *Handbook of Contemporary Syntactic Theory*, 562-603. Malden, MA: Blackwell.

Manfredi, Victor. 1992. A typology of Yorùbá nominalizations. In Chris Collins and Victor Manfredi, eds., *Proceedings of the Kwa Comparative Syntax Workshop*. Cambridge, MA: MIT Working Papers in Linguistics.

Szabolcsi, Anna. 2000. The syntax of scope. In M. Baltin and C. Collins, eds., *Handbook of Contemporary Syntactic Theory*, 607-634. Malden, MA: Blackwell.

Ọládiípo Jacob Ajíbóyè
Department of Linguistics and Nigerian Languages
University of Ilorin
P.M.B. 1515, Ilorin
Kwara State, Nigeria

oladiipo@yahoo.com

Divergent Structure in Ogonoid Languages

OLIVER BOND and GREGORY D. S. ANDERSON
University of Manchester and *University of Oregon*

0. Location and Demographics

The Ogonoid (Ogoni/Kegboid) language family of southeastern Nigeria (Benue-Congo, Niger-Congo) comprises five languages spoken in a contiguous area across the Niger Delta.[*] Despite shared ethnographic practices, frequent intermarriage and the immediate physical adjacency of these linguistic communities, Ogonoid languages (1) present with strikingly divergent morpho-syntactic structure. All the Ogonoid languages (except Baan) are spoken in an eponymous Local Government Area (LGA) of Rivers State; the political boundaries of each LGA largely reflect the linguistic boundaries. Baan (previously known as Ogoi) is spoken in eastern parts of Eleme LGA and western parts of Tai LGA.

(1)   Population figures for the Ogonoid languages

Language:	Gokana	Kana/Tai	Eleme	Baan
Population:	100,000	200,000	50,000	<5,000

1. Phonology and Lexis of the Ogonoid Languages

To date, classification of the Ogonoid languages as a linguistic family – as is typically the case – has been based on "systematic" correspondences in the phonology and lexicon of the individual varieties (2). Early classifications of the language family treated Tai as a dialect of Kana (Wolff 1964; Williamson 1985; Faraclas 1989, and in the first published comparison of the Ogonoid languages (Wolff 1964), Baan was likewise omitted, and then informally considered a dialect of Eleme.

    The table in (2), based on data from Williamson (1985) and supplemented with additional material from Tai (Nwí-Bàrì 2002), illustrates the differences in the phonological realization of cognate vocabulary across the languages. Generally speaking, Tai and Kana seem to cluster together, and these less so with Gokana. Baan and Eleme seem themselves to group together, appearing to

---

support the initial classification of Ogonoid listed in (3), adapted from Williamson and Blench (2000:33).

(2)   Cognate vocabulary in the Ogonoid languages

	Gokana	**Kana**	**Tai**	**Baan**	**Eleme**
dance	zòb	yèb	yèb	dʒee	dʒè
pound	kṹmĩ́	kṹm	kũm	kũũ	kũ̃
sweep	kpárí	kpáe[1]	kpɛɛ	kpari	kparí
story	lóg	lóó	lóó	yógi	èlóí
child	nẽn	nẽẽ	nẽẽ	nẽẽ	ɔnẽ̃
cooking pot	bã́	bã́	bã́	bã́ã́	àbã́
tree	té	té	té	té	èté
wife	va	wa	wa	wa	ɔwa
salt	ló̩	ló̩	ló̩	ǹdó̩	ǹló̩
rope	mã́	mã́	mã́	mĩ́mã́	m̀mĩ́mã́

(3)   Classification of the Ogonoid family
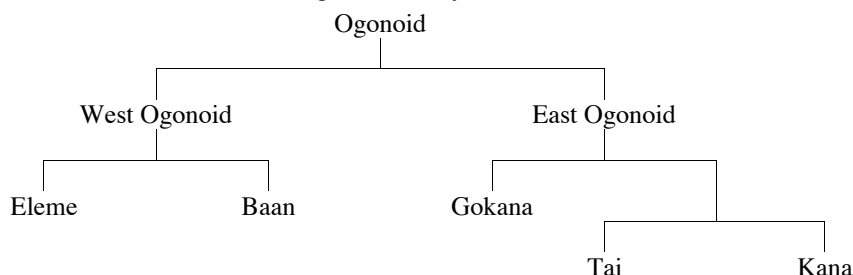
     This classification of Ogonoid is based purely on phonological aspects of lexical data rather than a systematic analysis of the languages crosswise. The comparative picture discussed in brief below is not nearly so clear when structural (morphological, syntactic) or even the ostensibly diagnostic "phonological" data are considered. On closer examination, it is uncertain that an East Ogonoid node exists *per se*, or that Gokana and Kana actually have a special relation to one another, as opposed to simply being "not West Ogonoid." Further, given the paucity of data on Baan, its position within a putative West Ogonoid subgroup with Eleme has yet to be adequately determined.

2.   Divergent Structure in Ogonoid
For the purposes of the following analysis, we limit ourselves to but a handful of the numerous features that characterize the Ogonoid languages both collectively

---

[1] Williamson (1985:434) notes that *[l] is lost in this citation form from Kana, but retained as [r] in the perfect verb form *è-kpár-a* 'he has swept'.

and individually. These include phonological features, morphological features of nouns and (in particular) verbs, and some features of noun phrase syntax. Within the domain of phonology, the types of coda restrictions exhibited among the Ogonoid languages are examined. The degree of retention of archaic noun class prefixes is also explored, as well as a range of verbal features, including systems of subject and object marking, tense and aspect encoding, use of reduplication in finite verb morphology and the presence of grammaticalized verbal negatives. Data used in the analysis comes from the available published literature on Gokana, Kana, Tai, and Baan, combined with the authors' field notes on Eleme.[2]

2.1.   Coda Restrictions

We start our discussion with a phonotactic feature that has figured prominently in the classification of the Ogonoid languages, namely the presence or absence of coda consonants. As it turns out, Eleme (and insofar as this can be determined, Baan as well) indeed differs from Gokana, Kana and Tai in having lost all original final consonants. However, as seen in (5), there is actually considerable variation in the permissible coda consonants of the different "East" Ogonoid languages.

(4)   Coda consonants in Ogonoid (Williamson 1985:431; Nwí-Bàrì 2002)

Gokana: *b, l, g, m, n, ŋ*
Kana: *b, g, m, ŋ*
Tai: *b, m, [ŋ], [s]*[3]

According to Faraclas (1986:40), coda consonants are common in Cross-River languages and indeed are to be reconstructed for proto-Cross-River (which was typified by *-CVC, *-CVV and *-CVVC roots).

(5)   Coda variation in Kana and Gokana (Ikoro 1996:192)

	forest	call	grass	town	reject
Gokana:	kɔl	kól	víl	bɔ̃n	kìn
Kana:	kùè	kúé	ábíé	bũɛ̃	kĩ̀ĩ

Even Kana and Gokana show considerable systematic differences in cognate lexemes despite being treated as an undifferentiated type in comparative Ogonoid studies (5). There is thus a tendency to restrict coda position that grows increasingly stronger from right to left in a continuum across the Ogonoid family. This can be represented by the cline: Gokana > Kana > Tai >> Eleme.

---

[2] Structural data on Baan is almost entirely lacking, and as such, Baan does not figure heavily into the discussion below. Further research on Baan may cause some or all of the statements contained in this paper to be modified or revised accordingly.
[3] The last two are found in exactly one word each in Nwí-Bàrì (2002), which appear to belong to either some sort of loan strata or "affective," "ideophonic," or "expressive" register or style.

2.2.   Noun Class Prefixes

In addition to coda restrictions in Ogonoid languages, another "diagnostic" feature that lumps Gokana and Kana together in opposition to Eleme is the absence vs. presence of noun class prefixes. These prefixes represent archaic features presumably inherited from Proto-Benue-Congo (6). As with coda restrictions, the correspondences among the Ogonoid languages are not nearly as straightforward as has been alleged in the literature. With noun class prefixes, there appears to be more or less the reverse hierarchy, where Eleme is the most archaic and preserves the greatest amount of prefixes while Gokana has lost them (completely or nearly so). Kana, on the other hand, appears to preserve prefixes to a greater degree than was previously realized.

Noun class prefixes in Ogonoid, when present, usually carry a low tone, or else the tone is copied from the initial syllable of the root. Data from Kana shows that this is merely a tendency, not a rigid absolute, for example *á-yo* 'onion' or *á-kɔ̀m* 'malaria' (Ikoro 1996:58). The shape of the prefix is either vowel or syllabic nasal. In Eleme the vowel is either *a*, *E* or *O* (with harmonic variants). Syllabic nasals show assimilation to the place of articulation of the following consonant.

(6)   Noun class prefixes (Faraclas 1986:47, 50; Nwí-Bàrì 2002)

	Eleme	**Baan**	**Kana**	**Gokana**	**Tai**	**Proto-Benue-Congo**
tooth	*à-dáː*	*dáː*	*dáː*	*dáː*	*dáː*	**li-* (sg) **à-* (pl)
tree	*è-té*	*té*	*té*	*té*	*té*	**ki-* (sg) **bi-* (pl)
ashes	*ǹ-tɔ̃́*	*ǹ-tɔ̃́*	*tɔ̃́*	*(-rù)*	*tɔ̃́*	
animal	*ǹ-nã̀*		*nãm*	*nɔ̃m*	*nãm*	**ì-* (sg) *í-* (pl)
goat	*m̀-bó*	*m̀-bó*		*ból*		**ì-* (sg) *í-* (pl)

Kana has preserved prefixes in a significantly greater number of lexemes than previously believed (7). It thus does not "clearly" pattern with Gokana in this innovation, but rather, as with coda consonants, occupies an intermediate position between the two. Tai[4] is closer to Gokana in this respect, and Baan appears to be between Kana and Eleme. Thus, Kana appears to preserve a vocalic noun class marker in numeral classifiers that have been lost in Baan, suggesting the following cline for the retention of noun prefixes in Ogonoid: Eleme >> Baan, Kana, Tai >> Gokana.

---

[4] There appear to be a small number of words that have preserved lexicalized noun class markers encoded by syllabic nasals in Tai, e.g. *ḿgbárà* 'small', *ńdè* 'type of cocoyam with heart-shaped leaves' (Nwí-Bàrì 2002).

(7)   Preservation of prefixes on numeral classifiers in Kana (Ikoro 1996:101)

	Baan	**Kana**	**Gokana**
?	*pã*	*ápã*	*pã*
skin	*kpá*	*ákpá*	*kpá*
piece	*péé*	*ápéé*	*péé*
grain	*súú*	*ásúú*	*súú*

2.3.   Verb Morphology

We have now examined the two basic features that have been used to classify and sub-group the Ogonoid languages and have demonstrated that the diachronic picture is far from as clear as would be desirable. This problem is further magnified when examining structural data from these languages. In Ogonoid languages there are bound, cliticized or structurally configured sets of object and subject pronominal elements, and a range of elements encoding tense, aspect, and mood categories. Many of the details correspond across the languages, but the diachronic picture is far from clear. One of the paradoxes of Ogonoid structure is that while lexically the languages seem more or less to reflect accepted *Stammbaum* differentiation (and to an extent exhibit consistent phonological differences as well), the situation when considering equally important morphosyntactic data from a comparative/diachronic perspective is quite different. As a group of languages with no known pre-historical data, argumentation in the analysis of the historical developments in Ogonoid is subject to circularity in the reasoning that attributes structures to the proto-language. Specifically, it is not (yet?) possible in principle to distinguish between shared innovations and common archaic retentions. Given the current geographic configuration of the Ogonoid languages, the most surprising of the structural commonalities is perhaps the significant structural correlations between Gokana and Eleme (their significant differences being perhaps less surprising), as outlined below. In particular, there appear to be virtually no structural innovations in East Ogonoid, questioning its homogeneity as a group that opposes Eleme-Baan in a tree structure.

2.3.1.   Object Encoding

Among the most obvious similarities and differences exhibited across the Ogonoid languages is the inflection of pronominal objects. Like many Benue-Congo languages, Ogonoid languages use a set of grammaticalized object pronominals which serve as either clitics or object.[5]

2.3.1.1. First Singular Object

The first singular object marker appears cognate across the Ogonoid languages, appearing with *m and a following front vowel, realized as non-high in Kana-Tai and high in Eleme. The Gokana form lacks a vowel. In any event, Kana-Tai

---

[5] These are represented as separate independent elements in Tai orthography, but we suspect they function as clitics in a manner similar to the object pronominals in Kana.

(*-mέ/-mε/-mE*), Eleme (*-mi*) and Gokana (*-m̋*) all have their own reflexes of a putative Proto-Ogonoid construct *mI (8-11).

(8) <u>Kana</u>
    nɛ- mɛ́ kpuɡ̀ì
    give- 1SG money
    'give me money'
    (Wolff 1964:44)

(9) <u>Gokana</u>
    né- m̋ kpὲge
    give- 1SG money
    'give me money'
    (Wolff 1964:44)

(10)   <u>Eleme</u>
    né- mi ekpὶ̀ì
    give 1SG money
    'give me money'
    (Wolff 1964:44)

(11) <u>Tai</u>
    núm beὲ agara nú a beὲ me dὸὸ
    NEG.1SG PST respond thing 3SG PST 1SG do
    'I did not respond to what he did to me'
    (Nwí-Bàrì 2002:44)

2.3.1.2. Second Singular Object

In contrast to first singular, the second singular object marking presents a more complex picture. At first glance, it might appear that none of the major languages have cognate forms, but it may turn out that Eleme and Gokana share the same suffix. Eleme exhibits alternations between *n* and *r* followed by a nasalized vowel. Note the second singular object suffixes in Gokana (*-nì/-ni*) and Eleme (*-rũ*), respectively, both with high vowels ((13) and (14)). These contrast starkly with the Kana (12) counterparts (*-à/-a/-á*). This suggests that Gokana and Eleme either share an archaic form or innovated a new second singular object suffix that differs from that of Kana (which seems perhaps rather to reflect an element found also in Kana imperatives). Thus, either Gokana and Eleme pattern together to the opposition of Kana, or all three show distinct developments.

(12) <u>Kana</u>
    lú sɔ̃́ m̋- kúé=à
    come when 1SG-call=2SG
    'come when I call you'
    (Wolff 1964:44)

(13) <u>Gokana</u>
    dù sé m= kor= nì
    come when 1SG= call= 2SG
    'come when I call you'
    (Wolff 1964:44)

(14) <u>Eleme</u>
    m̀- ʔí- rũ
    1SG- marry- 2SG
    'I married you'

2.3.1.3. "Syntax" (or Morphophonology) of Object Encoding

One of the most striking features of Kana (as well as Tai, apparently) is the placement of the object pronominal element. Emphatic pronominal and nominal objects follow the verb, while clitic objects precede it (15-18). This appears to be a reanalysis of a second position clitic to a proclitic on the verb following an

auxiliary. If no auxiliary is present, the object marker appears enclitic to the verb. As with all Tai forms cited herein, the non-emphatic pre-verbal object elements are probably proclitic to the verb or enclitic to the auxiliary.

(15) <u>Kana</u>
m̀- weè kúé alo
1SG-PST call 2SG
'I called you'
(Ikoro 1996:208)

(16) <u>Kana</u>
m̀-weè a- kue
1SG-PST 2SG-call
'I called you'
(Ikoro 1996:207)

(17) <u>Tai</u>
gbárà beè nè yààló irì
Gbara PST give coat 1PL
'Gbara gave us a coat'
(Nwí-Bàrì 2002:30)

(18) <u>Tai</u>
à- beè- wa nè tum
3SG- PST- 3PL give advice
'he gave them some advice'
(Nwí-Bàrì 2002:60)

2.3.2. TAM Categories

When viewed from the perspective of the "traditional" *Stammbaum* of Ogonoid, the tense/aspect/mood categories in the language family show an array of confusing correspondences. These features include marking of past action, imperatives, verbal reduplication, and negative verb constructions.

2.3.2.1. Past

In past function, Kana uses an auxiliary *weè* (19), while Tai (20) and Gokana (21) appear to have virtually identical auxiliaries (*beè*) to the one found in Kana. This is indeed one of the main structural similarities between Gokana and Kana, albeit one that seems to be fully grammaticalized in Kana (23) but perhaps only partially so in Gokana (24). Although Eleme has no such past formation, it does exhibit a cognate perfect(ive) auxiliary *bere* (22), and thus this auxiliary cannot be convincingly used as the foundation for a putative East Ogonoid node.

(19) <u>Kana</u>
m̀- weè lu
1SG-PST come
'I came'
(Ikoro 1996:202)

(20) <u>Tai</u>
à beè lu
3SG-PST come
'he came'
(Nwí-Bàrì 2002:12)

(21) <u>Gokana</u>
m-beè fɛ bol
1SG-PST kill goat
'I killed a goat.'
(Wolff 1964:47)

(22) <u>Eleme</u>
a- bere ʔù
3SG-PERF die
'he has died'

(23) <u>Kana</u>
m̀ beè-ba nam
1SG PST-eat meat
'I ate meat'
(Wolff 1964:41)

(24) <u>Gokana</u>
m̀ ba nɔm
1SG eat meat
'I ate meat'
(Wolff 1964:41)

(25) <u>Eleme</u>
m̀ ba ǹnã̄
1SG eat meat
'I ate meat'
(Wolff 1964:42)

2.3.2.2. Progressive

Among the prefixes encoding TAM categories found throughout the Ogonoid languages is one encoding a progressive. Eleme (28) and Tai (29) have near-identical forms (*ga-/ka-*), a possible archaic feature to be attributed to Proto-Ogonoid. Kana has long *áá-,* and assuming this element is cognate with that in Eleme and Tai, it differs in that it has lost the initial consonant and is frequently fused with the subject prefix (26). Gokana (27) has a front vowel in this element (*é-*), which optionally appears with an initial voiced velar plosive (*gé-*).

(26) <u>Kana</u>   (27) <u>Gokana</u>   (28) <u>Tai</u>   (29) <u>Eleme</u>
áá- lú   à é- dú   à ga lu   è- ka- dʒú
PROG-come   3SG PROG-come   3SG PROG come   3SG-PROG-come
'he is coming'   'he is coming'   'he is coming'   'he is coming'
(Wolff 1964:46) (Wolff 1964:46)   (Nwí-Bàrì 2002:22)

2.3.2.3. Habitual

Habitual appears to be a separate development in Gokana, Kana-Tai, and Eleme (30-33). In Kana-Tai, the auxiliary *wéè/wée* is found in a habitual function, while Gokana makes use of an auxiliary *ʔóro*. Eleme, on the other hand, uses a suffix *-a*.

(30) <u>Kana</u>   (31) <u>Tai</u>
m̀- wée lu   nee wée mènà
1SG-HAB come   men HAB reproduce
'I usually come'   'human beings reproduce'
(Ikoro 1996:171)   (Nwí-Bàrì 2002:43)

(32) <u>Eleme</u>   (33) <u>Gokana</u>
ǹ- de- á   ǹdʒa   m̀ ʔóro- bà nɔm
1SG-eat- HAB   food   1SG HAB- eat meat
'I usually eat food'   'I usually eat meat'
                            (Wolff 1964:47)

   Note that the auxiliary that gives rise to the Kana-Tai habitual (34) appears to be the same (although perhaps in a separate grammaticalization) to the one that underlies the Eleme perfect[ive] (35), and probably also the Kana and Gokana past auxiliaries.

(34) <u>Kana</u>   (35) <u>Eleme</u>
m̀- wée fà   nám   m̀- bere kɛ-   a   m̀bó
1SG-HAB weed farmland   1SG-PERF slaughter- HAB goat
'I used to weed farmland'   'I used to slaughter goats'
(Ikoro 1996:171)

What appears to be cognate with the Eleme habitual marker has the function of a perfect marker in Kana (36). In Tai, on the other hand (37), perfect is marked by an auxiliary in *nà*.

(36)	Kana			(37)	Tai		
	è-	fà-	a náḿ		à	nà	táa
	3SG.PERF-weed-PF farm				3SG PERF finish		
	'he has weeded a farmland'				'it has finished'		
	(Ikoro 1996:182)				(Nwí-Bàrì 2002:56)		

2.3.2.4. Plural Imperative

The last TAM category to be examined here is the plural imperative. This appears to be marked by a suffixed *-i* in Gokana and Eleme (40) but in Kana by a preverbal *búí* (38) or *bí-*, the latter followed by a postverbal *-aa* (39). Once again, Gokana and Eleme pattern together in distinction to Kana.

(38)	Kana		(39)	Kana			(40)	Eleme	
	búí	dé		bì-	dɔ-	aa		de-	i
	2PL	eat		2PL-	fall-	PL.IMP		eat-	2PL
	'eat (PL)!'			'fall (PL)!'				'eat (PL)!'	
	(Wolff 1964:46)			(Ikoro 1996:190)				(Wolff 1964:46)	

2.3.3. Verbal Reduplication

Eleme stands apart from Kana and Gokana in its extensive use of verbal reduplication in finite clauses, marking functions such as future (41) and continuation (42). Note that in Kana and Gokana, reduplication forms a non-finite gerund of the verb.

(41)	Eleme			(42)	Eleme			
	ń-	dé-	dé		àɲè	ka-	de-	dé ǹdʒa
	1SG-	RFUT-	eat		3SG	PROG-	RCNT-	eat food
	'I will eat'				'he is still eating'			

2.3.4. Negative Verbal Constructions

Negative verbal constructions vary considerably among the Ogonoid languages. Kana uses emphatic or lengthened forms of the subject pronouns, as illustrated in (43) and (44). The third person forms in Kana consist of a fusing of the negative particle *n-* and the lengthened pronominal element (45). It may be that the lengthening in the first and second person Kana forms reflects the absorption of the *n-* element seen in third person (perhaps originally 1 > 2/3).

(43) <u>Kana</u>   (44) <u>Kana</u>   (45) <u>Kana</u>

ḿm   maa   lu- na   ḿm= ye kue   naa   kúé=ye
1SG.NEG 1SG.PROG come-REP   1SG.NEG= 3SG call   NEG.3SG call=3SG
'I am not coming again'   'I did not call him'   'he did not call him'
(Ikoro 1996:157)   (Ikoro 1996:339)   (Ikoro 1996: 339)

In Tai, there are *n*-marked pronominals and a negative particle, depending on the conjugation (46-48). Tonal alternations are also evident in the Tai negative.

(46) <u>Tai</u>   (47) <u>Tai</u>   (48) <u>Tai</u>

m̀ ga   si   nú m ga   si   à   náa   beè gầầ
1SG PROG go   NEG 1SG PROG go   3SG NEG.3SG PST hesitate
'I am going'   'I am not going'   'he did not hesitate'
(Nwí-Bàrì 2002:42) (Nwí-Bàrì 2002:42)   (Nwí-Bàrì 2002:22)

In Gokana (49-51), there is a whole series of *n*-pronominals in the negative (with the tonal pattern MH for NEG.FUT, HL elsewhere). Note the differences in the form of the negative pronominal based on presence (50) and absence (51) of the past auxiliary.

(49) <u>Gokana</u>   (50) <u>Gokana</u>   (51) <u>Gokana</u>

néὲ   mɔn ὲ   náà   beὲ-mɔn   im náὲ   mɔn im
NEG.1PL see 3SG   NEG.3SG PST-see   1SG NEG.3SG see 1SG
'we did not see him'   'he did not see me'   'he did not see me'
(Wolff 1964:48)   (Wolff 1964:48)   (Wolff 1964:48)

Eleme presents a more complex historical picture. As in the other languages, there appears to be a **n* prefix (often realized, as is generally the case in Eleme, as *r* plus nasalized vowel) preceding the pronominal subject elements in the verbal complex in at least some paradigms. In the past, as in (52) and (53), and habitual (54) forms, there is a reduplication of the stem; in the negative future and negative progressive, a negative modal or irrealis inflection is found (55). The past and present are distinguished both tonally (MH in the past, HM in the habitual) and by the presence (past) or absence (habitual) of the negative "particle" *(*n/r)* element found across the Ogonoid languages in negative verb formations.

(52) <u>Eleme</u>   (53) <u>Eleme</u>

rĩ̃ m̀- be- bé òbe   ró- be- bé- i òbe
NEG 1SG- RNEG- fight fight   NEG.2- RNEG- fight-2PL fight
'I didn't fight'   'you didn't fight'

22

(54) <u>Eleme</u>

m̀-	bé-	be	òbe
1SG-	RNEG-	fight	fight

'I don't fight'

(55) <u>Eleme</u>

ré	m̄-	bá-	be	òbe
NEG	1SG-	MOD.NEG-	fight	fight

'I will not fight'

3.   Comparative Ogonoid Revisited

In summary, we offer the following table showing a range of structural similarities and differences across four members of the Ogonoid language family.

(56)   Partial Summary of Comparative Ogonoid Structural Features

	Eleme	**Kana**	**Tai**	**Gokana**
Syllable codas	-	+	+	+
Noun class prefixes	+[6]	±[7]	-?	-[8]
Verbal reduplication	+	±	-	-
Negative verb	+	-	-	+
Past auxiliary	-/+	+	+	±
Plural imperative suffix	+	-	-	+
Object in A/SVC	+	-	-	+
1SG subject	+	-	?	+
***g*-progressive**	+	-	+	+

Eleme clearly differs from the "East" Ogonoid languages in a number of ways, but this is not to be understood as Kana and Gokana's forming a coherent whole in opposition to Eleme. Among the salient features found in Eleme that characterize this language vis-à-vis its attested sister languages are the complete lack of coda consonants, the relatively common if lexicalized occurrence of noun class prefixes, and the use of reduplication in negative and other finite verbal clauses. Kana differs from Eleme and Gokana in its system of object marking in auxiliary and serial verb constructions and in the formation of plural imperatives. Kana, Gokana, and Eleme all differ in progressive auxiliary/prefix and instrumental verbal suffix allomorphy, the system of future marking, the 2nd singular object marker (although see above for arguments that Gokana and Eleme may pattern together here) and in negative formations. Conversely, Kana, Gokana and Eleme all show the same structure in 1st singular object marking (the form of the marker, not its placement) and subject marking. Gokana differs from the rest of Ogonoid in having no trace of reduplication or emphatic/prosodic lengthening in negatives and the near complete lack of even vestigial noun class prefixes. Also, Tai differs from Kana in preserving the *g*-initial progressive marker and in its perfect(ive) construction. Finally, what little is known of Baan suggests that it differs from Eleme in that it has no syllabic nasal noun class prefixes preserved,

---

[6] Vestigial noun class prefixes in very small number of forms.
[7] Remnant forms with noun class prefixes occur to a small degree.
[8] Noun class prefixes in perhaps a half-dozen total words.

only vocalic ones. In summary, the details of the actual internal relationships of the Ogonoid subgroups of Benue-Congo languages remain a subject for further research.

Abbreviations

1	first person	**PERF**	perfect	**RCNT**	continuous (reduplication)
2	second person	**PF**	perfective		
3	third person	**PL**	plural	**REP**	repetitive
HAB	habitual	**PROG**	progressive	**RFUT**	future (reduplication)
IMP	imperative	**PRTCL**	particle	**RNEG**	negative (reduplication)
MOD	modal	**PST**	past	**SG**	singular
NEG	negative				

References

Blench, R., and K. Williamson. 2000. Niger-Congo languages. In B. Heine and D. Nurse, eds., *African Languages*, 11-42. Cambridge: Cambridge UP.

Faraclas, Nicholas G. 1986. Cross River as a model for the evolution of Benue-Congo nominal class/concord systems. *Studies in African Linguistics* 17(1):39-54.

--. 1989. Cross River. In John Bendor-Samuel, ed., *The Niger-Congo Languages*, 377-99. Arlington: SIL.

Ikoro, Suanu. 1996. *The Kana Language*. Leiden: CNWS Publications.

Nwí-Bàrì, Wilson Kpàkpàn. 2002. *Tèè-English Dictionary. Tabtàb Tèè Turé bu Bèkéé*. Port Harcourt: Onyoma Research Publications.

Williamson, Kay. 1985. How to become a Kwa language. In Adam Makkai and Alan K. Melby, eds., *Linguistics and Philosophy: Essays in Honor of Rulon S. Wells*, 428-443. Amsterdam: John Benjamins.

Wolff, H. E. 1964. Synopsis of the Ogoni languages. *Journal of African Linguistics* 3:38-51.

Oliver Bond
Endangered Languages Academic Program
Department of Linguistics, SOAS
Thornhaugh Street, Russell Square
London WC1H 0XG
UK

oliver.bond@soas.ac.uk

Gregory D. S. Anderson
Department of Linguistics
1290 University of Oregon
Eugene, OR 97403-1290
USA

gdsa@darkwing.uoregon.edu

A Fixed Hierarchy for Wolof Verbal Affixes

LESTON BUELL and MARIAME SY
University of California, Los Angeles

0.   Introduction

This article examines the valence-changing verbal suffixes of Wolof, a West Atlantic language spoken primarily in Senegal and the Gambia. The challenge is to account for the various attested and ungrammatical suffix orders in forms where two or more suffixes are combined. It will be shown that a straightforward head movement account is inadequate, if Baker's (1998) Mirror Principle is assumed. Instead, an analysis using phrasal movement will be. Wolof verbal extensions are taken to have a derivational pattern similar to that of verbal complexes and to be amenable to a similar analysis, as in Koopman and Szabolcsi (2000). The idea behind this article is taken from an analysis sketched out in a Koopman (2004), which itself is a response to Buell and Sy (2004).

1.   Characteristics of the Analysis

A close correspondence between syntax and morphology is assumed. Words are taken to be built in the syntax, using only the usual syntactic mechanisms. Furthermore, with respect to head movement, the surface order of morphemes in a complex head is initially assumed to reflect the order of head adjunction operations, as required by the Mirror Principle. Establishing the underlying hierarchy of Wolof's valence-changing morphemes will lead us to observe Mirror Principle violations, forcing us to abandon a head movement approach. Instead, phrasal movement must be appealed to, which will generate all the attested morpheme orders.

Blocking unattested affix orders will depend on the concept of complexity filters (Koopman and Szabolcsi 2000). A complexity filter can be described as a lexically specified constraint on a head which requires its specifier at spell-out to not exceed a particular degree of complexity, defined here as the degree to which the feature to be checked is embedded. The concept of complexity filter will be taken up in greater detail at the point where it becomes relevant.

We argue that the range of Wolof valence-changing morpheme orders points to a single underlying hierarchy, namely: Impersonal Causative > Causative > Benefactive > Instrumental > Verb. The analysis presented here thus contrasts

with Buell and Sy (2004), which uses a variable hierarchy to account for the various orders.

2. Simple Cases of Verbal Extensions

Here is a simple Wolof sentence, which we will use as a basis to illustrate the valence-changing suffixes:

(1)   Faatu togg -na   jën   wi.
        Faatu cook -3sg. fish   the
        'Faatu cooked the fish.'

The valence-changing suffixes we are concerned with are illustrated below, with the abbreviations that will be used in the glosses:

(2)   Benefactive (Ben, al_{Ben})
        Faatu togg -*al*   -na   Gàllaay jën wi.
        Faatu cook -BEN   -3sg. Gàllaay fish the
        'Faatu cooked the fish for Gàllaay.'

(3)   Instrumental (Instr, e_{Instr})
        Faatu togg -*e*   -na   jën   wi (ag) diwtiir.
        Faatu cook -INSTR -3sg. fish   the with palm.oil
        'Faatu cooked the fish with palm oil.'

(4)   Causative (Caus, loo_{Caus})
        Faatu togg -*loo* -na   Gàllaay jën wi.
        Faatu cook -CAUS -3sg. Gàllaay fish the
        'Faatu had/made Gàllaay cook the fish.'
        (causer = Faatu, causee = Gàllaay)

(5)   Impersonal causative (ImpCaus, $lu_{ImpCaus}$)
        Faatu togg -*lu*   -na   jën wi.
        Faatu cook -IMP.CAUS -3sg. fish the
        'Faatu had someone cook the fish.'

We will assume that these affixes are merged somewhere above the lexical verb, as illustrated below with the benefactive suffix:

(6)   BenP
        /\
          Ben'
          /\
        al   VP

While it is these suffixes we are concerned with in this article, other suffixes

exist, including a locative, subject and object argument absorbers, and a different causative suffix used only with stative verbs. The remainder of the article focuses on deriving the orders of different combinations of the Ben, Instr, Caus, and ImpCaus illustrated above. (7) contains an exhaustive list of the grammatical combinations of these affixes.

(7)   V-*al*$_{Ben}$-*e*$_{Instr}$, V-*e*$_{Instr}$-*loo*$_{Caus}$, V-*loo*$_{Caus}$-*al*$_{Ben}$, V-*lu*$_{ImpCaus}$-*al*$_{Ben}$,
       V-*lu*$_{ImpCaus}$-*loo*$_{Caus}$, V-*e*$_{Instr}$-*loo*$_{Caus}$-*al*$_{Ben}$, V-*al*$_{Ben}$-*e*$_{Instr}$-*loo*$_{Caus}$,
       V-*lu*$_{ImpCaus}$-*al*$_{Ben}$-*e*$_{Instr}$, V-*lu*$_{ImpCaus}$-*loo*$_{Caus}$-*al*$_{Ben}$,
       V-*lu*$_{ImpCaus}$-*loo*$_{Caus}$-*al*$_{Ben}$ -*e*$_{Instr}$ V-*lu*$_{ImpCaus}$ -*e*$_{Instr}$[1]

3.   Establishing the Underlying Hierarchy

For the moment, we will ignore the ImpCaus suffix, whose morphological position is constant, and consider just the Caus, Ben, and Instr suffixes. Taking into account two-affix combinations to establish the structural hierarchy for these affixes, assuming that they are attached to the verb by head movement leads to a contradiction, as evidenced in the following cases.

First consider combinations of the Instr and Caus suffixes. The order for this combination is always V-*e*$_{Instr}$-*loo*$_{Caus}$:

(8)   a.   Gàllaay dóór -*e*   -*loo* -na Faatu xeer bi (ag) bant.
             Gàllaay hit   -INSTR -CAUS -3sg. Faatu stone the with stick
             'Gàllaay made Faatu hit the stone with a stick.'
       b.   *   … dóór -*loo*   -*e*   -na ...
                   hit   -CAUS -INSTR -3sg.

In a head movement account, this requires the Caus^0 head to be higher than the Instr^0 head in the structure. This is because, by the Mirror Principle, the Caus^0 suffix is farther from the verb root than the Instr^0 suffix and hence higher in the structure than Instr^0, and Instr^0 must hence attach to the verb root first. The derivation is shown here:

(9)   a.   CausP   b.   CausP   c.   CausP

            loo   InstrP   *loo*   InstrP   [V$_i$+*e*]$_j$+*loo*   InstrP

                e   VP   V$_i$+*e*   VP   t$_j$ VP

                     V   t$_i$ t$_i$

---

[1] This combination and order of *lu*$_{ImpCaus}$ and -*e*$_{Instr}$ occurs in *wh*-questions questioning the instrument, but not in statements. There seems to be a co-occurrence restriction between these two affixes in certain types of structures; as these restrictions have not been investigated further, we leave them for future research. For our purposes, it only matters which order they appear in whenever they co-occur.

Now let's turn to the Ben and Instr suffixes, where the surface order is always V-al_{Ben}-e_{Instr}:

(10) a. Gàllaay togg -*al* -*e* -na Faatu yàpp diwtiir.
          Gàllaay cook -BEN -INSTR -3sg. Faatu meat palm.oil
          'Gàllaay cooked Faatu some meat with palm oil.'
      b. *…togg -*e* -*al* -na …
              cook -INSTR -BEN -3sg.

In a head movement account, and by the same Mirror Principle considerations just explained, this requires $Instr^0$ to be higher than Ben^0 in the structure: InstrP > BenP. Finally consider Caus and Ben, with only the surface order loo_{Caus}-al_{Ben}:

(11) a. Gàllaay bind -***loo*** ***-al (-lool)*** -na gan gi xale yi taalif.[2]
          Gàllaay write -CAUS -BEN -3sg. visitor the child the poem
          'Gàllaay made the children write the visitor a poem.'
      b. * … bind -***al*** ***-loo*** -na …
              write -BEN -CAUS -3sg.

In precisely the same way as before, in a head movement account, this requires Ben^0 to be higher than $Caus^0$ in the structure: BenP > CausP.

At this point, then, it would appear that we have established three facts: BenP > CausP, CausP > InstrP, and InstrP > BenP. However, by transitivity, these "facts" lead to a contradiction. Assuming BenP > CausP and CausP > InstrP, by transitivity we deduce BenP > InstrP, which contradicts our earlier conclusion InstrP > BenP. Therefore, we need to look for some other diagnostic to find the underlying hierarchy. In so doing, we might want to take into account the fact that in four-affix combinations only one order is possible:

(12) V -$lu_{ImpCaus}$ -loo_{Caus} -al_{Ben} -e_{Instr}

Given that we expect the increased number of affixes to lead to a greater degree of restrictiveness, it seems reasonable to take this order as indicative of the following hierarchy of merger for all Wolof verb forms:

(13) $lu_{ImpCaus}$ > loo_{Caus} > al_{Ben} > e_{Instr} > V

Further evidence in support of the idea that the hierarchy in (13) is correct can be found in the relative semantic scope of the Caus and Instr predicates.

---

[2] Certain combinations of affixes which result in the concatenation of two vowels undergo syncope or coalescence, depending on the types of vowels that come into contact. In such cases, pronunciation is indicated in parentheses.

(14) a. Jàngalekat bi bind -lu   -loo -al   -e   (luloole)
          teacher   the write -IMPCAUS -CAUS -BEN -INSTR

          -na xale yi gan gi taalif bi **kereyon**.
          -3s child the guest the poem the pencil
          'The teacher made the children have someone write the poem
          for the guest **with a pencil**.'
          [force [someone write the poem *with a pencil*]]

     b. *   … **yar**.
              whip
          '…**with a whip**.'
          *[force *with a whip* [someone to write a poem]]

As indicated by the bracketed paraphrases, the instrumental predicate can modify the lower predicate WRITE A POEM in (14a), but cannot modify the higher, causative predicate as attempted in (14b). Thus, there seems to be a fixed hierarchy for ImpCaus and Instr: ImpCaus > Instr. What this example shows is that the instrument modifies the predicate which is embedded by the predicate of causation. This would be entirely unexpected if $Instr^0$ were actually merged higher than $Caus^0$. Assuming head movement would thus lead to admitting a Mirror Principle violation.

The form in (12) cannot be formed by head movement. Rather, it must be formed by moving the verb root to a position preceding the $ImpCaus^0$ head, the highest head in our hierarchy:

(15) ImpCausP
      VP_i ImpCaus'
          lu   CausP
              loo   BènP
                  al   InstrP
                      e t_i
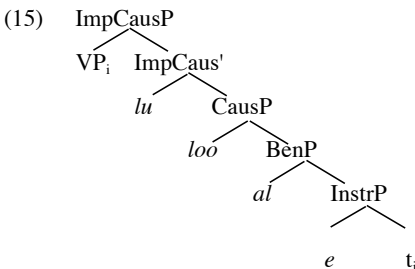
Most of the attested affix orders can, in fact, be derived in the very same fashion, as shown in the table (16), in which each of the morphemes used is numbered to make clearer the way the V (always the first head merged and hence the highest number) has moved alone to precede any higher affixes. These orders are called "stranding orders" because the verb strands its affix rather than pied-piping it.

(16)   Pure stranding orders

		Surface Order	ImpCaus	Caus	Ben	Instr	V
			lu	*loo*	*al*	*e*	V
a.	51234	V-*lu*$_{ImpCaus}$-*loo*$_{Caus}$-*al*$_{Ben}$ -*e*$_{Instr}$	1	2	3	4	5
b.	4123	V-*lu*$_{ImpCaus}$-*loo*$_{Caus}$-*al*$_{Ben}$	1	2	3		4
c.	4123	V-*lu*$_{ImpCaus}$-*al*$_{Ben}$-*e*$_{Instr}$	1		2	3	4
d.	312	V-*lu*$_{ImpCaus}$-*loo*$_{Caus}$	1	2			3
e.	312	V-*lu*$_{ImpCaus}$-*al*$_{Ben}$	1		2		3
f.	312	V-*lu*$_{ImpCaus}$-*e*$_{Instr}$	1			2	3
g.	312	V-*loo*$_{Caus}$-*al*$_{Ben}$		1	2		3
h.	312	V-*al*$_{Ben}$-*e*$_{Instr}$			1	2	3

4.   Successive Inversion and Complexity Filters

With a verbal head and valence-changing suffixes, we expect the verbal head to raise stepwise, with each successive suffix appearing at the end of the head which adjoins to it, as in the following case from Zulu (numbers in glosses refer to noun classes):

(17)   a.   Ba-   zo-   fihl$_3$-   el$_2$-   an$_1$-   a   imali.
            2.SBJ-   FUT-   hide-   APPL-   RECIP-   FV   9.money
            'They will hide money from each other.'
            [from each other [hide money]]

      b.   Ba-   zo-   fihl$_3$-   an$_2$-   el$_1$-   a   amaphoyisa
            2.SBJ-   FUT-   hide-   RECIP-   APPL-   FV   6.police
            'They will hide each other from the police.'
            [from the police [hide each other]]

In Zulu, as in other Bantu languages, the reciprocal suffix *-an* normally appears to the immediate right of the head whose argument it encodes. In (17b), *-an* encodes the direct object of the simple verb *fihl* 'hide' and thus appears to its immediate right, while in (17a), *-an* encodes an applicative object and thus appears to the immediate right of the applicative suffix *-el*. These facts are easily captured by assuming that *-an* is underlyingly lower than applicative *-el* in (17b), but higher than *-el* in (17a). The observed morpheme orders are obtained by

successive inversion (which could technically be implemented with either head movement or phrasal movement). For example, in (17a) we have the underlying hierarchy $Recip_1 > Appl_2 > V_3$. The verb first inverts with the applicative head, giving us $[V_3 \, Appl_2]$. Then this internally inverted constituent inverts with the higher $Recip_1$, yielding $[[V_3 \, Appl_2] \, Recip_1]$. It is this process we are terming successive inversion, which creates structures that obey the Mirror Principle.

Why should the Bantu verb root and suffixes undergo successive inversion while their Wolof counterparts do not? This contrast resembles a word order contrast found in verbal complexes in Germanic languages:

(18)   a.   to want$_1$ to have$_2$ to sing$_3$ (English)
          b.   singen$_3$ müssen$_2$ wollen$_1$ (German)
          c. $WANT_1 > MUST_2 > SING_3$ (underlying hierarchy)

The English and German phrases in (18) are assumed to be derived from the same underlying hierarchy in (18c), but the German inversion displays successive inversion while the English version does not exhibit any inversion at all.

The full range of patterns of such verbal complexes in Germanic languages and Hungarian is treated in Koopman and Szabolsci (2000). In that analysis, successive inversion of verbal complexes is assumed to take place unless something like a complexity filter prevents it. Our proposal is to derive all the Wolof orders in the same way as verbal complexes. Now that the hierarchy has been established, our next task is to find all the cases of successive inversion in Wolof. Using our hierarchy, there is only one case of total successive inversion in Wolof, namely $V\text{-}e_{Instr}\text{-}loo_{Caus}$:

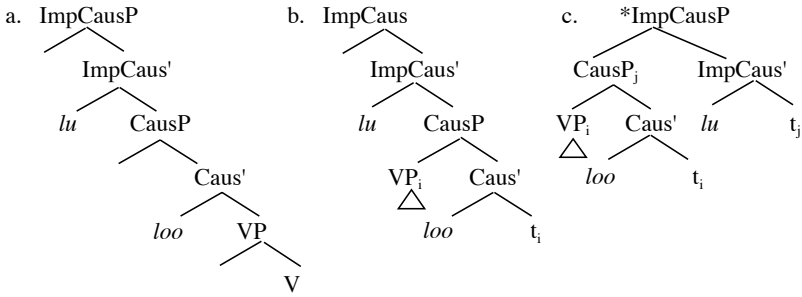(19) $V\text{-}e_{Instr}\text{-}loo_{Caus}$
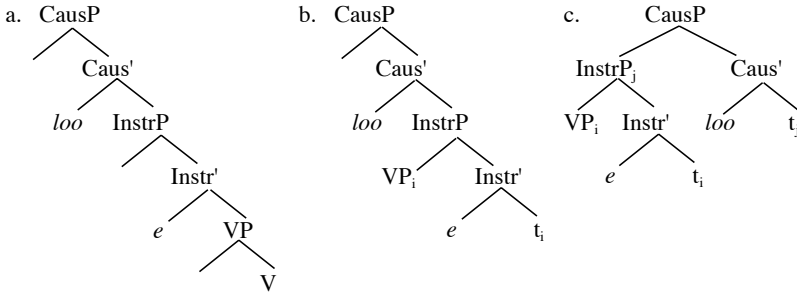
In this system, availability of successive inversion is expressed as the absence of complexity filters on any head to which a portion of the verbal complex raises. For our purposes here, a complexity filter constrains the degree to which the V^0 head can be embedded in this constituent. In (19c), the V^0 is not the head of the InstrP constituent that raises to the specifier of $Caus^0$. Rather, the V^0 is in the specifier of this InstrP. And yet the surface structure in which this V^0 is embedded

in the specifier of a specifier is grammatical. This indicates that loo_{Caus} is not lexically specified for a complexity filter. Complexity filters are assumed to be constraints on specifiers, which is why the successive inversion in (19) has been implemented with phrasal movement rather than head movement.

Now consider all of the forms containing $lu_{ImpCaus}$ above in (16). In all of these forms, $lu_{ImpCaus}$ immediately follows the verb stem. In other words, none of these forms use any degree of successive inversion. For example, for (16d), using successive inversion, we would get the morpheme order *V-loo_{Caus}-$lu_{ImpCaus}$, derived as follows:

(20)   *V-loo_{Caus}-$lu_{ImpCaus}$

a. ImpCausP
   ImpCaus'
   lu   CausP
          Caus'
          loo   VP
                  V

b. ImpCaus
   ImpCaus'
   lu   CausP
          VP$_i$ Caus'
          loo   t$_i$

c. *ImpCausP
   CausP$_j$ ImpCaus'
   VP$_i$ Caus'   *lu*   t$_j$
   loo   t$_i$

Such an ungrammatical derivation is ruled out by assuming that $lu_{ImpCaus}$ has a complexity filter, meaning that while it requires a verb in its specifier at some point in the derivation, as in (21a), at spell-out the verb cannot be embedded more deeply in its specifier, as in (21b):

(21)   a. VP not embedded

   ImpCausP
   VP   ImpCaus'
        lu   …

   b. VP too embedded for $lu_{ImpCaus}$ complexity filter

   *ImpCausP
   XP   ImpCaus'
   VP X' *lu*   …
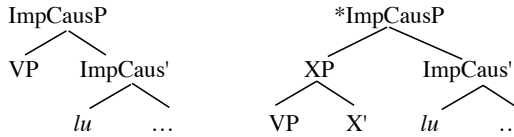
Similarly, in the form combining al_{Ben} and e_{Instr}, the former must immediately follow the verb.

(22)   V-*al*$_{Ben}$-*e*$_{Instr}$

		Surface Order		ImpCaus	Caus	Ben	Instr	V
				lu	*loo*	*al*	*e*	V
a.	312	V-*al*$_{Ben}$-*e*$_{Instr}$				1	2	3
b.	*321	*V-*e*$_{Instr}$-*al*$_{Ben}$						

Because the successive inversion order is not available, we conclude that *al*$_{Ben}$ also has a complexity filter of the same form as the one for *lu*$_{ImpCaus}$.

   To summarize, then, we have isolated the following complexity filters for the four affixes under consideration:

(23)   a.   *lu*$_{ImpCaus}$ has a complexity filter.
         b.   *al*$_{Ben}$ has a complexity filter.
         c.   *loo*$_{Caus}$ does not have a complexity filter.
         d.   Instr is the lowest head in our hierarchy; no complexity filter could be detected. (The verb will never be embedded enough to trigger an effect.)

8.   Problems

All the grammatical orders can be obtained with the fixed hierarchy. However, issues do arise in ruling out a few of the ungrammatical orders. One such issue concerns the way in which *al*$_{Ben}$ and *loo*$_{Caus}$ combine.

(24)   V-*loo*$_{Caus}$-*al*$_{Ben}$

		Surface Order		ImpCaus	Caus	Ben	Instr	V
				lu	*loo*	*al*	*e*	V
a.	312	V-*loo*$_{Caus}$-*al*$_{Ben}$			1	2		3
b.	*321	*V-*al*$_{Ben}$-*loo*$_{Caus}$						

   The ungrammatical order in (26b) is problematic for the simple reason that we have already established that *loo*$_{Caus}$ does not have a complexity filter. So, we must say that *al*$_{Ben}$ forces specifier extraction of VP, but why it should do so is not understood, given the ungrammatical structure in (25) and the grammatical orders in (26a).

(25)   *V-al_{Ben}-loo_{Caus}


```
                    CausP
              _____/_____
         BenP_j            Caus'
        __/\__           __/\__
      VP_i   Ben'      loo      t_j
            __/\__
          al     t_i
```
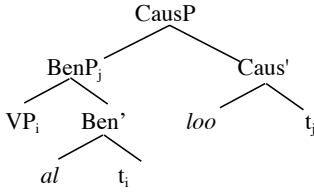
(26)   Caus+Ben+Instr

		Surface Order		ImpCaus	Caus	Ben	Instr	V
				lu	*loo*	*al*	*e*	V
a.	4231	V-al_{Ben}-e_{Instr}-loo_{Caus}						
b.	4312	V-e_{Instr}-loo_{Caus}-al_{Ben}						
c.	*4321	*V-e_{Instr}-al_{Ben}-loo_{Caus}		1	2	3	4	
d.	*4123	*V-loo_{Caus}-al_{Ben}-e_{Instr}						

First we must note that the choice between two grammatical orders in (26a) and (26b) cannot be a purely morphological reordering independent of syntax, because the preferred argument orders used with the two forms are different:

(27) Gàllaay door -**al** -**e**   -**loo** -na Faatu xale yi bant xeer.
     Gàllaay hit   -BEN -INSTR -CAUS -3sg. Faatu child the stick stone
     a.   'Gàllaay made the children hit a stick with a stone for Faatu.'
           …V -al_{Ben} -e_{Instr} -loo_{Caus} Beneficiary Causee…
     b.   ?? 'Gàllaay made Faatu hit a stick with a stone for the children.'
           ?? …V -al_{Ben} -e_{Instr} -loo_{Caus} Causee Beneficiary…
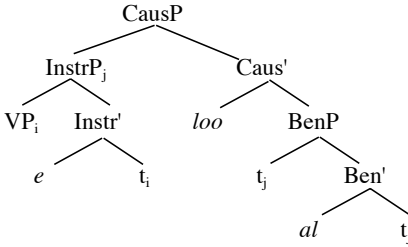
(28) Gàllaay door -**e**   -**loo** -**al** -na Faatu xale yi bant xeer.
     Gàllaay hit   -INSTR -CAUS -BEN -3sg. Faatu child the stick stone
     a.   ?? 'Gàllaay made the children hit a stick with a stone for Faatu.'
           ?? …V -e_{Instr} -loo_{Caus} -al_{Ben} Beneficiary Causee…
     b.   'Gàllaay made Faatu hit a stick with a stone for the children.'
           …V -e_{Instr} -loo_{Caus} -al_{Ben} Causee Beneficiary…

Turning to the ungrammatical orders, the pure successive inversion form in (26c) is ruled out because al_{Ben} has a complexity filter, while the form would require the complex phrase [V+e_{Instr}] in its specifier. The pure stranding form in
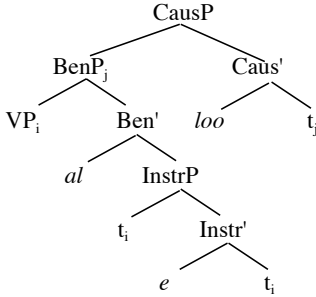
(26d) is ruled out because partial successive inversion is possible, employed in the (26a,b) forms. As for (26b), V-e_{Instr}-loo_{Caus}-al_{Ben}, assuming our fixed hierarchy, the bracketing must be [V-e_{Instr}]-loo_{Caus}-al_{Ben}. This form respects the complexity filter on al_{Ben}. (Recall that complexity filters are filters on surface representations, not on derivations.) [V-e_{Instr}] is extracted from the specifier of al_{Ben} and moved to the specifier of loo_{Caus}:

(29)   V-e_{Instr}-loo_{Caus}-al_{Ben}

```
                    CausP
                   /     \
            InstrP_j      Caus'
            /    \        /    \
          VP_i  Instr'  loo    BenP
          /      \            /    \
         e       t_i         t_j   Ben'
                                   /    \
                                  al    t_j
```

It is the grammatical order V-al_{Ben}-e_{Instr}-loo_{Caus} in (26a) that is problematic. Given our hierarchy, the bracketing must be [V-al_{Ben}-e_{Instr}]-loo_{Caus}, with this structure:

(30)   V-al_{Ben}-e_{Instr}-loo_{Caus}

```
                    CausP
                   /      \
            BenP_j         Caus'
            /    \         /    \
          VP_i  Ben'     loo    t_j
                /   \
               al   InstrP
                    /    \
                  t_i    Instr'
                         /    \
                        e     t_i
```

The problem is that (30) is structurally identical to the simpler, but ungrammatical order *V-al_{Ben}-loo_{Caus}, seen above in (25). It is not clear how V-al_{Ben}-e_{Instr}-loo_{Caus} can be allowed without also allowing the ungrammatical *V-al_{Ben}-loo_{Caus}.

9.   Conclusion

We have seen that using head movement to derive valence-changing affix orders in Wolof leads both to contradictions and to incorrect scopal predictions if the Mirror Principle is assumed. An alternative account employing phrasal movement and a fixed structural hierarchy was shown capable of deriving the attested orders.

Problems encountered using the phrasal approach involved ruling out a few ungrammatical orders. However, this is not a particular disadvantage of our account over one which employs either head movement or a variable hierarchy, as any type of account will face the same problem.

Conversely, the phrasal account seems to have certain advantages. First, it allows us to maintain a close correspondence between morphological order and syntactic structure. Second, it allows us to treat cross-linguistic variation in valence-changing morpheme orders in the same way as verbal complexes, in which a similar range of orders is observed. And finally, a phrasal fixed hierarchical account presents the learner with fewer options, restricting the number of possible underlying hierarchies (a real problem if Mirror Principle violations are admitted), and, assuming the possibility of forming words in syntax, perhaps restricting the domain in which the learner must choose between derivations employing head movement and those employing phrasal movement.

References

Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Buell, Leston, and Mariame Sy. 2004. Affix ordering in Wolof applicatives and causatives. Paper presented at the Annual Conference on African Linguistics (ACAL) 35. Submitted for publication in *Trends in African Linguistics*.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Koopman, Hilda. 2004. Some aspects of the syntax-phonology interface. UCLA Phonology Seminar presentation handout.

Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.

Leston Buell
Leiden University Centre for Linguistics
Van Wijkplaats 4
2311 BX Leiden, The Netherlands

bulbul@ucla.edu

Mariame Sy
Columbia University Institute of African Studies
1103 International Affairs Building
420 West 118th Street
NY, NY 10027-3331

sms2168@columbia.edu

S-O-V-X Constituent Order and Constituent Order Alternations in West African Languages

DENIS CREISSELS
Université Lumière (Lyon 2)

0.   Introductory Remarks

In current presentations of constituent order typology, the notion of *S-O-V* constituent order is implicitly treated as more or less equivalent to the notion of verb final language. The reason for this confusion is that, in most languages, objects and obliques occupy the same position in relation to the verb, and differ only in a tendency of objects to stand closer to the verb, so that *S-V-O*, *S-O-V*, *V-S-O* and *V-O-S* can generally be considered equivalent to *S-V-O-X*, *S-X-O-V*, *V-S-O-X* and *V-O-X-S* (X = oblique). This, however, does not hold for languages in which core syntactic terms (S and O) precede the verb, and all obliques follow it (*S-O-V-X* constituent order), as in Bambara (1), or in Soninke (2).

(1)   Bambara[1]

    a.   sékù   bɛ́   mǎdù kálán tùbàbùkán ↓ná
        Sékou PM   Madou teach   French   Po
        'Sékou is teaching French to Madou.'

    b.   sékù   jé   mǎdù délí wárí   ↓lá   kúnùn
        Sékou PM   Madou ask   money Po   yesterday
        'Sékou asked Madou for money yesterday.'

    c.   sékù   jé   mǎdù nèní à   mùsó ↓ɲɛ́ná sìɲɛ́ fìlà
        Sékou PM Madou teach 3s   wife   Po   time two
        'Sékou insulted Madou twice in the presence of his wife.'

    d.   ù   bɛ́nà fántà dí à   mà   mùsó ↓yé
        3P PM   Fanta give 3S Po   wife Po
        'They will give him Fanta as a wife.'

---

[1]Abbreviations: APPL=applicative, CAUS=causative, DAT=dative, DEF=definite, DEM= demonstrative, P(L)=plural, PM=predicative marker, Po=postposition, Pr=preposition, S=singular, TAM = tense-aspect-modality marker.

(2)   Soninke

    a.   fààtú   dà   t́ijè-n   q́obó   sá χà-n   ŋá
        Fatou   PM   meat-DEF   buy   market-DEF Po
        'Fatou has bought meat at the market.'

    b.   fààtú   dà   t́ijè-n   ýigándí   lémínè-n   ŋá
        Fatou   PM   meat-DEF   eat-CAUS child-DEF Po
        'Fatou had the child eat meat.'

    c.   ó   dà   χáálìsí ké   kínì à   ýi
        1P   PM   money DEM   give 3S   Po
        'We gave him/her the (aforementioned) money.'

It has often been claimed that the *S-O-V-X* type of constituent order is one of the morphosyntactic features that concern a proportion of African languages significantly higher than that observed at world level, and that within the limits of the African continent, this type of constituent order pattern shows a particular concentration in West Africa, where in addition to the whole Mande family it is found also, at least to some extent, in languages belonging to the Kwa, Gur, Kru, Songhay, and Atlantic families.

This is certainly not entirely false, in the sense that the canonical variety of the *S-O-V* pattern (with the verb in clause-final position) is particularly rare in West Africa, where its only representatives are Ijo and Dogon, whereas constituent order patterns that do not fit the current *S-V-O* vs. *S-O-V* distinction are common in West Africa. But the West African languages whose constituent order patterns cannot be straightforwardly identified as belonging to the *S-V-O* or to the *S-O-V* type show in some important respects a variety that may cast some doubt on Heine's claim that all these languages can be grouped into a single type (his "type B"), and on more recent claims that, taken as a whole, they provide converging evidence supporting the reconstruction of a constituent order of the Mande type in Proto-Niger-Congo.

The point is that most authors seem to take for granted that the *S-O-V-X* pattern found in the languages of the Mande family, and in a few other languages that share with Mande the absence of constituent order alternations, is also found as the alternative to *S-V-O-X* in the West African languages that have *V-O ~ O-V* alternations. What I would like to show is that things are not so simple, and that in the languages that have *V-O ~ O-V* alternations, the variant characterized by the anteposition of the object to the verb differs in some important respects from the *S-O-V-X* constituent order that constitutes the only possible constituent order in Mande languages. In other words, what I would like to show is that West African languages with *V-O ~ O-V* alternations cannot be straightforwardly characterized as having an alternation between a canonical *S-V-O(-X)* constituent order and a *S-O-V-X* constituent order of the Mande type. An important conclusion will be that data from West African languages with *V-O ~ O-V* alternations should be used with extreme caution in attempts at reconstructing Proto-Niger-Congo constituent order.

1.   The *S-O-V-X* Constituent Order in Mande Languages
1.1.   Rigidity of the *S-O-V-X* Constituent Order

In Mande languages, the *S-O-V-X* constituent order is absolutely rigid in the sense that it is neither restricted to particular types of clauses, nor conditioned by certain characteristics of the object NP. Moreover, Mande languages are extreme "configurational" languages, in which changes in the position of the NP representing a given participant always imply other changes in the construction. In cases when the same participant can optionally be represented by a term inserted between *S* and *V*, or by a term following *V*, the second construction implies the use of adpositions. Consequently, the change in constituent order is not the result of a mere possibility of choice within the frame of the same construction, and must be analyzed as the consequence of a valency alternation of the antipassive type (demotion of the second argument of transitive verbs to the status of oblique), as illustrated by ex. (3).

(3)   Bambara

a.   mǎdù   bɛ́nà   dúmúní   ꜜbán   ($S_A O_P V$)
       Madou   PM   food   finish
       'Madou will finish the food.'

b.   mǎdù   bɛ́nà   bán   dúmúní   ꜜná   ($S_A V X_P$)
       Madou   PM   finish   food   Po
       'Madou will finish the food.'

1.2.   Never More than One Nominal Term Inserted Between *S* and *V*

In Mande languages, the patient of typical transitive verbs regularly occurs inserted between the subject and the verb, i.e., in the syntactic role of object, but Mande languages do not have constructions in which a second nominal term would be inserted between the subject and the verb, either as a second object in a double object construction, or in an oblique role.

This means in particular that, in the construction of verbs of giving, which cross-linguistically are the most typical ditransitive verbs, one participant only (either the thing given, or the recipient) can be encoded as an object, and the only possibility for the third participant is to appear as an oblique, as illustrated by ex. (4) to (6).

(4)   Bambara

a.   sékù   jé   wári   ꜜdi   mǎdù   mà
       Sékou   PM   money   give   Madou   Po
       'Sékou gave money to Madou.'

b.   sékù   jé   mǎdù   sɔ́n   wári   ꜜlá   ~ ... ní   wári   ꜜyé
       Sékou   PM   Madou   give   money   Po   Pr   money   Po
       'Sékou gave money to Madou.'

(5)   Soso

    a. ń   nìngéé fíi-mà   í   má
        1S   cow   give-TAM   2S   Po
        'I will give you a cow.'

    b. ń   í   kíi-mà   nìngéé   rá
        1S   2S   give-TAM   cow   Po
        'I will give you a cow.'

(6)   Mende (Innes 1971)

    a. mbɛí ve   kpaná   wɛ
        rice   give Kpana   Po
        'Give the rice to Kpana.'

    b. kpanâ gɔ́   a   mbɛí
        Kpana give Pr   rice
        'Give the rice to Kpana.'

The only apparent exceptions to the ban on the insertion of more than one nominal term between *S* and *V* are instances of incorporation, as in ex. (7). Note that, in Bambara, incorporation is made apparent by the tonal contour of the sequence *incorporated noun + verb*.

(7)   Bambara

    a. dókótóró ꜜmá   fántà   fúrá-ké
        doctor   PM   Fanta   medicine-do
        'The doctor did not attend Fanta.'

    b. ù   yé   cɛ̀kɔ̀rɔbá   wùlù-fáá
        3P   PM   old man   dog-kill
        'They killed the old man like a dog.'

1.3.   Predicative Markers and Verbal Inflection

Most Mande languages have a very reduced verbal inflection and make wide use of auxiliaries (called predicative markers in many descriptions of Mande languages) that obligatorily follow the subject. For example, as illustrated by ex. (8), Bambara, like the majority of Central Mande languages, marks tense, aspect, and mood by predicative markers following the subject and expressing also the *positive vs. negative* distinction. The only exceptions are the imperative singular positive, which has no marker at all, and the intransitive past/perfective positive, marked by a verbal suffix.

(8)   Bambara

a. woman   PM   go
    mùsó   ꜜbɛ́   táá   'The woman is going.'
    mùsó   ꜜtɛ́   táá   'The woman is not going.'
    mùsó   ꜜtáá-rá   'The woman went/has gone.'
    mùsó   ꜜmá   táá   'The woman didn't go/hasn't gone.'
    mùsó   ꜜbɛ́nà ~ ꜜná   táá   'The woman will go.'
    mùsó   ꜜténà   táá   'The woman will not go.'
    mùsó   ꜜká   táá   'The woman should go.'
    mùsó   kàná   ꜜtáá   'The woman should not go.'

    táá   pl. á yé táá   'Go!'
    kàná ꜜtáá   pl. á kàná ꜜtáá   'Don't go!'

b. woman PM   fish   buy
    mùsó ꜜbɛ́   jɛ́gɛ́   sàn   'The woman is buying fish.'
    mùsó ꜜtɛ́   jɛ́gɛ́   sàn   'The woman is not buying fish.'
    mùsó ꜜyé   jɛ́gɛ́   sàn   'The woman bought/has bought
                                          fish.'
    mùsó ꜜmá   jɛ́gɛ́   sàn   'The woman didn't buy/hasn't
                                          bought fish.'
    mùsó ꜜbɛ́nà~ꜜná   jɛ́gɛ́   sàn   'The woman will buy fish.'
    mùsó ꜜténà   jɛ́gɛ́   sàn   'The woman will not buy fish.'
    mùsó ꜜká   jɛ́gɛ́   sàn   'The woman should buy fish.'
    mùsó kàná   ꜜjɛ́gɛ́   sàn   'The woman should not buy fish.'

    jɛ́gɛ́ sàn   pl. á yé jɛ́gɛ́ sàn   'Buy fish!'
    kàná ꜜjɛ́gɛ́ sàn   pl. á kàná ꜜjɛ́gɛ́ sàn   'Don't buy fish!'

The situation of Mande languages is, however, far from uniform in this respect. On the one hand, Koyaga, a Manding dialect spoken in Ivory Coast, differs from Bambara by using a predicative marker in the intransitive past/perfective positive too, and therefore has no verbal inflection at all. But on the other hand, several Mande languages (for example, Soso) have a relatively developed verbal inflection, and mark several TAM values without making use of predicative markers.

An important characteristic of the Mande systems of predicative markers is that most of them cannot be analyzed as auxiliary verbs. Moreover, for some of them, there is evidence that they developed from other categories than verbs, in particular, from postpositions (see Bird and Kendall 1986, Bearth 1995, Creissels 1997a, Kastenholtz 2003).

It is therefore not entirely correct to characterize Mande constituent order as *S-Aux-O-V-X*, as often proposed. An accurate representation of the Mande pattern not limited to the lexical verb and the nominal terms of its construction should rather be something like *S-neg(tam)-O-V(tam)-X*.

1.4.   No Evidence of a Different Pattern of Constituent Order in Proto-Mande
Claudi (1994) claims that, originally, Mande languages had the *S-V-O-X* order at clause level, but the order *G-N* (*genitival dependent + head noun*) in the noun phrase, and that the *S-O-V-X* order is an innovation resulting from the reanalysis of constructions of the type *auxiliary + nominalized verb*, in which the NP that would have constituted the object of a finite form of the nominalized verb was treated as a genitival dependent. This is undoubtedly a possible scenario, which considerably weakens previous claims according to which Mande constituent order provides evidence that the constituent order of Proto-Niger-Congo was *S-O-V*. However, Claudi's proposal is entirely speculative, since

(a) in Mande languages, the uniformity of word order and constituent order patterns is total, and therefore cannot provide the slightest evidence of the previous existence of a constituent order other than *S-O-V-X*, and

(b) the auxiliarization processes postulated by Claudi can also operate within the frame of a constituent pattern identical with that of present-day Mande languages without inducing any change in the linearization rules.

Moreover, other equally plausible grammaticalization processes can lead to the same shift from *S-V-O-X* to *S-O-V-X*, for example the replacement of a transitive construction conforming to the pattern *S-V-O-X* by a serial verb construction *S take O V X*, followed by the decategorialization of *take*, a process widely attested for example among Kwa languages.[2]

In other words, internal evidence from Mande languages leaves us with the default hypothesis that Proto-Mande had word order patterns identical to those attested in the modern Mande languages. To what extent this can be considered as an argument for reconstructing *S-O-V-X* at Proto-Niger-Congo level, as suggested in several recent studies,[3] is another question, to which we will return after examining the constituent patterns of other West African languages currently viewed as having a constituent order pattern of the Mande type in at least certain conditions.

2.   Non-Mande Languages with a Rigid *S-O-V-X* Constituent Order
Non-Mande languages with exactly the same pattern of constituent order as Mande are not very numerous. This situation seems to be restricted to languages spoken in areas shared with Mande languages. Senufo languages constitute the best known case. Carlson (1994) provides a detailed and precise description of the morphosyntax of a Senufo language, which shows that this language shares with Mande not only the absolute rigidity of the *S-O-V-X* constituent order, but also the

---

[2] See Lord (1993). Given the universal affinity between the syntactic notions of subject and object and the discursive notions of topic and focus, another possible scenario is the syntacticization of a pragmatically driven constituent order *T-F-V-others* (T=topic, F=focus) of the type attested for example in Basque or in Hungarian. An additional argument against Claudi's hypothesis is that if the emergence of the *S-O-V-X* constituent order could be the mere consequence of auxiliarization in *S-V-O-X* languages in which the genitival dependent precedes its head, the Mande pattern should be common elsewhere in the world, since *S-V-O-X* languages anteposing the genitival dependent are not rare, and auxiliarization processes are universal.
[3] See Gensler (1994), Gensler (1997), Gensler & Güldemann (2003).

impossibility of inserting more than one nominal term between *S* and *V*. As in Mande, even the most typical ditransitive verbs cannot be constructed with two nominal terms inserted between S and V (9a-b), and whenever the second argument of a transitive verb can be represented by a nominal term in postverbal position, the presence of a postposition indicates that this is an instance of a valency alternation of the antipassive type, whereby the object has been demoted to oblique (9c-d).[4]

(9)   Supyire (Carlson 1994, Carlson 2000)

     a. mìi   a   sém̄-píi   cyɛ̀ɛ̀   ali   na
        1S   PM   papers-DEF   show   Ali   Po
        'I have shown the papers to Ali.'

     b. mìi   a   ù   kàn   nùjìrìmɛ̀   na
        1S   PM   3S   give   milk   Po
        'I have given him/her milk.'

     c. u   a   m̀pàa-bíi   bɔ̀   (S_A O_P V)
        3S   PM   sheep.PL-DEF   kill
        'He/she has killed the sheep (pl.).'

     d. u   a   bo   m̀pàa-bíl-ê   (S_A V X_P)
        3S   PM   kill sheep.PL-DEF-Po
        'He/she has killed some of the sheep.'

3.   The Constituent Order Pattern of Eastern Songhay

In Eastern Songhay, illustrated here by Zarma and Gao Songhay,[5] the *S-V-O-X* constituent order (which is the only possible constituent order in Western Songhay and in Northern Songhay) has a marginal status. For typical transitive verbs, the *S-O-V-X* order is either the only possible order (in Gao Songhay), or the preferred order (in Zarma). As in Mande, in clauses with a constituent order of the *S-O-V-X* type, no more than one nominal term can be inserted between *S* and *V*, as illustrated by ex. (10).

(10)   Zarma (Oumarou Yaro 1993)

     a. ábdù   nà   fèèjì   wíi   yàwòó   sɛ̀
        Abdou   PM   sheep   kill   guest.DEF Po
        'Abdou has killed a sheep for the guest.'

---

[4] As discussed in Carlson (2000), the possible meanings of the antipassive construction of Supyire are "partially affected undergoer" (as in the example reproduced here) and "participatory agent."
[5] On Zarma, see Oumarou Yaro (1993). On Gao Songhay, see Heath (1999).

b. múúsà nà ŋ́gà mòótàà nóó káynòó sè
   Moussa PM 3S car.DEF give younger brother.DEF Po
   'Moussa has given his car to his younger brother.'

Apart from the fact that the *S-O-V-X* pattern of Eastern Songhay is virtually identical to that of Mande, a particularity that sharply distinguishes Eastern Songhay from Mande is the existence of a limited class of semantically bivalent verbs whose second argument must occur in postverbal position, but shows no evidence of an oblique status, as illustrated by ex. (11). Oumarou Yaro argues that the behavioral properties of the second argument of such verbs do not differ from those of the second argument or prototypical transitive verbs, and that consequently it must be recognized as a variety of object.[6] The same analysis is proposed by Heath for Gao Songhay.

(11)   Zarma (Oumarou Yaro 1993)

a. ábdù gá hìmá bààbò / *ábdù gá bààbóó hìmà
   Abdou PM resembles father.DEF
   'Abdou resembles his father.'

b. ábdù gà báá hǎysà / *ábdù gá hǎysà bâ
   Abdou PM love Aïssa
   'Abdou loves Aïssa.'

c. ábdù díí zànkày / *ábdù nà zànkày dí
   Abdou see child.PL.DEF
   'Abdou saw the children.'

Moreover, in Zarma (but apparently not in Gao Songhay), the *S-O-V-X* constituent order is not the only possible constituent order in clauses headed by a prototypical transitive verb. In Zarma, with prototypical transitive verbs, the *S-V-O-X* order is not frequent, but it is possible, and without any apparent conditioning, as illustrated by ex. (12).

(12)   Zarma (Oumarou Yaro 1993)

a. à nà gòrŋòò wíí yàwòó sè
   3S PM chicken.DEF kill guest.DEF Po
   'He has killed the chicken for the guest.'

---

[6] Oumarou Yaro (1993) gives the following list of transitive verbs whose object cannot occur in preverbal position: *máá* 'hear', 'feel', *díí* 'see', *dòòná* 'be accostumed to', *dùù* 'get', 'have', *hín* 'surpass', *hìmá* 'resemble', *màànù* 'approach', *báá* 'like', *wáání* 'know'.

b. à   wíí   gòrŋòó   yàwòó   sè [7]
   3S   kill chicken.DEF guest.DEF Po
   'He has killed the chicken for the guest.'

In Zarma, contrary to Mande or Senufo, this variation between two possible positions of a term representing the second argument or prototypical transitive verbs seems to be nothing more than a free variation in the linearization of the transitive construction, since no postposition marks the second argument of prototypical transitive verbs in postverbal position, and more generally, there seems to be no evidence that this variation should be analyzed as the result of a construction change of the antipassive type. Oumarou Yaro explicitly states that nothing in his observations or in his native speaker intuitions suggests any explanation of this variation.

Heath signals an interesting particularity of the constituent order pattern of Gao Songhay which is related to the existence of postverbal objects, and therefore has no equivalent in Mande syntax. As already indicated, Eastern Songhay, like Mande, forbids the insertion of two non-conjoined NPs between the subject and the verb. But the existence of a second position for objects (to the right of the verb) results in the possibility of double object constructions in which one of the two objects follows the verb, and this is precisely what happens with the verb *noo* 'give'. In Gao Songhay, this verb has two possible constructions: either the thing given is encoded as a preverbal object, and the recipient as an oblique, as in ex. (13a), or the recipient is encoded as a preverbal object, and the thing given as a postverbal object, as in ex. (13b).

(13)   Gao Songhay (Heath 1999)

a.   a   na   atteyoo   noo   yane
      3S   PM   tea.DEF   give 1S.DAT
      'She gave me the tea.'

b.   a   na   ey   noo   atteyoo
      3S   PM   1S   give tea
      'She gave the tea to me.'

4.   Other West African Languages Showing Alternations in the Constituent Order
4.1.   General Remarks

We now turn to the case of West African languages which share with Zarma the existence of alternations in the constituent order involving *S-V-O-X* as one of the two alternative orders, but in which the pattern that alternates with *S-V-O-X* is not entirely identical with (and sometimes very different from) the *S-O-V-X* pattern of

---

[7] Note that, in Zarma, the predicative marker of the perfective positive occurs only in the *S O V X* construction, not when the verb immediately follows the subject. The other predicative markers are not sensitive to this distinction.

Mande. Such alternations are a widespread phenomenon in the Gur, Kwa, and Kru families, and are attested in some Atlantic languages too. I will argue in section 5 that recent studies tend to overestimate their frequency. But before examining this question, what I would like to show first is that, in contrast with the situation found in West African languages that invariably put the object in preverbal position, the constituent order patterns of the languages examined in this section are not uniform. They have in common that the variation (in contrast with the situation observed in Eastern Songhay) is conditioned by TAM or polarity: the use of a constituent pattern other than *S-V-O-X* is commonly restricted to clause types characterized by the presence of overt predicative markers immediately after the subject. But the details of the conditioning greatly differ from one language to another, and no generalization is possible concerning the TAM and polarity values that trigger a constituent order other than *S-V-O-X*. There are also important differences in the range of nominal terms involved in the alternation, with the result that treating them indistinctly as intances of a variation between the canonical *S-V-O* pattern and the Mande pattern implies some dose of oversimplification.

The three cases examined below are not intended to give a comprehensive view of the question, but only to illustrate the heterogeneity of the constituent order patterns found among West African languages, and to emphasize the need to gather more detailed and more precise information on this matter before any serious attempt is made to establish a detailed typology of the constituent order patterns found in West Africa.

4.2. Kisi

Kisi,[8] an Atlantic language spoken in Sierra Leone and Guinea, illustrates an alternation in constituent order triggered by the presence of an auxiliary immediately after the subject, in which the pattern alternating with *S-V-O-X* differs from the Mande type on the following points:

First, Kisi has double object constructions in which two (or even three[9]) nominal terms take part in the alternation, as in ex. (14).

(14)   Kisi (Childs 2003)

a.   ò   ké   yá   tòòlúláŋ
       3S   give 1S   support
       'She gave me support.'

b.   à   wá   ndú   kòówáŋ   kìóó
       3P   PM   3S   medicine give
       'They were giving him medicine.'

---

[8] Childs (1995).

[9] In Kisi, constructions with three objects are possible with the applicative form of ditransitive verbs.

Second, in Kisi, adpositional phrases are not involved in the alternation: as a rule, they remain in postverbal position. Similarly, in Kisi, locative arguments are not considered objects, and invariably remain in postverbal position. However, the alternation is not limited to objects in typical multiple object constructions: time adverbs can occur between the predicative marker and the verb too, as in ex. (15).

(15)   Kisi (Childs 2003)

    ò   cò   nĩŋ yá   mààlóŋ   hùŋgùlló
    3S PM now 1S   rice   beat.APPL
    'He is beating the rice for me now.'

The differences with the Mande pattern are striking, more especially as Kisi is surrounded by Mande languages and is considered to have been influenced by Mande languages: the fact that Kisi puts in preverbal position terms that Mande languages would put in postverbal position is difficult to reconcile with the hypothesis of a development under Mande influence. Since Kisi, like other Atlantic languages, has the order *N-G* in noun phrases, Claudi's reanalysis scenario is not available either as the explanation of a possible independent shift from *S-V-O-X* to the pattern attested in clauses including an auxiliary.

4.3.   Attie

Attie,[10] a Kwa language spoken in Ivory Coast, attests a situation in which the range of terms involved in the constituent order alternation includes not only the two objects of fairly typical double object constructions, but also locative terms whose status as objects or adjuncts is not entirely clear. For example, in sentence (16a), the two objects precede the verb, but the locative term 'at the market' follows it. By contrast, in sentence (16b), the locative term 'in the forest' precedes the verb. The solution proposed by Kouadio is that in Attie, oblique arguments take part in the alternation in the same way as objects, whereas adjuncts invariably remain in postverbal position. A more detailed description of Attie syntax would however be necessary in order to evaluate this hypothesis.

(16)   Attie (Kouadio 1996)

   a. mɛ̄ ɟi-ī   jàpí ʃìkā   dzé jábò   lɵ̄
       1S father-PM Yapi money give market there
       'My father is giving money to Yapi at the market.'

   b. jàpí-ī   kpɔ̄ɛ̂   pjà̄ nɵ̄ɛ̄
       Yapi-PM forest.DEF   in   walk
       'Yapi is walking in the forest.'

---

[10] Kouadio (1996).

4.4. Neyo

Neyo[11] is an Eastern Kru language spoken in Ivory Coast.

Alternations between *S-V-O-X* and a constituent order in which the object precedes the verb are general in Kru languages, and they are conditioned by the presence of an auxiliary inserted immediately after the subject. It is also general in Kru languages that, much in the same way as in Kisi or in Attie, and in contrast to the situation observed in Mande, the alternation may involve more than one nominal term, and is not restricted to objects. The available data suggests that, at least in some Eastern Kru languages, the alternative constituent pattern can be characterized as *S-O-X-V*, i.e. as verb-final, but with an interesting particularity: in most verb-final languages, the default position of the object is immediately before the verb. By contrast, in Eastern Kru languages, the final position of the verb does not seem to affect the relative order *O-X*. In her description of Neyo, Grah explicitly states that *S-O-X-V* is the canonical constituent order triggered by a set of six predicative markers. She adds, however, that the relative ordering of objects and obliques is not rigid, and that in verb-final clauses, obliques may precede the object, or even occur in postverbal position (17).

(17)   Neyo (Grah 1983)

    a. kónî nî   sáká   jàlé   ĺi
        Koni PM   rice   kitchen eat
        'Koni has not eaten rice in the kitchen.'

    b. làĺi yā   mágìtī kú   ĺiēplū yé
        Lali PM   market Po   scarf   see
        'Lali has seen a scarf at the market.'

    b. kónî níkā ɲú   mlā   zīmlē̄
        Koni PM   water drink   today
        'Koni will not drink water today.'

5.   Nominal and Pronominal Objects

I would like to react now against a tendency to treat the position of pronominal objects on a par with the position of object NPs in constituent order studies concerning the Niger-Congo family. This confusion leads in particular to a drastic overestimation of the importance of the Mande type of constituent order. In the languages of the world, weak object pronouns prefixed to verbs are extremely common in otherwise robust *S-V-O* languages. Therefore, there is nothing particularly strange in the fact that so many Niger-Congo languages are strict *S-V-O* languages as regards the ordering of NPs, but have weak object pronouns prefixed to verbs. In other words, it is contradictory to present the Mande type of word order as an exotic constituent order pattern, virtually non-existent outside Africa, and to use at the same time data from languages in which pronominal

---

[11] Grah (1983).

objects only precede the verb to demonstrate the alleged pervasiveness of this type of constituent order among Niger-Congo languages.

This has important consequences concerning the use of typological data in reconstruction, since proposals to reconstruct a constituent order with the object before the verb (either canonical *S-O-V*, or a pattern of the Mande type) in Proto-Niger-Congo crucially rely on the assumption that, when pronominal objects do not occur in the same position as object NPs, the position of pronominal objects can be assumed to reflect the position of object NPs in an ancient state of the language.

In languages with a flexible constituent order, the default position of pronominal objects may be different from that of object NPs, due to their high degree of inherent topicality. For example, in Russian, a language with a particularly flexible constituent order, the less marked order is very clearly *S-V-O* with object NPs, but rather *S-O-V* with personal pronouns in the role of object.

Morevover, pronouns tend to cliticize, and it is well known that prosodic factors independent from constituent order typology are crucial for the evolutions in the positioning of clitics. For example, second position clitics are not restricted to languages having a particular pattern of constituent order. Consequently, nothing ensures that the position occupied by object clitics or affixes at some stage in the evolution of a language should reflect the position occupied by object NPs at a more ancient stage. The well-known and often-quoted slogan "Today's morphology is yesterday's syntax" certainly does not hold for pronominal affixes.

For example, in modern Romance languages, object NPs invariably follow the verb, but in most of them (Portuguese being the main exception), weak object pronouns invariably attach to the left of finite verb forms. In French, this rule of left-attachement extends to non-finite verb forms, and the imperative positive is the only exception. Following argumentation of the type developed in attempts to reconstruct the constituent order of Proto-Niger-Congo that take the position of object clitics or affixes as evidence of the position previously occupied by object NPs, a constituent order of the Mande type should be reconstructed in Proto-Romance, which is certainly not correct. Latin had a flexible constituent order with the verb in final position as the less marked option, and modern Romance languages have more or less flexible patterns of constituent order with a clear predominance of *S-V-O-X*, but there is no evidence that *S-O-V-X* played a role as an intermediate stage in the shift from the Latin pattern of constituent order to that of modern Romance languages. Moreover, the history of Romance languages is well documented enough to establish that the position of pronominal objects in modern Romance languages results from evolutions that cannot be characterized as the maintenance of the position occupied by object NPs at some stage in the history of Romance languages.

6. Conclusion

In the attempts to reconstruct the history of constituent order patterns in the West African language families belonging to the Niger-Congo phylum, the only hypothesis relying on a firm empirical basis is that Proto-Mande already had a constituent pattern of the type attested by modern Mande languages. By contrast,

the historical interpretation of the similarities and differences between the Mande constituent order pattern and the alternating constituent order patterns found in other West African languages families is far from clear.

In my opinion, most studies dealing with this question have greatly underestimated the differences between the Mande constituent order pattern and the alternating patterns found in Kwa, Gur, Kru, and Atlantic languages. In section 4, I have tried to show that the differences are of two types that, from a historical point of view, seem to point in opposite directions:

On the one hand, the absolute rigidity of the position of the object in the Mande pattern suggests a Mande influence in the diffusion or maintenance of constituent order patterns in which the object can be inserted between the subject and the verb;

On the other hand, in West African languages with alternating constituent order patterns, the range of nominal terms that can be inserted between the object and the verb is considerably wider than in Mande languages, which leads one to reject the hypothesis of an areal diffusion from Mande languages, and casts serious doubts on the reconstruction of a constituent order pattern of the Mande type in Proto-Niger-Congo too, since it is difficult to explain why so many West African languages would have at the same time shifted to *S-V-O* in certain conditions, and widened the range of nominal terms inserted between the subject and the verb in other conditions.

I have no solution to propose to this puzzle, but I think it is important to emphasize that a fine-grained typology of constituent order patterns in West Africa does not confirm the current view according to which, in languages with alternating constituent order patterns, the variant with the object between the subject and the verb can be identified with the Mande type of constituent order.

References

Bearth, T. 1995. Nominal periphrasis and the origin of the predicative marker in Mande languages: an alternative view. *Afrikanistische Arbeitspapiere* 41:89-117.

Bird, C. and M. Kendall. 1986. Postpositions and auxiliaries in Northern Mande: Syntactic indeterminacy and linguistic analysis. *Anthropological Linguistics* 28(4):389-403.

Carlson, R. 1991. Grammaticalization of postpositions and word order in Senufo languages. In E. C. Traugott & B. Heine, eds., *Approaches to Grammaticalization*. Vol. 2, 201-223. Amsterdam/Philadelphia: Benjamins.

Carlson, R. 1994. *A Grammar of Supyire*. Berlin: Mouton de Gruyter.

Childs, T. 1995. *A Grammar of Kisi*. Berlin: Mouton de Gruyter.

Childs, T. 2003. It's everywhere! More on the S-Aux-O-V-Other in Atlantic. Paper presented at CALL 33, University of Leiden.

Claudi, U. 1994. Word order change as category change: the Mande case. In W. Pagliuca & G. Davis, eds., *Perspectives on Grammaticalization*, 201-241. Amsterdam: Benjamins.

Creissels, D. 1997a. Postpositions as a possible origin of certain predicative markers in Mande. *Afrikanistische Arbeitspapiere* 50:5-17.

Creissels, D. 1997b. Une tentative d'explication de particularités typologiques de la négation en mandingue. *Mandenkan* 32:3-22.

Gensler, O. 1994. On reconstructing the syntagm S-Aux-O-V-Other to Proto-Niger-Congo. In K. E. Moore, D. A. Peterson & C. Wentum, eds., *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistic Society, February 18-21, Special Session on Historical Issues in African Linguistics*, 1-20. Berkeley: BLS.

Gensler, O. 1997. Grammaticalization, typology, and Niger-Congo word order: Progress on a still-unsolved problem. *Journal of African Languages and Linguistics* 18-1. 57-93.

Gensler, O. and T. Güldemann. 2003. *S-Aux-O-V-Other in Africa: Typological and areal perspective*. Paper presented at WOCAL 4. New Brunswick, NJ: Rutgers University.

Grah, C. 1983. *Approche systématique du nı́wɔ̄lı̄*. University of Paris.

Heath, J. 1999. *A grammar of Koyraboro (Koroboro) Senni*. Köln: Rüdiger Köppe.

Heine, B. 1976. *A Typology of African Languages, Based on the Order of Meaningful Elements*, Berlin: Reimer.

Heine, B. and M. Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.

Kastenholz, R. 2003. Auxiliaries, grammaticalization, and word order in Mande. *Journal of African Languages and Linguistics* 24(1):31-53.

Kouadio, N. J. 1996. *Description systématique de l'attié de Memni*. Thèse de doctorat d'état. University of Grenoble.

Lord, C. 1993. *Historical Change in Serial Verb Constructions*. Amsterdam/Philadelphia: Benjamins.

Oumarou Yaro, B. 1993. *Eléments de description du zarma*. Ph.D. thesis, University of Grenoble.
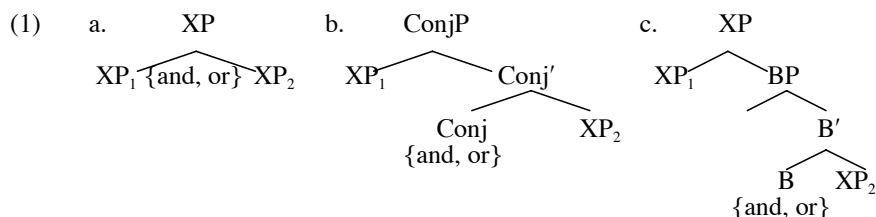
denis.creissels@univ-lyon2.fr

Nupe Coordinate Structures: A Syntactically Heterogeneous Class

JASON KANDYBOWICZ
University of California, Los Angeles

0.   Introduction: The Many Flavors of Nupe Coordination

In the wake of Kayne (1984, 1994), the traditional flat structures associated with coordinate constructions (1a) have been widely reanalyzed as asymmetric binary branching structures (1b,c). Conceptually speaking, flat representations such as (1a) are non-endocentric and thus do not conform to X-bar theory. Empirically, one can show that the coordinator and the second coordinate form a constituent (Ross 1967:90-91), which is asymmetrically c-commanded by the preceding coordinate (Munn 1993). Regarding the asymmetric structure of coordination, there are two leading analyses, both of which are held to regulate all varieties of coordinate constructions within and across languages. According to Kayne (1994), Johannessen (1998), and Zoerner (1999), the coordinator projects a phrase whose specifier is filled by the initial coordinate and whose complement hosts the second in cases of binary coordination (cf. (1b)). Munn (1993), Büring and Hartmann (1998), and Velde (2000) analyze the coordinator as heading a functional projection (Boolean phrase (BP) in Munn 1993), which adjoins to the initial coordinate and takes the second coordinate as its complement (cf. (1c)).

(1)   a.   XP   b.   ConjP   c.   XP

XP_1 {and, or} XP_2 XP_1 Conj′ XP_1 BP

Conj XP_2 B′
{and, or}

B XP_2
{and, or}

Does coordination in every language exclusively pattern with respect to one of these analyses or is the phenomenon syntactically heterogeneous? In this paper, we restrict ourselves to the phenomenon of conjunction in Nupe (Niger-Congo: Kwa) and show that it is not reducible to a single homogeneous syntactic analysis.

As in a number of languages, the identity of the coordinating morpheme in Nupe is determined by the size/categorical status of the conjoined constituents. As such, conjunction comes in three flavors in the language: phrases outside the extended projection of the verb (e.g. DP, PP, AP) are conjoined by the morpheme *tò* (2a-c); verb phrase conjunction proceeds via covert coordination (2d); and the conjunction of clausal constituents (TP, CP) is mediated either by *ma* or *ci* (2e,f).

(2)   a.   Musa si [$_{DP}$bise   nana] tò   [$_{DP}$nakàn sási].
           Musa buy chicken this   CONJ   meat some
           'Musa bought this chicken and some meat.'
      b.   Musa ba nakàn [$_{PP}$bè   foki yin] tò [$_{PP}$bè   èbi   yin].
           Musa cut meat   with fork PRT CONJ with knife PRT
           'Musa cut the meat with a fork and with a knife.'
      c.   Musa pa   eci [$_{AP}$dzúrú tò bókùn] zì [$_{AP}$yína tò   tsúwó].
           Musa pound yam red   CONJ white PL   today CONJ yesterday
           'Musa pounded the red and white yams today and yesterday.'
      d.   Eza   guba [$_{VP}$ba nakàn] [$_{VP}$lo dzukó] à.
           person two   cut meat   go market NEG
           'Two people didn't (both) cut the meat and go to the market.'
      e.   [$_{TP}$Musa à   ba   nakàn] [$_{TP}$ Gàná ma/ci à   gi eci].
             Musa FUT cut meat   Gana CONJ FUT eat yam
           'Musa will cut the meat and Gana will eat the yam.'
      f.   [$_{CP}$Ze ba nakàn o] [$_{CP}$ke   ma/*ci Gàná gi o]?
             who cut meat FOC what CONJ   Gana eat FOC
           'Who cut the meat and what did Gana eat?'

This paper argues that the proper analysis of the conjunction of smaller categories such as DP, PP, and AP in Nupe is as in Kayne (1994)/Johannessen (1998) and others (cf. (1b)), whereas the syntax of clausal conjunction involves adjunction as in Munn (1993), among others (cf. (1c)). We argue that covert coordination of verb phrases represents a third type of coordination in the language, despite the fact that it admits neither structural analysis. These constructions are shown to be syntactically akin to serial verb constructions, where the coordination takes place in the semantics rather than in the syntax. Nupe thus provides evidence that coordination is not amenable to a unitary syntactic analysis and thus that efforts to reduce all instances of it to a single clausal variety are misguided.

1.   Small Category Conjunction
1.1.   Basic Properties
As shown in (2a-c), the morpheme *tò* is used to conjoin constituents below the VP level, namely, DPs, PPs, and APs. The particle may not coordinate verb phrases (3a) or clauses (3b), nor can the clausal conjunctions *ma* and *ci* combine small categories in the language (3c). Coordination of DPs, PPs, or APs may not be

covert as in verb phrase coordination (3d). Lastly, the order of the conjuncts is reversible in the case of *tò* coordination (cf. (2a) and (3e)).

(3)　　a.　　*Musa à [$_{VP}$ba nakàn] tò [$_{VP}$gi eci].
　　　　　　　Musa FUT cut meat CONJ eat yam
　　　　b.　　*[$_{TP}$Musa à ba nakàn] [$_{TP}$Gàná tò à gi eci].
　　　　　　　　Musa FUT cut meat Gana CONJ FUT eat yam
　　　　　　　Also BAD as: *[Musa à ba nakàn] tò [Gana à gi eci].
　　　　c.　　*Musa si [$_{DP}$bise nana] ma/ci [$_{DP}$nakàn sási].
　　　　　　　Musa buy chicken this CONJ meat some
　　　　　　　Also BAD as: *Musa si [bise nana] [nakàn ma/ci sási].
　　　　d.　　*Musa si [$_{DP}$bise nana] [$_{DP}$nakàn sási].
　　　　　　　Musa buy chicken this meat some
　　　　e.　　Musa si [$_{DP}$nakàn sási] tò [$_{DP}$bise nana].
　　　　　　　Musa buy meat some CONJ chicken this
　　　　　　　'Musa bought some meat and this chicken.'

1.2.　　Syntax of Nupe *tò* Conjunction

The proposals in (1b) and (1c) make a number of parallel claims and predictions. For instance, in line with the Linear Correspondence Axiom (LCA - Kayne 1994), they both claim that the first conjunct asymmetrically c-commands the following constituent, given the precedence relation that holds between them. The truth of this claim is easily substantiated in Nupe, given binding facts. Bound variable readings are possible when the quantified DP appears in the first conjunct, but not when in the second.

(4)　　a.　　[Bagi ndondo]$_i$ tò [egi u$_i$] dzò eyì.
　　　　　　　man every CONJ child 3rd.SG plant corn
　　　　　　　'Every man$_i$ and his$_i$ child planted corn.'
　　　　b.　　[Egi u$_i$] tò [bagi ndondo]$_{*i/j}$ dzò eyì.
　　　　　　　child 3rd.SG CONJ man every plant corn
　　　　　　　'His$_i$ child and every man$_{*i/j}$ planted corn.'

The second commonality between the two approaches is that they both predict the impossibility of coordinating heads. This prediction is clearly borne out in Nupe.

(5)　　a.　　*Musa yébó bagi [nana] tò [wuncìn]
　　　　　　　Musa love man this CONJ that
　　　　　　　*'Musa loves this and that man.'
　　　　　　　OK as: Musa yébó [bagi nana] tò [bagi wuncìn].
　　　　b.　　*[nda] tò [nna] Musa
　　　　　　　father CONJ mother Musa
　　　　　　　*'the mother and father of Musa'
　　　　　　　OK as: [nda Musa] tò [nna Musa]

Both approaches also predict the reversibility of the conjuncts, given that neither conjunct contains the other structurally.

There is a direct syntactic reason for rejecting the structure in (1c) in favor of a Kaynian theory of small category conjunction in Nupe. The argument comes from patterns of agreement with coordinated subjects. In a limited range of constructions in the language (e.g. subjunctives and optatives), reduced pronominal elements pro-cliticize to embedded verbs (Smith 1967:9). These clitics agree with the φ-features (person and number) of the subject. This is illustrated below with singular subjects.

(6)   a.   **Mi** è wá cènkafa **n-** gi.
           1st.SG PRES want rice   1st.SG-eat
           'I want to eat rice.'

      b.   **Wo** è wá cènkafa **o-** gi.
            2nd.SG PRES want rice   2nd.SG-eat
           'You want to eat rice.'

      c.   **Wun** è   wá cènkafa **u-** gi.
           3nd.SG PRES want rice   3rd.SG- eat
           'S/he wants to eat rice.'

In many languages, agreement is triggered by elements in the first conjunct, but not in the second (Munn 1993, Johannessen 1998).

(7)   a.   There are [three unicorns] and [a man] in the garden.
      b.   *There is [three unicorns] and [a man] in the garden.
      c.   There is [a man] and [three unicorns] in the garden.
      d.   ?There are [a man] and [three unicorns] in the garden.

Munn (1993) argued that the second coordinate is invisible for purposes of agreement because it is not generated in an A-position. We can determine whether or not the second conjunct in Nupe is an adjunct or an argument on the basis of the identity of the agreement marker that surfaces when DP subjects with conflicting φ-features are conjoined. The following data involving conjoined first and third person DPs show that there is no asymmetry with regard to the conjunct that triggers agreement in Nupe. Unless one of the conjuncts is a quantified DP (8c,d), first person agreement is triggered regardless of whether the first person pronominal element is generated in the first conjunct or the second.

(8)   a.   [Mi tò wun] è wá cènkafa **n/** *u- gi.
           1st.SG CONJ 3nd.SG PRES want rice   1st.SG/*3rd.SG-eat
           'Me and her/him want to eat rice.'

      b.   [Wun tò mi] è wá cènkafa *u/ **n-** gi.
           3rd.SG CONJ 1st.SG PRES want rice   *3rd.SG/1st.SG-eat
           'S/he and I want to eat rice.'

    c.   [Mi   tò   eza   ndondo] è   wá   cènkafa *n/   **u**-   gi.
          1st.SG CONJ person every   PRES want rice   *1st.SG /3rd.SG-eat
          'I, along with everyone, want to eat rice.'

    d.   [Eza   ndondo tò   mi] è   wá   cènkafa **u**/   *n-   gi.
          person every   CONJ 1st.SG PRES want rice   3rd.SG/*1st.SG-eat
          'Everyone and I want to eat rice.'

Thus, the φ-features of *both* DP conjuncts must be syntactically visible for purposes of agreement computation. This argues against the adjunct analysis represented in (1c) because the φ-features of the adjunct-internal DP conjunct cannot percolate up outside the adjunct to the maximal projection containing both conjuncts. The reason for this is that feature movement, like constituent movement, is island-sensitive. Hence, only those φ-features of the first conjunct should be visible for agreement under Munn's analysis.

2. Clausal Conjunction
2.1. Basic Properties

Both root TPs and CPs can be conjoined using *ma* (9). In both cases, the conjunction head is a second-position particle that is preceded by a nominal element, typically the subject (cf. Latin and Somali). Once again, the linear order of the conjoined constituents is reversible (9b,c).

(9)   a.   Musa à   du   nakàn_i Gàná ma à   gi   wun_i.
          Musa FUT cook meat Gana CONJ FUT eat 3rd.SG
          'Musa will cook the meat and Gana will eat it.'
          BAD as: *Musa à du nakàn **ma Gana** à gi wun.

      b.   Ze   du   nakàn o   ke   ma Gàná gi o.
          who cook meat FOC what CONJ Gana eat FOC
          'Who cooked the meat and what did Gana eat?'

      c.   Ke   Gàná gi o   ze   ma   du   nakàn o.
          what Gana eat FOC who CONJ cook meat FOC
          'What did Gana eat and who cooked the meat?'

However, although root CPs may be conjoined (9b,c), embedded CPs may not.

(10)   *Musa kpe [ganan Gàná bé] mi   ma [ganan ba nakàn].
      Musa know that   Gana come 1st.SG CONJ that   cut meat
      *'Musa knows (two things) that Gana came and that I cut the meat.'
      Also BAD as: *Musa kpe [ganan Gàná bé] **ma** [ganan **mi** ba nakàn].
      OK as: Musa_i kpe ganan Gàná bé wun_i ma kpe ganan mi ba nakàn.

    A variant of *ma* exists in the language that is used to conjoin clauses and relate them temporally. When two sentences are conjoined by *ci*, the eventuality denoted by the second conjunct is understood to hold at a time posterior to that of

the first conjunct (11a); otherwise, the resulting sentence is semantically ill-formed (11b). The temporal interpretation of the second conjunct in this construction is thus parasitic on the prior coordinate, unlike the case of *ma* conjunction. Similar facts hold in Maasai (Caponigro 2003), Korean, and Turkish.

(11) a. Musa à   ba nakàn yínna Gàná ci   à   gi eci   èsun.
          Musa FUT cut meat today Gana CONJ FUT eat yam tomorrow
          'Musa will cut the meat today and Gana will eat the yam tomorrow.'
          BAD as: *Musa à ba nakàn yínna **ci Gàná** à gi eci èsun.
      b. #Musa à   ba nakàn èsun   Gàná ci   à   gi eci yínna.
           Musa FUT cut meat tomorrow Gana CONJ FUT eat yam today
          #'Musa will cut the meat tomorrow and Gana will eat the yam today.'
          (Compare with: ✓Musa à ba nakàn èsun Gàná **ma** à gi eci yínna.)

The *ci* particle is thus similar to English *then*. Semantically, it is a function from pairs (or n-tuples) of events to temporal orderings. When conjoining exactly two TPs, for instance, its denotation is given by the following.

(12) $[[ci]] = \lambda e.\lambda e'[(e \wedge e') \wedge (e' < e)]$

As such, the arguments of *ci* are events/event variables (e.g. TPs). The inability of *ci* to coordinate constituents smaller than TP (cf. (3c)) immediately follows from the fact that event variables in such constituents are unbound and thus outside its scope. Because propositions also lie outside the scope of *ci*, the particle cannot conjoin CP constituents (13), unlike *ma* (cf. (9b,c)).

(13)   *Ze du   nakàn o   ke   ci   Gàná gi o.
         who cook meat FOC what CONJ Gana eat FOC
        *'Who cooked the meat and (then) what did Gana eat?'

## 2.2.   Syntax of Nupe Clausal Conjunction
Many of the syntactic properties of Nupe clausal conjunction fall into place once we consider the syntax of the *ma* and *ci* heads. That these heads are obligatorily preceded by a nominal expression (cf. (9a), (11a)) is not unique in the language. Infinitival *yin* and conditional *gá* impose a similar requirement.

(14) a. Musa yá   nakàn **yin** ba.
          Musa begin meat PRT cut
          'Musa began to cut the meat.'
          BAD as: *Musa yá **yin** ba **nakàn**.
      b. Musa **gá**   è   ba nakàn, Gàná à   pa   eci.
          Musa COND PRES cut meat Gana FUT pound yam
          'If Musa is cutting the meat, then Gana will pound the yam.'
          BAD as: ***Gá Musa** è ba nakàn, Gàná à pa eci.

One analysis that immediately suggests itself is that these heads bear [+D] EPP features that trigger the movement of a nominal to their specifier, in accordance with economy principles such as Shortest Move (Chomsky 1995). In fact, Kandybowicz (in prep) argues that all functional heads in Nupe bear generalized EPP features which (for the most part) guarantee that the specifier position of each functional head is filled by overt material at some point in the derivation. Suppose this is the correct analysis of the *ma* and *ci* heads. We then have an explanation for why clausal conjunctions are always preceded by subject DPs and why *ma* cannot conjoin embedded clauses. Because EPP features are uninterpretable and the attraction of subjects yields a shorter movement chain than the attraction of objects, the derivation of a clausal conjunction construction will not converge unless the coordinator is immediately preceded by the subject DP. Similarly, the derivation of a sentence in which embedded clauses are conjoined will not converge because the EPP features of *ma* will go unchecked. In this case, due to the fact that the second conjunct is a strong phase (CP), the only elements accessible to operations outside that phase (e.g. to the EPP-induced attraction of *ma*) are those syntactic occurrences in C^0 and Spec,CP by the Phase Impenetrability Condition (Chomsky 2001). Since Spec,CP is phonetically empty, the EPP features of *ma* cannot be eliminated and the derivation crashes.

 With this analysis of the clausal conjunction heads in place, we can mount decisive arguments for a Munn-style analysis of Nupe clausal conjunction (1c/15b) over the Kayne/Johannessen approach (1b/15a). The first argument against (1b/15a) is conceptual; adopting such an analysis will force several stipulations which are otherwise unmotivated by the facts. Unlike the representation in (1b/15a), the structure in (1c/15b) makes available a position for the nominal expression to move into in order to check the EPP features of the conjunction head, namely Spec, BP. Because TP_1 occupies the specifier of *ma* under the analysis in (1b/15a), we would be forced to assume that a) Nupe allows multiple specifiers, an otherwise unmotivated assumption given the word orders observed in the language, and b) the EPP-driven movement of the nominal element "tucks-in" to an inner specifier (Richards 1997), another stipulation lacking adequate motivation in the language. These innovations are illustrated in (15a) below.

(15) a. ConjP   b. TP
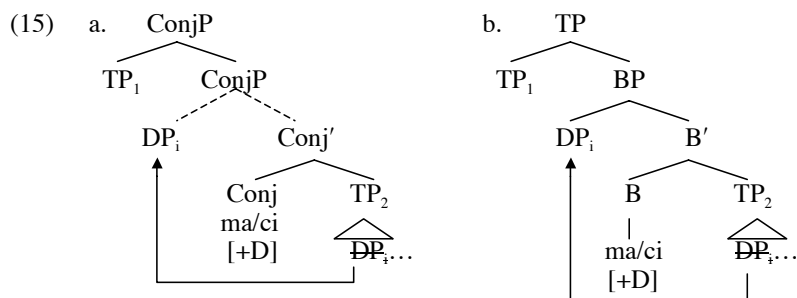
Empirical considerations also favor the analysis in (1c/15b). The subject of the first clausal conjunct can bind into the second conjunct, as shown below.

(16)   [Bagi ndondo]$_i$ á   lele [egi   u$_i$]   ma   lo makanta.
         man every   PRF sleep child 3rd.SG CONJ go school
         'Every man$_i$ slept and his$_i$ child went to school.'

This can be accounted for under the Munn analysis (1c/15b) alone, given a "first-branching category" definition of c-command (Kayne 1994). Unlike the Kayne/Johannessen approach, the c-command domain of the subject DP is the entire BP constituent containing the coordinator and the second sentential conjunct, given that the lower TP in (1c/15b) (TP$_1$) is not a *category*, but rather a *segment* of the TP category (May 1985).

A second (but less direct) empirical consideration favoring the Munn structure over the Kayne/Johannessen structure is extractability. Unfortunately, this is a sub-optimal diagnostic in this case because embedded clauses cannot be conjoined (10a) and economy principles such as Shortest Move favor conjunct-internal extraction. What we need, then, is a test construction where extraction is forced into a conjunct-external position. Under the Ross (1967)/Postal (1974) tradition, Right Node Raising is such a case. Without making any additional assumptions about Right Node Raising (e.g. the nature of the extraction/derivation), the data in (17) below minimally show that extraction from conjunct one is permitted (17a), while extraction from conjunct two (17b) is not.

(17)   a.   Musa à   ba __ Gana ma   à   gi nakàn.
             Musa FUT cut   Gana CONJ FUT eat meat
             'Musa will cut and Gana will eat the meat.'
        b.   *Musa à   ba nakàn Gana ma   à   gi __
              Musa FUT cut meat Gana CONJ FUT eat

The Kaynian/Johannessen structure makes the opposite prediction. On that analysis, extraction from conjunct one should be ruled out by the Left Branch Condition/CED and only extraction from conjunct two should be possible. In contrast, the Munn analysis correctly predicts that only extraction from the second conjunct is blocked due to its status as an adjunct island.

3.   Verb Phrase Conjunction
3.1.   Basic Properties
Verb phrases may not be conjoined by any of the overt linkers previously discussed. Coordination must be covert in this case (cf. Edo (Stewart 2001)).

(18)   *Musa ba nakàn **tò/ma/ci** lo dzukó.
          Musa cut meat CONJ   go market
          ALSO *: Musa ba nakàn dzukó (ma/ci) lo.

It is often difficult to distinguish verb phrase conjunction from clausal coordination with subject ellipsis. Two facts suggest that when a pair of adjacent verb phrases surface without marking of coordination in Nupe, the verb phrases themselves (and not two clauses containing the verb phrases) are being covertly conjoined. First, although an overt tense marker may precede the initial verb phrase, one may not precede the second (19a). Second, quantifiers do not distribute over each verb phrase individually, but rather take wide scope (19b).

(19) a.   Eza   sasi   è   ba nakàn (*à)   lo dzukó.
           person some   PRES cut meat   FUT   go market
           'Some person is (both) cutting the meat and going to the market.'
      b.   i. Lf(19b) = ∃x(*cut-meat*(x) ∧ *go-market*(x))
           ii. Lf(19b) ≠ ∃x∃y(*cut-meat*(x) ∧ *go-market*(y))

Speakers judge (19a) to be true exclusively in contexts where the intersection of [[cut-meat]] and [[go-market]] is non-empty (cf. (19b.i)). Thus, in contrast to (19b.ii), [[(19a)]] = 0 in contexts where [[cut-meat]] = {A} and [[go-market]] = {B}.

Semantically, covert coordinations are pseudo-purposives. In (20a), for example, the subject is understood as taking the knife for the purpose of using it as an instrument for cutting the meat. (20b) is semantically anomalous because meat-cutting cannot naturally be construed as a purpose for yam-eating.

(20) a.   Musa   lá   ebi   ba   nakàn.
           Musa   take knife cut   meat
           'Musa took the knife and cut the meat.'
           Alternatively, 'Musa used the knife to cut the meat.'
      b.   #Musa ba   nakàn gi   eci.
            Musa cut   meat   eat yam

Unlike in purposives, however, the event denoted by the second verb phrase conjunct is asserted. Given the semantic relationship between the two conjuncts, it is not possible to reverse the order of the verb phrase conjuncts, as it is in other instances of conjunction in the language.

(21) a.   #Eza   sasi   lo   dzukó   ba   nakàn. (compare with (19a))
            person some   go   market cut   meat
      b.   #Musa ba   nakàn lá   ebi.   (compare with (20a))
            Musa cut   meat   take knife

3.2.   Syntax of Nupe Verb Phrase Conjunction

A number of facts suggest that neither structural analysis previously considered adequately characterizes the syntax of Nupe covert coordination. Following Kandybowicz and Baker (2003), we assume that in Nupe, object Case-licensing occurs outside of the VP and that verbs raise to v^0 to support its affixal features.

Given these assumptions, we analyze covert coordination as vP conjunction, rather than conjunction of VPs.

Consider first the Kayne/Johannessen structure shown in (1b). Despite the fact that neither vP is subordinate to/contained within the other, the linear order of the conjuncts is not reversible (cf. (21)). More serious grounds for rejecting this type of analysis, however, comes from the existence of various movement operations which should be impossible given the structure in (1b), setting aside the Coordinate Structure Constraint. Because vP_1 occupies the specifier position of the Conjunction Phrase, any movement of a sub-constituent to a position outside vP_1 is predicted to be impossible, given the CED. Nonetheless, vP-internal subject raising for EPP/Case, predicate clefting (22a), and object focus (22b) are all possible from within the first conjunct of Nupe covert coordination constructions, suggesting that vP_1 does not occupy the specifier position of a phonetically null CONJ head.

(22) a. Bi-ba Musa ba nakàn lo dzukó o.
          cutting Musa cut meat go market FOC
          'It's cutting that Musa did to the meat as well as go to the market.'
      b. Nakàn Musa ba lo dzukó o.
          meat Musa cut go market FOC
          'It's meat that Musa cut and (he) went to the market.'

Lastly, the structure in (1b) incorrectly predicts that right-adjoining adverbs should be able to appear between the coordinated verb phrases and modify the first conjunct exclusively (23a), given that neither verb phrase contains the other. Although adverbs may follow the second verb phrase, they cannot be interpreted as modifying vP_2 alone (23b). This too runs contrary to the predictions of (1b).

(23) a. *Musa lá ebi sanyin/ karayin ba nakàn.
           Musa take knife quietly/carefully cut meat
      b. Musa lá ebi ba nakàn sanyin/ karayin.
          Musa take knife cut meat quietly/carefully
          'Musa quietly/carefully took the knife and cut the meat.'

Similar types of considerations suggest that a Munn-style analysis is also inadequate. That is, vP_2 cannot be generated in a projection that is an adjunct of vP_1. Because OV word orders are not attested in the second vP conjunct (18), we needn't assume that the silent Boolean head triggers an EPP movement of a nominal expression to its specifier as in Nupe clausal conjunction. Thus, structures like (1b), when applied to verb phrase conjunction, violate the general condition that all Nupe specifier positions be filled (Kandybowicz in prep). Additionally, the adjunction analysis incorrectly predicts the impossibility of predicate cleft and object focus from vP_2, an island on this analysis.

(24) a. Bi-ba Musa lá ebi ba nakàn o.
          cutting Musa take knife cut meat FOC
          'It's cutting that Musa took the knife and did to the meat.'
      b. Nakàn Musa lá ebi ba o.
          meat Musa take knife cut FOC
          'It was meat that Musa took the knife and cut.'

Furthermore, the Munn approach shares the shortcomings of the Kaynian analysis with regard to adverb placement and conjunct reversibility.

   We thus have solid syntactic evidence that neither of the structures assumed for small category conjunction and clausal coordination adequately accounts for the syntactic properties of Nupe verb phrase coordination. Given that the order of the verb phrases is fixed (i.e. they are not commutable), that right adjoining adverbs cannot come between the two verb phrases, and that extraction of all constituents within each verb phrase is possible, we propose that vP_2 is merged as a complement of V_1 and that the coordination of the two constituents takes place in the semantics rather than in the narrow syntax, as in (consequential) serial verb constructions (Stewart 2001:ch. 2). We analyze the identity of the subject in both verb phrases to be mediated by Control, as opposed to ATB movement of two identical VP-internal DPs. Nupe covert coordination is thus *asyntactic* conjunction on this analysis. The structure we assume is provided below.

(25)
```
                vP₁
          ┌──────┴──────┐
      DP_i-SUBJ         v′
                   ┌─────┴─────┐
                   v          VP₁
                         ┌─────┴─────┐
                      DP_OBJ        V′
                               ┌─────┴─────┐
                               V          vP₂
                            ┌──────────┴──────────┐
                          PRO_i                  VP₂
```
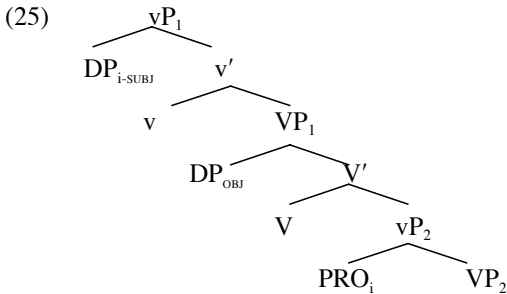
4.   Conclusion

Our study of Nupe conjunction has yielded a number of theoretical insights. We have argued that there are three distinct conjunction strategies in the language, each exhibiting distinct syntactic properties. That is to say, there is evidence for a typology of coordinate structures. This conclusion runs counter to several recent trends in the literature that strive to provide a unified syntax for all coordinate constructions (Munn 1993, Kayne 1994, Johannessen 1998, etc.). Although we have found evidence for each of the leading types of structures proposed in this literature, it is clear that there is no homogeneous analysis of coordination, at least for Nupe. This discovery carries with it the corollary that coordination is not universally reducible to a single clause-level phenomenon supplemented with

ancillary transformational operations such as "conjunction reduction" (Johannessen 1998, among others). This paper thus bears witness to the encouraging amount of progress made in the field of coordination, but at the same time highlights the fact that our understanding of the phenomenon is still limited.

References

Büring, Daniel, and Katharina Hartmann. 1998. Asymmetrische Koordination. *Linguistische Berichte* 174:172-201.

Caponigro, Ivano. 2003. Unbalanced coordination in Maasai. In Jason Kandybowicz, ed., *UCLA Working Papers in African Linguistics 3*:1-16.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

--. 2001. Derivation by phase. In Michael Kenstowicz, ed., *Ken Hale: A Life in Language*, 1-52. Cambridge, MA: MIT Press.

Johannessen, Janne Bondi. 1998. *Coordination*. Oxford: Oxford University Press.

Kandybowicz, Jason. In prep. Conditions on multiple copy spell-out and the syntax-phonology interface. Ph.D. diss, UCLA.

Kandybowicz, Jason, and Mark C. Baker. 2003. On directionality and the structure of the verb phrase: Evidence from Nupe. *Syntax* 6(2):115-166.

Kayne, Richard. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

--. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

May, Robert. 1985. *Logical Form*. Cambridge, MA: MIT Press.

Munn, Alan. 1993. Topics in the syntax and semantics of coordinate structures. Ph.D. diss., University of Maryland, College Park.

Postal, Paul. 1974. *On Raising*. Cambridge, MA: MIT Press.

Richards, Norvin. 1997. What moves where when in which language? Ph.D. diss., MIT.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman, ed., *Elements of Grammar*, 281-337. Dordrecht: Kluwer.

Ross, John Robert. 1967. Constraints on variables in syntax. Ph.D. diss., MIT.

Smith, Neil V. 1967. *An Outline Grammar of Nupe*. London: Luzac.

Stewart, Osamuyimen Thompson. 2001. *The Serial Verb Construction Parameter*. New York, NY: Garland.

Velde, John R. 2000. Assumptions about the structure of coordination. In Kerstin Schwabe and Niina Zhang, eds., *Ellipsis in Conjunction*, 322-341. Tübingen: Max Niemeyer Verlag.

Zoerner, Ed. 1999. One coordination for all. *Linguistic Analysis* 29:322-341.

Jason Kandybowicz
University of California, Los Angeles
3125 Campbell Hall, Box 951543
Los Angeles, CA 90095-1543

jkandy@humnet.ucla.edu

Emai's Aspect-Causative Interaction[*]

RONALD P. SCHAEFER[1] and FRANCIS O. EGBOKHARE[2]
Southern Illinois University Edwardsville[1] and University of Ibadan[2]

Talmy (2000) articulates a typology of lexicalization potentials for how verb roots incorporate semantic elements of aspect, causation, and their interaction. The latter, for instance, reveals three semantic types – stative (being in a state), inchoative (entering a state), and agentive (being put in a state), with verb roots cross-linguistically registering restricted combinations of types and achieving others via grammatical augmentation. As this typology unfolds, Talmy notes that languages often lexicalize posture and position events asymmetrically, although no clear boundary is evident and overlap might thus occur. For this paper, we explore how posture and position verbs manifest aspect-causative (A-C) interaction in the Edoid language Emai (Elugbe 1989, Williamson and Blench 2000). Typologically, Emai exhibits relatively strict SVO word order and employs verbs in series. Its intransitive posture verbs and transitive position verbs exhibit unmarked and marked syntactic constructions that reveal a shift from one A-C type (stative or agentive) to another (inchoative). To signal an inchoative change in orientation ("move up/down") for a theme participant, each adds a second verb to a construction whose grammatical properties are then highly restrictive.

Transitive position verbs occur in unmarked and marked constructions. In unmarked constructions, the verbs *khuae* 'raise' (1a) and *gbe* 'fell' (1b) are preceded by an agent subject and followed by a theme direct object.

(1) a. òjè khúáé ó̩lì ùkòdò.[1]
    Oje raise the pot
    'Oje raised the pot.'

[1] Orthographic conventions for Emai are consistent with those in Schaefer (1987) and Schaefer and Egbokhare (1999), where <o̩> represents a lax mid-back vowel, <e̩> a lax mid-front vowel, and <vb> a voiced bilabial approximant. High tone is marked by an acute accent, low tone by a grave accent and high downstep by an acute accent followed by an apostrophe.

b. òjè gbé ọ́lí óràn.
   Oje fell the tree
   'Oje felled the tree.'

Each position verb appears in a corresponding marked construction where its theme participant has entered or assumed a positional end state. This inchoative construction is formed by adding a verb in series (*nwu* with *khuae* in (2a) and *fi* with *gbe* in (2b)) and by changing word order so that the erstwhile direct object now precedes the position verb and follows the added construction verb.

(2) a. òjè nwú   ọ́lì ùkòdò khúáé.
       Oje pick.up the pot   raise
       'Oje raised the pot up at arm's length.' / 'Oje got the pot raised up.'

   b. òjè fí   ọ́lí óràn gbé.
      Oje drop the tree fell
      'Oje got the tree down (flat out).' / 'Oje felled the tree (down) flat out.'

Added construction verbs can also appear as the primary predicate of a simple clause, particularly in non-literal structures with body-part subjects or direct objects. *nwu* has the sense 'pick up' (3a-b) and *fi* the meaning 'drop, dangle' (3c-d). Each reflects movement and a corresponding orientation: "upward" for *nwu* and "downward" for *fi*.

(3) a. òjè nwú   úkpórán vbì   òtọ̀ì.[2]
       Oje pick.up stick   LOC   ground
       'Oje picked up a stick on the ground.'

   b. òjè ò ó nwù   égbè.
      Oje SC C pick.up body
      'Oje is getting prepared.'

   c. òjè fí   òè ó vbì   òò.
      Oje drop foot CL LOC   hole
      'Oje stepped into the hole.'

   d. áwẹ́ ísì òjè ò ó fí.
      legs ASS Oje   SC C drop
      'Oje's legs are dangling.'

---

[2] Abbreviations used throughout this study include the following: ANT=anterior, ANTI=anticipative, ASS=associative, C=continuous, CL=change of location, CON=conative, D=displacement, F=factative, LOC=locative, NEG=negative, PF=positive focus, PR=prohibitive, SC=subject category, TEMP=temporal perspective.

Marked constructions with position verbs reveal a stringent grammatical character. They require perfective tense/aspect (*òjè* completive present and *ójé* completive past (4a)), and reject imperfective continuous (4b).

(4) a. òjè / ójé nwú   ólì ùkòdò khúáé.
      Oje Oje pick.up   the pot   raise
      'Oje has raised up the pot.' / 'Oje raised up the pot.'

   b. * òjè <u>ò</u> <u>ó</u> nwù   òlí úkódó khùàè.
       Oje SC C pick.up the pot   raise
       'Oje is raising the pot up (at arm's length).'

Marked constructions do not accept any aspectual or temporal forms. For instance, they disallow particles for preverbal aspectualizers that access temporal phases of event onset (<u>óó</u>' CON 'go to' (5a)), postverbal temporal perspective (*léé* TEMP 'already' (5b)) as well as adjuncts of temporality (*òdè̩* 'yesterday' (5c)).

(5) a. * òjè <u>óó</u>'   nwù   ólì ùkòdò khúáé.
      Oje CON   pick.up the pot   raise
      'Oje went to raise the pot up (at arm's length).'

   b. * òjè nwú   ólì ùkòdò khúáé léé.
       Oje pick.up the pot   raise   TEMP
       'Oje already raised the pot up.'

   c. * ójé nwú   ólì ùkòdò khúáé òdè̩.
       Oje pick.up the pot   raise yesterday
       'Oje raised the pot up (at arm's length) yesterday.'

Marked constructions limit particles that bear on event assertion/polarity and speaker attitude toward event occurrence. They reject predicate negation and prohibition (6a-b) in addition to intentionality preverbs (*dobo̩* 'mistakenly' (6c)).

(6) a. * òjè í   ì   nwù   òlí úkódó khùàè.
      Oje SC NEG pick.up the pot   raise
      'Oje did not raise the pot up (at arm's length).'

   b. * é   è   nwú   ólì ùkòdò khúáé.
       you PR pick.up the pot   raise
       'Don't raise the pot up (at arm's length).'

   c. * ójé dóbò̩   nwú   ólì ùkòdò khúáé.
       Oje mistakenly pick.up the pot   raise
       'Oje mistakenly raised the pot up (at arm's length).'

Ronald P. Schaefer and Francis O. Egbokhare

Intransitive posture verbs exhibit construction pairs highlighting their shift from stative to inchoative. In unmarked constructions with a stative character, *muzan* 'stand' and *m̲e̲h̲e̲n* 'lie' appear with a locative complement.

(7) a. ólí ómó múzán-í vbì ìtébù.
     the child stand-F LOC table
     'The child stood on the table.'

    b. ólí ómó méhén-í vbì ìtébù.
     the child lie-F LOC table
     'The child lay on the table.'

Posture verbs also appear in marked constructions where body orientation of the subject referent has assumed a positional end state. Inchoative "stand up" requires *muzan* in series with the verb *daa* 'raise' (8a) and "lie down" demands *m̲e̲h̲e̲n* with the verb *dee* 'lower' in series (8b).

(8) a. ólí ómó dáá múzán.
     the child raise stand
     'The child stood up.'

    b. ólí ómó déé méhén.
     the child lower lie
     'The child lay down.'

Verb forms in series with posture verbs can serve as the primary predicate of a simple clause. Transitive *daa* 'raise' occurs with a body-part complement, while *dee* 'lower' appears with or without a body-part complement (9a-b).

(9) a. òjè ò ó dàà àgbàn.
     Oje SC C raise chin
     'Oje is raising his chin.' / 'Oje is searching for a better view.'

    b. òjè déé ré / déé óbò ré.
     Oje lower D lower hand D
     'Oje lowered himself / lowered his hand.'

Marked constructions with posture verbs exhibit severely restricted syntax. They permit perfective tense/aspect (*ólí ómò* completive present and *ólí ómó* completive past (10a)) but not imperfective continuous (10b).

(10) a. ólí ómò / ólí ómó dáá múzán.
      the child the child raise stand
      'The child has stood up.' / 'The child stood up.'

   b. * <u>ó</u>lí <u>ó</u>mò <u>ó</u>   <u>ò</u> dàá   mùzán.
        the child SC   C raise stand
        'The child is standing up.'

They do not permit aspectual and temporal grammatical forms: neither aspectualizer preverbs (<u>óó</u>' 'go to' (11a)), temporal perspective postverbal particles (*léé* 'already' (11c)), nor temporal adjuncts (*òd<u>è</u>* 'yesterday' (12d)).

(11) a. * <u>ó</u>lí <u>ó</u>mò <u>óó</u>'   dàá   múzán.
            the child CON   raise stand
            'The child went to stand up.'

   b. * <u>ó</u>lí <u>ó</u>mò dáá   múzán léé.
        the child raise stand TEMP
        'The child stood up already.'

   c. * <u>ó</u>lí <u>ó</u>mó dáá múzán òd<u>è</u>.
        the child raise stand   yesterday
        'The child stood up yesterday.'

As well, marked posture verb constructions do not permit the speaker to register polarity (negation and prohibition particles, (12a-b)) or attitude toward event occurrence (intentionality preverbs like *dob<u>o</u>* 'mistakenly' (12c)).

(12) a. * <u>ó</u>lí <u>ó</u>mò í   ì   dàá   mùzán.
            the child SC NEG raise stand
            'The child did not stand up.'

   b. * é   è   kè   dáá   múzán.
        the child ANT raise stand
        'Don't stand up anymore.'

   c. * <u>ó</u>lí <u>ó</u>mó dób<u>ò</u>   dáá   múzán.
        the child mistakenly raise   stand
        'The child mistakenly stood up.'

   Marked constructions do not seem constrained primarily by their position or posture verbs and their lexical properties. Rather, their stringent syntax appears to be a function of the inchoative construction. Position verbs in unmarked constructions, for example, entertain the range of grammatical forms rejected by their marked counterparts. They permit imperfective continuous tense/aspect (13a) as well as completive present and completive past (13b).

(13) a. òjè <u>ò</u> <u>ó</u> khùàè <u>ó</u>lì ùkòdò.
     Oje SC C   raise the pot
     'Oje is raising the pot.'

   b. òjè / ójé khúáé <u>ó</u>lì ùkòdò.
      Oje   Oje raise the pot
      'Oje has raised the pot.' / 'Oje raised the pot.'

Unmarked position-verb constructions allow grammatical forms denoting an event's aspectual and temporal properties. They accept preverbs that access aspectual phases of event onset (<u>óó</u>' (14a)), postverbal particles of temporal perspective (*léé* (14b)), and adjuncts of temporality (*òd<u>è</u>* (14c)).

(14) a. òjè <u>óó</u>'   khùàè <u>ó</u>lì ùkòdò.
     Oje CON   raise the pot
     'Oje went to raise the pot.'

   b. òjè khúáé <u>ó</u>lì ùkòdò léé.
      Oje raise the pot   TEMP
      'Oje has finished raising the pot.' / 'Oje already raised the pot.'

   c. ójé̩ khúáé <u>ó</u>lí úkódó òd<u>è</u>.
      Oje raise the pot   yesterday
      'Oje raised the pot yesterday.'

They also admit forms for event assertion/polarity and speaker attitude toward event occurrence. Predicate negation and prohibition auxiliaries (15a-b) as well as preverbs of intentionality (*dob<u>o</u>* 'mistakenly' (15c)) are acceptable.

(15) a. òjè í ì   khùàè <u>ó</u>lì ùkòdò.
     Oje SC NEG   raise the pot
     'Oje did not raise the pot.'

   b. é   è   khúáé <u>ó</u>lì ùkòdò.
      you PR raise the pot
      'Don't raise the pot.'

   c. ójé dób<u>ò</u>   khúáé <u>ó</u>lì ùkòdò.
      Oje mistakenly raise   the pot
      'Oje mistakenly raised the pot.'

   Unmarked position-verb constructions and their posture neighbors thus exhibit far less restrictive behavior in contrast to the stringent syntactic behavior of marked inchoative constructions. This suggests that the inchoative's rigid

grammatical character derives from a structural template more complex than the lexical projection of a posture or position verb. What is the nature of that template? In the African context, one might look to the frequently noted consecutive construction. Niger-Congo languages often signal a temporal sequence of actions by marking the second verb in series overtly: Nupe's *ci*/'and' (16a; Hyman 1971) and Tswana's à/CONS (16b; Creissels 2000).

(16) a. u   lá   dùku   ci   bé.
       he   take   pot   and   come
       'He took the pot and came.'

   b. kì̠-ìlé   tù̠rópó-ng   k-à-ré̠ká   dítlhàkú̠.
      SM.1S-go.ANT   town-LOC   SM.1S-CONS-buy   C18.shoes
      'I went to town and bought shoes.'

Emai exhibits no formal marking of consecutive constructions. However, it does show constructions with a similar function that also reference a temporal sequence of events (17a). Functional consecutives and transitive inchoative constructions exhibit some commonality. Both manifest a surface syntax consisting of NP1-V1-NP2-V2 and observe a linear-order constraint. Functional consecutive constructions strictly limit syntactic ordering of their verbs (17b) and position verbs never precede their construction verb in inchoative sentences (17c).

(17) a. òjè   dé̠   émà   é.
       Oje   buy   yam   eat
       'Oje bought yam and ate it.'

   b. * òjè   é   émà   dé̠.
       Oje   eat   yam   buy
       'Oje ate yam and bought it.'

   c. * òjè   khúáé   ó̠lì   ùkòdò   nwú.
       Oje   raise   the pot   pick.up
       'Oje raised the pot and picked it up.'

Moreover, inchoative constructions reflect the orientation of their theme participant through verb selection. Contrasting orientations of the repositioned theme correlate with distinct verbs in series: *nwu khuae* 'raise up' and *fi gbe* 'fell down' among position verbs and *daa muzan* 'stand up' and *dee me̠he̠n* 'lie down' among posture verbs. Alternative verb combinations are disallowed; one cannot for example pair *fi* with *khuae* or *dee* with *muzan*.

   Emai's functional consecutives and positional transitive constructions are distinct in other respects, however. Functional consecutives deconstruct in a strict

fashion. In a mono-verbal clause, each verb of the consecutive is found with the construction's grammatical subject and direct object.

(18)   òjè dé̱ émà / òjè é̱   émà.
       Oje buy yam   Oje eat   yam
       'Oje bought yam.' / 'Oje ate yam.'

Transitive inchoative constructions do not deconstruct consistently. Although V2 from (19a) is obviously compatible with its preceding noun phrase (19b), V1 most often is not (19c). Simple transitive clauses composed of *fi* and the theme *ó̱lí óràn* 'the tree' from inchoative constructions are ungrammatical. *fi* allows only the meaning 'throw' in simple transitive clauses with theme direct objects (19d).

(19) a. òjè fí   ó̱lí óràn gbé.
        Oje drop the tree   fell
        'Oje got the tree down (flat out).' / 'Oje fell the tree (down) flat out.'

     b. òjè gbé ó̱lí óràn.
        Oje fell the tree
        'Oje felled the tree.'

     c. * òjè   fí   ó̱lí óràn.
          Oje   drop the tree
          'Oje dropped the tree.'

     d. òjè fí   ó̱lí úkpóràn.
        Oje threw the stick
        'Oje threw the stick.'

    As an alternative, one might consider the inchoative construction as a classic resultative, which also exhibits the surface syntax NP1-V1-NP2-V2, and seek to explicate V2 as intransitive. Emai has few resultatives, but they have been noted in analyses of other West African languages (Durie 1997). NP2 in resultative constructions functions as grammatical direct object for V1 but logical subject for V2 (20a). And like consecutives, the resultative deconstructs strictly (20b-c).

(20) a. òjè hó̱ó̱ ó̱lí úkpùn fúán.
        Oje wash the cloth   clean
        'Oje washed the cloth clean.'

     b. òjè hó̱ó̱ ó̱lí úkpùn.
        Oje wash the cloth
        'Oje washed the cloth.'

    c. <u>ó</u>lí úkpùn fúán-ì.
       the cloth be.clean-F
       'The cloth is clean.'

By analogy, one might construe a V2 like *khuae* in an inchoative construction as intransitive. *<u>ó</u>lì ùkòdò* 'the pot' would then serve as *khuae*'s logical subject, where English translation in (21a) suggests this analysis. A major obstacle for this interpretation is that *khuae*, like other inherently transitive position verbs, does not occur as a simple intransitive (21b). Moreover, it is important to recognize that Emai has no syntactic passive construction whereby *khuae*, and others of its ilk, might become syntactically intransitive.

(21) a. òjè nwú   <u>ó</u>lì ùkòdò khúáé.
      Oje pick.up the pot   raise
      'Oje picked up the pot and it raised up (at arm's length).'

    b. * <u>ó</u>lì ùkòdò khúáé-ì.
        the pot   raise-F
        'The pot was / got raised.'

    Additional facts with a body-part subject show that an inchoative construction with a position verb like *khuae* requires co-event expression by a verb in series. That is, *khuae* in (22a) requires the verb in series *tin* 'fly' (22b) in order to convey the notion 'become erect.' *khuae* alone is not sufficient (22c).

(22) a. úkp<u>é</u>g<u>é</u>l<u>é</u> ísì   òjè tín khúáé.
      penis   ASS Oje fly raise
      'Oje's penis became erect.' / 'Oje's penis flew up.'

    b. <u>ó</u>lí áfìánmì <u>ò</u> <u>ó</u> tín.
       the bird   SC C fly
       'The bird is flying.'

    c. * úkp<u>é</u>g<u>é</u>l<u>é</u> ísì   òjè khúáé-ì.
        penis   ASS Oje raise-F
        'Oje's penis became erect. / Oje's penis raised up.'

    Transitivity and word order properties thus suggest that marked constructions for position and posture verbs do not clearly align with widely recognized functional consecutive or resultative constructions. What about other grammatical properties? Let's consider additional facets of functional consecutives, beginning with tense/aspect. Functional consecutives accept perfective tense/aspect (*òjé* completive present and *ójé* completive past (23a)) but reject imperfective continuous (23b), as did marked inchoative constructions.

(23) a. òjè / ójé dé̞ émà é.
     Oje   Oje buy yam eat
     'Oje has bought yam and eaten it.' / 'Oje bought yam and ate it.'

   b. * òjè <u>ò</u> <u>ó</u> dè̞ èmá è.
       Oje SC C buy yam eat
       'Oje is buying yam and eating it.'

Emai's functional consecutive constructions allow a range of aspectual and temporal forms. They admit aspectualizer preverbs accessing event phases (<u>óó</u>' (24a)), postverbal temporal perspective particles (*léé* (24b)) and temporal adjuncts (*òdè̞* (24c)). Recall that inchoative constructions accepted none of these.

(24) a. òjè <u>óó</u>'   dè̞ émà é.
     Oje CON buy yam eat
     'Oje went to buy yam and eat it.'

   b. òjè dé̞ émà é léé.
     Oje buy yam eat TEMP
     'Oje already bought yam and ate it.'

   c. ójé dé̞ émà é òdè̞.
      Oje buy yam eat yesterday
     'Oje bought yam and ate it yesterday.'

And unlike inchoatives, functional consecutive constructions permit particles that bear on event assertion/polarity (predicate negation and the prohibition (25a-b)) and speaker attitude toward event occurrence (intentionality preverbs of the type *dobo̞* 'mistakenly' (25c)).

(25) a. òjè í ì   dè̞ èmá è.
     Oje SC NEG buy yam eat
     'Oje did not buy yam and eat it.'

   b. é è dé̞ émà é.
     you PR buy yam eat
     'Don't buy yam and eat it.'

   c. ójé dóbò̞   dé̞ émà é.
     Oje mistakenly buy yam eat
     'Oje mistakenly bought yam and ate it.'

    We thus conclude that functional consecutive constructions and marked position-verb constructions (as well as marked posture verb constructions) are not

of the same grammatical type. Building on this, we turn to the consecutive construction's temporal "and then" interpretation. Emai reveals semantic relations other than temporal sequence for constructions that might otherwise appear to be functional consecutives. Temporal sequence fails to capture their semantic identity. For example, Emai has serial constructions exhibiting purposive (26a) or extentive (26b) relations. The first verb event in the purposive is construed as actualized while the second is not (*hua ka* 'carry to dry'), whereas both verb events in the extentive are actualized (*vie̱ me̱he̱n* 'cry until asleep').

(26) a. òjè   húá   ìwàwà   ká.
         Oje   carry   cooking.pots dry
         'Oje carried cooking pots to dry.' /
         * 'Oje carried cooking pots and they dried.'

     b. òjè   víé̱   mé̱hé̱n.
         Oje   cry   sleep
         'Oje cried until asleep.' / * 'Oje cried and then slept.'

Using this richer set of semantic relations, we suggest that position and posture inchoatives are more akin to extentive constructions than to purposives. The latter reveals an actualization pattern of events not shown by marked position and posture constructions. That is, Emai inchoative constructions express entering a state of posture or position by incorporating an orientation-bearing verb whose actualization is required over a temporal extent that concludes with a change in postured or positioned state: "event X until end state Y." Hence, the marked inchoative constructions in (2a) and (8a) might be rendered explicitly and symmetrically as 'Oje moved the pot upward until it entered a raised state' and 'Oje moved himself upward until he entered a standing state'. It is this extentive ('until') parameter that precludes grammatical augmentation by forms bearing on event aspectuality (e.g. *óó̱'* 'go to') and temporality (e.g *òdè̱* 'yesterday'). Apparently, it is also this parameter that rules out expression of polarity (negation and prohibition) and speaker attitude (*dobo̱* 'mistakenly'). One of the major tasks confronting further investigation of Emai concerns the identification of semantic relation types in serial constructions (e.g. temporal consecutive, resultative, purposive, extentive) and how each correlates with a particular set of grammatical properties. This applies in particular to events of position and posture not explicitly considered in this paper and to broader syntactic issues of construction verb transitivity as it relates to body-part complements. More importantly, we need to clarify the principles that constrain semantic relation types in the overall serial verb system and that link individual types to specific syntactic patterns.

References

Creissels, Denis. 2000. Typology. In B. Heine and D. Nurse, eds., *African Languages: An Introduction*, 231-258. New York: Cambridge University Press.

Durie, Mark. 1997. Grammatical structures in verb serialization. In A. Alsina, J. Bresnan and P. Sells, eds., *Complex Predicates*, 289-354. Stanford: CSLI.

Elugbe, Ben. 1989. *Comparative Edoid: Phonology and Lexicon*. Port Harcourt: University of Port Harcourt Press.

Hyman, Larry. 1971. Consecutivization in Fe'fe'. *Journal of African Languages* 10:29-43.

Schaefer, Ronald. P. 1987. *An Initial Orthography and Lexicon for Emai: an Edoid Language of Nigeria*. Bloomington: Indiana Linguistics Club.

Schaefer, Ronald P. and Francis O. Egbokhare. 1999. *Oral Tradition Narratives of the Emai People, Parts I and II*. Hamburg: LIT Verlag.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics, Vols. 1 and 2*. Cambridge, MA: MIT Press.

Williamson, Kay. 1989. Benue-Congo Overview. In J. Bendor-Samuel, ed., *The Niger-Congo Languages*, 247-274. New York: University Press of America.

Williamson, Kay and Roger Blench. 2000. Niger Congo. In Bernd Heine and Derek Nurse, eds., *African Languages: An Introduction*, 11-42. New York: Cambridge University Press.

Ronald P. Schaefer
Department of English
Southern Illinois University Edwardsville
Edwardville, IL 62026-1431

rschaef@siue.edu

Francis O. Egbokhare
Department of Linguistics and African Languages
University of Ibadan
Ibadan, Nigeria

foegbokhare@yahoo.com

Yobe State, Nigeria as a Linguistic Area

RUSSELL G. SCHUH
University of California, Los Angeles

0.   Introduction

In late summer 1969, I arrived in Potiskum, Nigeria for the first time. I was a Research Assistant gathering data on the Ngizim language for a comparative Chadic syntax project[1] and for my dissertation. I could not have imagined that 30 years later I would be returning to Potiskum, essentially to pick up where I left off.

From 2001-2004, I worked in collaboration with Dr. Alhaji Maina Gimba of the University of Maiduguri and speakers of five languages of Yobe State, Nigeria to document these languages. The primary focus was lexical and morphological documentation and collection of texts. An important aspect of the project was to have native-speaking participants do most of the data collection, and resulting documents have been printed and distributed locally as a stimulus for members of the respective communities to continue adding to the documentation after the end of the project.

Though the Nigerian states are somewhat artificial political creations, Yobe State, fortuitously, has an interesting linguistic composition. There are six distinct modern languages that are indigenous to the area that is now Yobe State—Duwai, Ngizim, Bade, Karekare, Bole, Ngamo—and with the exception of Bole, these languages are spoken almost entirely within the confines of Yobe State (Bole has a large number of speakers to the south, in Gombe State). The other three languages with large resident populations in Yobe State are Kanuri, Hausa, and Fulfulde, all of which have spread into Yobe State area from elsewhere, though

the eastern and far northern parts of the state have been Kanuri-speaking for several centuries. Figure 1 locates the languages discussed in this paper.

The Yobe languages project has revealed yet another unifying linguistic trait of Yobe State, namely the indigenous languages in this state form a *Sprachbund*—an area in which languages share a cluster of typologies that are absent, at least as a cluster, outside the area. Although the indigenous languages of Yobe State are all members of the West Branch of the Chadic family, they fall into two distinct groups. Karekare, Bole, and Ngamo are members of the "A" group of West Chadic, whereas Duwai, Ngizim, and Bade are members of the "B" group.[2] These two language groups, however, share properties that must result from areal diffusion rather than inheritance from ancestral languages of the respective subgroups or from a common ancestral language. This paper describes areal lexical features, one areal feature from morphology, and one from syntax.

Figure 1: Yobe State, Nigeria and the Yobe languages

Yobe State within Nigeria   The locations of the Yobe languages

In order to discuss the nature of Yobe State as a linguistic area, some additional information about the linguistic situation is necessary. The map in Figure 1 shows that Ngizim, Karekare, Ngamo, and Bole are geographically contiguous whereas Bade and Duwai[3] are geographically separated from this

---

[2] See Newman (1977) for the classification of the Chadic languages. Dialectally, Ngizim and Karekare are relatively uniform. Bole has a major dialect split, roughly defined by the Gongola River that forms part of the southwestern border between Yobe and Bauchi States. This paper considers only the Fika dialect, which is that of Yobe State. Ngamo has a major split between the Gudi dialect to the east, indicated Ngamo (G) here, and the Yaya dialect, to the west, indicated Ngamo (Y) here. Bade is dialectally so diverse that it might be considered a group of closely related languages (Schuh 1981). This paper has data from the Western variety, Bade (W), and the Gashua variety, Bade (G).

[3] The Duwai area, which is not separately designated on the map, is contiguous to the Bade area, east and southeast of Gashua. Study of Duwai was not part of the Yobe languages project. While

group by a Kanuri/Hausa speaking zone.[4] I will refer to the former group as the "Potiskum area" languages since Potiskum is the largest city in the vicinity and many speakers of all the languages live in Potiskum. I will refer to the latter group as the "Bade-Duwai area" languages. In fact, each of these areas comprises its own *Sprachbund*. In this paper I will focus on the Potiskum area.

Although Ngizim is clearly more closely related to Bade than either of those languages is to Duwai, Bade and Duwai share certain properties not found in Ngizim and vice versa. Some of these properties must be inherited from the common ancestral language but have been lost in Ngizim, e.g. 'nine' is Duwai *wàarìyà*, Bade (Gashua) *wuliyà*, Bade (Western) *wurayà*, but Ngizim *kuɗkûvdà*. Others are a result of diffusion within the Bade-Duwai area that must have taken place after their geographical separation from Ngizim, e.g. a change of original initial CəCV to əCCV as in the word for 'thirst': Duwai *ə̀gjì*, Bade (Gashua) *ə̀gjì*, Bade (Western) *ə̀gjaan*, but Ngizim *gə̀jì* (Schuh 1978a, 1981). Bade-Duwai thus provide a way to identify properties in Ngizim that result from areal diffusion within the Potiskum area, i.e., where Bade-Duwai share features not found in Ngizim or Ngizim-Karekare-Bole-Ngamo share features not found in Bade-Duwai, a likely explanation is diffusion of those features in the respective areas.

All of the existing close linguistic relatives of Karekare, Ngamo, and Bole lie to the south, in Gombe and Bauchi states. Two of the better known and better documented languages are Tangale and Kanakuru. Others are Kirfi, Galambu, Gera, Kwami, and Pero.[5] These languages can provide evidence regarding features that seem special to the Yobe State languages, and more specifically, the Potiskum area languages, such as words like Ngizim *buucî*, Karekare *buucì*, Ngamo *bùushî*, Bole *buushì* all meaning 'palm leaf mat' (an old loanword from Kanuri), which does not seem to be found in West Chadic languages outside Yobe State.

1.   Lexicon
1.1.   Shared Lexicon in the Potiskum Area
The most readily visible (or better, audible) criterion for believing that the Yobe area is a *Sprachbund* is a large number of shared lexical items that have probably

---

working on Duwai in 1973-75, I had the impression that it was in the process of being replaced by Bade and Hausa, or, at its eastern extremities, by Kanuri. However, during the Yobe project I discovered that my prediction of Duwai's demise was premature! It still seems to be actively spoken by a significant population.

[4] Fulfulde is spoken throughout northern Nigeria by nomadic herders and in Fulfulde-speaking villages and neighborhoods of larger towns. In Yobe State and contiguous areas, Fulfulde has exerted virtually no linguistic influence on the languages with which Fulfulde speakers come in contact. Though the Fulb'e retain a strong cultural identity and typically speak Fulfulde within their own communities, they all become fluent speakers of one or more of the local languages, whereas it is rare to find speakers of other languages who acquire speaking ability in Fulfulde.

[5] The largest Chadic language speaking group immediately south of Bole is Tera, with whom the Boles have a special historical connection (Newman 1969/70). Tera belongs to the Biu-Mandara Branch of Chadic and is typologically quite different from the West Chadic languages.

not been inherited directly from a Chadic ancestor language but where the original source cannot be pinpointed with certainty, as would be the case of obvious loanwords (see below). More specifically the languages of the Potiskum area (Karekare, Ngamo, Bole, and Ngizim) share many words that are apparently not shared by Bade nor are they obvious loanwords from outside the Yobe area.[6] Particularly significant is the fact that such words are shared between Ngizim and the other three languages, to which Ngizim is not closely related, but not between Ngizim and its close cousin, Bade. Also significant are words where Karekare is phonologically nearly identical to Bole and Ngamo. Though all three languages are in the Bole-Tangale group, Karekare is not closely related to Bole and Ngamo within that group.

The tables in (1) present some examples, roughly grouped into social/semantic categories. All items include Ngizim and at least one of the Bole-Tangale languages. The fact that items are missing for particular languages means only that those words did not come up in the current project. In most cases, corresponding words probably do exist in those languages. In fact the significance of these lists is that the words came up with no expectation on my part of which items WOULD be shared across languages. Ngamo examples are from the Gudi dialect unless otherwise noted.

(1)   Examples of Yobe areal words

	KAREKARE	NGAMO (G)	BOLE	NGIZIM
Shared material culture (foods, clothes, household items)				
'steamed cake'		làmbà	làmbà	làmbà
'locust bean cake'			gìskirmi	gə̀skərmi
'woman's loincloth'	diidàm	dìidâm	diidàm	diidàm
'basket'	dàaɓir	dàa'ùr	dàaɓur	dàaɓə̀r̃
'stalk door panel'	gwamper		gompor	gwampèr̃
Animals and birds				
'baboon'	bangài	bàngêi	bangè	bangâi
'hairy goat or sheep'	bàzaa	bàzâ	bàza	bàzà
'domestic pigeon'	bàru	bàrù	bàru	bàri
'grey hornbill'	tiilaakò	tilàakǒ		tiilaakòk
Farming				
'type of sorghum'	dàashàa	dàashà (Y)	dàashà	dàashà
'young corn'	kuncàu	kùnshô	kunsho	kunco
'buried granary'		daadàfiiɗà	daadafiiɗà	daatəfiiɗòk

---

[6] Some of these items do have counterparts outside the Yobe region. For example, the word defined as 'locust-bean cake' in the first group below was identified by Yobe speakers as Hausa *gàskamii*, defined by Bargery (1934) as 'the mealy pulp from the inside of locust-bean pods'. Unlike typical loanwords, this word cannot be related phonologically in a straightforward way across languages, and indeed, the direction of borrowing itself is not obvious.

	KAREKARE	NGAMO (G)	BOLE	NGIZIM
Shared aspects of culture (implements, customs, occupations)				
'large pressure drum'		kànjâu	kanjàu	kanjâu
'widow'		gudgùm	gudugùm	gudùgùm
Descriptive terms				
'cold'	layi	leilei	lai	layi
'sour'	zhiimu 'smell sour'	shòmshòm	shòmshòm	còmcòm
Expressions and grammatical markers				
'until; even'		kaba	kapa	kapa
'that's it, OK'	anyà	anyà	anyà	anyà

Shared lexical traits extend above the word level. These include idioms, such as Ngamo (Y) *fàna sàra*, Bole *'yùwa sara*, Ngizim *jɔ̀bà amài* 'offer condolences', all meaning "catch hand" in the respective languages, and many proverbs and riddles that seem to be regional rather than general to (West) Africa.

1.2. Kanuri as an Areal Force in the Yobe Languages

In northeastern Nigeria, Kanuri, a language of the Nilo-Saharan family, was the dominant cultural and linguistic force for centuries, extending into the 20th century. This is evident in the large number of Kanuri loanwords in the Yobe languages. Although it is possible to identify many loanwords introduced directly from Kanuri into one or more of the Yobe languages, there is an extensive group of words of Kanuri origin that typify the languages of the Yobe area but for which it is not possible to identify the language or languages that served as the path by which the words were introduced. Moreover, these words set the Yobe languages apart from related languages, such as Tangale and Kanakuru, which are outside the Yobe area. A few such words are the following:

(2)   Some Kanuri loanwords shared across the Yobe languages

Kanuri	Bade (W)	Bade (G)	Ngizim	Karekare	Ngamo (G)	Bole	
jirè	jərén	jiĩrâi	jiĩrèewa	jirè	jìrê	jirè	'truth'
karè	karen	kaĩrài	kaĩrê	karài	kàrêi	karài	'stuff'
ngudì	ǹgudíin	ǹgudì	ngudì	ngudì	ngùdî	ngudì	'lazy p.'
ngalwò	ngalkò	ǹgalkò	ngalkò	ngalkò	ngalkò	ngalkò	'better that...'

The influence of Kanuri has been much greater on Bade-Ngizim as a group than on the Bole-Tangale languages, suggesting that many of these words entered Bade-Ngizim while the languages of the group still occupied a contiguous geographical zone separate from the Bole-Tangale languages. (See Schuh 2003 for discussion of Kanuri loanwords in Bade-Ngizim.) The table below shows percentages of identified Kanuri loanwords in currently available lexicons:

Table 1: Percentages of Kanuri loanwords in the Yobe languages[7]

Western Bade:	12%	Karekare:	9%
Gashua Bade:	15%	Ngamo:	4.6%
Ngizim:	15.6%	Bole:	7%

There is evidence that Bole has borrowed directly from Kanuri rather than introducing Kanuri loanwords via the Bade-Ngizim languages. For example, Bade-Ngizim languages adapt Kanuri verbs by borrowing the Kanuri verbal noun, which has a *t* suffix, then adding their own inflections. Bole does not have a single pattern for borrowing Kanuri verbs, but the most common pattern is to add a suffix *n* to the root.[8] Compare loans such as the following:

(3)   Patterns of Kanuri loan verbs in Ngizim and Bole

Kanuri verbal noun	Ngizim verb	Bole verb	
njès-tə	ngèstu	ngèsùnu	'be late'
wàa-tə	wàakàatu	wàkàanu	'happen'

Karekare and Ngamo, however, seem to have introduced most Kanuri loanwords via Ngizim and Bole respectively. Compare Karekare *ngèstu* 'be late', with *-t-* as the final consonant, as in Ngizim in (3), vs. Ngamo (G) *ngèsnâ* 'be late', with *-n-* as the final consonant, as in Bole in (3). Likewise, Karekare *bìlân* 'beautiful' < Kanuri *bəlân* is shared with Ngizim and Bade but not with Bole or Ngamo, whereas Ngamo *gède* 'different'< Kanuri *gàde* is shared with Bole but not with Ngizim or Bade. The introduction of Kanuri loanwords into Karekare from Ngizim and into Ngamo from Bole is part of larger lexical and cultural borrowing patterns within the respective language pairs.

1.3.   Hausa as a Recent Source of Loanwords

The influence of Hausa provides a useful contrast with Kanuri. The previous section argued that one of the *Sprachbund* aspects of the Yobe languages is the result of a long period of Kanuri influence. Though Hausa has had, and continues to have, massive influence on the Yobe languages, its influence cannot be viewed as *areal* in nature. As is the case throughout northern Nigeria, everyone in Yobe State, regardless of heritage language, speaks Hausa. Younger people, aside from those who have grown up in very rural areas, generally seem to feel more comfortable speaking Hausa than their heritage languages, and in conversations

---

[7] The figures, esp. for the Bade-Ngizim languages, are probably too low, first, because I am not an authority on Kanuri and have certainly failed to identify many Kanuri loans, and second, because the currently available lexicons of Yobe languages do not include extensive numbers of items in lexical arenas which would consist largely of Kanuri loans, such as religious, legal, and philosophical concepts.

[8] I do not know the source of the *n* suffix. There is no native verb derivation suffix of this form in Bole. It may come from one of the complex set of Kanuri inflectional suffixes.

between speakers of any of the Yobe languages, code-switching between that language and Hausa is the standard mode of speech. Not surprisingly, all the Yobe languages have incorporated many Hausa loanwords. The table below shows percentages of identified Hausa loanwords in currently available lexicons:

Table 2: Percentages of Hausa loanwords in the Yobe languages

Western Bade:	3.8%	Karekare:	5.6%
Gashua Bade:	5%	Ngamo:	3.7%
Ngizim:	4.2%	Bole:	7.1%

These figures are surely on the low side for all the languages because at this stage of research, there has been no effort made to collect words in lexical arenas where virtually all words are Hausa loanwords, such as imported trade goods, modern occupations, modern media and communications, modern mechanical objects, and the like. Many Hausa loanwords, like Bole *àyàbà* 'banana' or *kèeke* 'bicycle', themselves originally loanwords into Hausa, refer to introduced items that would have not had traditional names, and in many cases where a Hausa loanword refers to something that would have preexisted contact with Hausa, it is hard to know whether it is really an integrated lexical item or a sort of code-switch, introducing a Hausa word even though a native word exists. For example, is Ngizim *kûȓkudù* 'sandhopper' (Hausa *kûȓkudìu*) the word that all Ngizims would now use for this insect or has the speaker who provided the word just forgotten (or never heard) the native word?

Even allowing for the fact that the figures in Table 2 would be larger with more complete lexicons, two facts of interest emerge: (1) the low numbers compared to the percentages of Kanuri loanwords seen in the previous section, especially in Bade-Ngizim, and (2) the uniformity of the percentages across the languages (the Bole number is skewed high because the currently available lexicon for this language is the largest). The relatively low numbers as a percentage of vocabulary must be a result of the lengths of the contact periods of the Yobe languages with Kanuri vs. Hausa. While continuous contact with Kanuri must have lasted over many centuries, intensive contact with Hausa probably does not extend back much more than a century, and the linguistic dominance that Hausa exerts today would extend over even less time. The uniformity of the percentages must have to do with the way Hausa loanwords are being introduced. Everyone speaks Hausa, and every language is independently introducing Hausa loanwords through the same processes, viz. frequent code-switching and the universal practice of using Hausa words when "native" words are not readily available, either because the concepts are new or because the speakers cannot immediately access existing native words.

In short, though the lexicons of Yobe languages have been and are being shaped by the introduction of loanwords from non-Yobe languages, the legacies are quite different. Kanuri loans have been, for the most part, integrated and

adapted such that they give the Yobe languages a regional flavor not found in related languages outside the region. Hausa loans are a recent and relatively superficial, albeit highly prominent part of the lexicon that does not differentiate the Yobe languages from virtually any other minority language of northern Nigeria.

2. Morphology: Gender and Feminine Gender as Default

Proto-Chadic inherited a three-way noun classification system from Proto-Afroasiatic: *masculine* singular, *feminine* singular, and common gender *plural*. At least some languages in all the major sub-branches of West Chadic retain this system. In Yobe State, gender remains a robust lexical feature of nouns in Western Bade and in Ngamo, as shown by the demonstrative agreement patterns in (4). Other gender/number sensitive morphemes in these languages are genitive linkers and personal pronouns.

(4)   Grammatical gender agreement in Western Bade and Ngamo

	Bade (W)	Ngamo (G)	
masculine	gwàmaa-<u>mso</u>	gâm <u>wòye'e</u>	'this ram'
feminine	tə̀màku-<u>mco</u>	tèmshi <u>wònse'è</u>	'this ewe'
plural	tə̀màkùnə̂-<u>mdo</u>	tèmkà <u>màaye'è</u>	'these sheep'

Gender as an idiosyncratic property of individual lexical items is in various states of decay in other Yobe languages. Gashua Bade still retains gender agreement in its demonstrative system, genitive linkers, and pronouns, but aside from nouns with inherent natural gender (humans, some domestic animals), choice of agreement is shifting toward predictability on the basis of phonological form of the noun—essentially, nouns ending in *-i*, *-u*, *-a* take feminine agreement and others take masculine, though there is fluctuation, especially in vowel-final nouns (Schuh 1977). Bole differentiates gender/number agreement forms in demonstratives and pronouns, though not genitive markers. Most nouns other than those with inherent natural gender take masculine agreement, but nouns in certain semantic groups have specific genders, e.g. fruit-bearing trees are feminine but the fruits are masculine (Gimba 2005). Karekare demonstratives are not gender sensitive, but Karekare does distinguish gender in personal pronouns and in genitive linkers used in N+N genitives, e.g. *mizi <u>ma</u> Jilwaye* 'the husband <u>of</u> Jilwaye' vs. *men <u>ta</u> idihu* 'the wife <u>of</u> the king'. More investigation is needed to discover whether there are consistent patterns of gender agreement, but for many nouns agreement fluctuates. Ngizim has lost gender as a lexical category. There is only one set of demonstratives, one paradigm for genitives, and pronoun agreement is like that of English, with regular masculine/feminine agreement only for referents with inherent natural gender.

Despite this range of systems related to gender, one striking Yobe areal feature emerges: THE DEFAULT GENDER IS FEMININE. This claim is supported by

the languages where gender remains a robust lexical category and, ironically, by languages that are losing or have lost this lexical property.

Evidence for feminine as the default gender in Western Bade and Ngamo, the languages with robust systems of lexical gender, comes from loanwords. As noted in section 1.2, the most significant source for loanwords in Yobe languages until fairly recent times has been Kanuri. Kanuri does not have grammatical gender. Hence, languages that do categorize nouns for gender must have strategies for assigning gender to nouns borrowed from Kanuri. Semantic properties play a key role, in particular when the noun is human and hence has inherent natural gender. For example, Western Bade *kàadùkùmaan* (m) 'messenger' (Kanuri *kàdùnomà*) would normally refer to a role played by a male whereas *duukəramə́n* (f) 'pot maker' (Kanuri *duwùràm* [*njèmâ*]) refers specifically to women who make pots. In Ngamo (G) *mà'ì* (m) 'king' (Kanuri *mâi*) refers to a position always occupied by males whereas *kilaakì* (f) 'prostitute' would refer to a female (the latter is ultimately from English 'clerk' and may have come into Ngamo via Hausa, though it is also used in Kanuri). In Western Bade, all mass nouns take plural agreement (cf. *sàasáà-mdo* 'this meat' with the plural noun in the table above), and as expected, loanwords with mass referents fall into this category, e.g. *bərbərən* 'dust' (Kanuri *bərbər*). Sometimes gender has been assigned by semantic association with another noun, e.g. *zèenaan* (m) 'gourd ladle' (Kanuri *jènyi*) may be masculine because of association with the native word *maɗakwáan* (m) 'gourd cup'.[9]

For most borrowed nouns, however, natural gender, semantic association, and/or phonological form do not determine assignment of lexical gender. This is true both for loanwords from Kanuri, which does not have grammatical gender, and from Hausa, which does have grammatical gender. As might be expected, Hausa gender does sometimes coincide with gender assigned by the borrowing language, but it is just as often overridden, e.g. Hausa *goõr̃ò* (m) 'kola nut' borrowed as Western Bade *goorón* and Ngamo (G) *goorò*, which are feminine in the respective languages. Overriding Hausa feminine with masculine in the borrowing language does take place, e.g. Hausa *maaraa* (f) 'food scoop made of a calabash fragment' borrowed as Western Bade *maaraan* (m), but this is less common than overriding masculine with feminine. Counts of borrowed nouns according to gender assignment clearly show the skewing of gender assignment toward feminine. In Table 3, the Bade figures exclude mass nouns (see above) and nouns with paired masculine and feminine counterparts, such as *àapənon* 'Hausa man', *àapənàakon* 'Hausa woman' (Kanuri *àafùno* for either gender). The overall greater numbers in Bade reflect the fact that the currently available lexicon for Bade is about twice as big as that for Ngamo.

---

[9] In providing gender for loanwords, Bala Dagona Wakili, the primary Western Bade speaking participant in this project, would often say something like, "It has to be masculine (or feminine) because X is masculine (or feminine)." Unfortunately, I seem not to have cited any such comments in my notes and cannot recall specific cases.

Table 3: Counts of gender assignment for loanwords

	Western Bade		Ngamo (G)	
	Kanuri	Hausa	Kanuri	Hausa
Masculine	58	28	15	13
Feminine	98	59	32	27

A few typical examples of borrowed nouns that have been assigned feminine lexical gender show that there is no obvious correlation between form or meaning and gender.

(5)   Some loanwords classified as feminine in Western Bade and Ngamo (G)

Kanuri	Bade (W)		Kanuri	Ngamo (G)	
àshîr	àasirən	'secret'	kàshaàr	kàskâr	'sword'
bəndəgə̀	bə̀ndəgín	'gun'	jìnadə̀	jìnaadi	'flint stone'
r̰iwà	riipan	'carrion'	sərgə̀	sìrkâ	'poison'
Hausa			Hausa		
dàliilìi (m)	dàliilín	'reason'	hankàlii (m)	hànkalì	'sense'
goor̰ò (m)	goorón	'kola nut'	goor̰ò (m)	goorò	'kola nut'
makaranta (f)	makarantan	'school'	àyàbà (f)[10]	àyàbà	'banana'

Evidence for the default nature of feminine gender in languages that have lost gender as an idiosyncratic feature of lexical items is more equivocal, but at the very least this evidence shows that the automatic default is not masculine. As languages lose a multi-parameter category such as gender, they must decide how to dispose of the marking system that they have inherited from the time when the category was active. In the case of Ngizim, we can see what decisions were made by comparing Ngizim and Gashua Bade, the Bade dialect that most closely resembles Ngizim. The demonstrative systems of Bade and Ngizim have three parameters comprising 'near' (= proximal), 'far' (= distal), and a third parameter, translatable as something like "that very one mentioned." Comparing the forms in (6), it is evident that Ngizim, having lost lexical gender as an active category, has extended the original FEMININE forms to all nouns, including plurals. (We return to the fate of the original masculine forms in a later section.) It appears that as lexical gender began to shift to natural gender in Ngizim (as is now the case in pronoun agreement), nouns where gender was not an inherent feature were assigned what was felt to be the unmarked or default form—the feminine. Eventually, the erstwhile feminine became so overwhelmingly frequent that it was extended to nouns with referents of masculine natural gender and even plurals.

---

[10] The word for 'banana' is borrowed into Western Bade as *àyàbaan*, which is categorized as a mass noun.

(6)   Bade and Ngizim evidence that Ngizim has extended feminine forms

	Gashua Bade			Ngizim		
	Near	Far	Mentioned	Near	Far	Mentioned
Masculine 'bull'	kwàm-aau	kwàm-áàni	kwàm-áànau	kwàm-tku	kwàm-tiiwú	kwàm-tənu
Feminine 'cow'	tlà-tku	tlà-tiiwú	tlà-tənáu	tlà-tku	tlà-tiiwú	tlà-tənu
Plural 'cattle'	ùktl-áàndau	ùktl-áàndiiwú	ùktl-áàndənáu	tlàaɗin-tku	tlàaɗin-tiiwú	tlàaɗin-tənu

With pronoun possessors, Bade has a three-way system for marking genitives, one for masculine and plural possessed nouns, one for feminine possessed nouns, and a non-gender marked form used with objects of nominalized verbs, prepositions, and a handful of "inalienable" nouns, such as *ŋgwa* 'household' (both languages have a single non-gender sensitive system for marking for N+N genitives). Ngizim has only one method of marking pronominal genitives. The table below shows 2nd feminine singular and 3rd masculine singular possessive pronouns, whose morphology is representative of other pronouns in the system.[11]

(7)   Extension of non-gender sensitive genitive to all nouns in Ngizim

Gashua Bade			Ngizim		
'son'	'daughter'	VN 'catching'	'son'	'daughter'	VN 'catching'
wun-n-ə̀m	wunya-tk-ə̀m	gə̀f-aa-gə̀m	wun-kə̀m	wuny-aa-kə̀m	gaf-aa-kə̀m
wul-lì /wun-lì/	wunya-tkə-lì	gə̀f-aa-lì	wùn-gəɽì	wunyà-gəɽì	gafa-gəɽì

A comparison of these paradigms shows that the Ngizim genitive most closely resembles the non-gender marked Gashua Bade forms, exemplified with the verbal noun+object. The 2nd feminine singular pronoun is virtually identical in the two languages, modulo some minor phonological differences, and both languages insert a long *-aa-*, rather than a gender-sensitive linker, when the first noun ends in a vowel. The 3rd masculine singular form in Ngizim has a formative *-gə-*, which is the common gender linker seen with nouns. Notably, Ngizim lacks either of the gender sensitive linkers *-n-* and *-tk-* seen in Gashua Bade. In short, in losing grammatical gender, Ngizim abandoned all erstwhile gender-sensitive genitive linkers.

Karekare has moved in a somewhat different direction with respect to its genitive linkers and demonstratives. As noted above, Karekare has linkers *ma* and *ta* in N+N genitive constructions that agree with at least referents having masculine and feminine natural gender, respectively. This agreement pattern is

---

[11] See Schuh (1977) for a description of genitive systems of languages of the Bade-Ngizim group.

inherited from Proto-Bole-Tangale and comparative evidence shows that at one time it also functioned in constructions with pronoun possessors, as it does today in a number of Bole-Tangale languages—cf. Kirfi *kaɣala mi-shi* 'your (f) ram' vs. *wùcci shi-shi* 'your (f) female goat' (Schuh 1978b:35). Karekare, however, has only a single paradigm of linked possessive pronouns for masculine, feminine, and plural referents.[12]

(8) Extension of feminine genitive pronoun forms to all nouns in Karekare

Possessor	'husband'	'wife'	'wives'
1 sg.	mìzìi hǹno < *tìno	mèn hǹno < *tìno	mendèe hǹno < *tìno
2 m. sg.		mèn tìko	mendèe tìko
2 f. sg.	mìzìi ci < *tìci		

but cf. nominal genitives

mizi ma Jilwaye	men ta idihu
'the husband of Jilwaye'	'the wife of the king'

On the other hand, independent genitive pronouns use the original MASCULINE forms, regardless of referent, viz. *mìno* 'mine', *mìko* 'yours (m. possessor)', *mìci* 'yours (f. possessor)', etc.

Formatives **m-* masculine and **t-* feminine also served as bases for demonstratives in Proto-Bole-Tangle—cf. Kanakuru *lowòi me* 'that' boy' vs. *gunyòi she* 'that' girl' (Newman 1974:87). Karekare uses one gender/number neutral base *-m-*, undoubtedly from the original masculine, though in Karekare demonstratives have taken on a different appearance from those of Kanakuru, which probably looks more like the original:

(9) Karekare demonstratives

Near	Far			Forms with gender marked genitives
koor-aayam	kooroo âm	'donkey'		cf. *kooro mà rìya* 'donkey of the bush'
kwàr-àayam	kwàrà âm	'house'		cf. *kwar ta ba'ato* 'house of her father'
kwarcin-àayam	kwarcinòo âm	'houses'		

To summarize, the strongest claim would be that the default gender for Yobe languages is FEMININE, which runs against the more common tendency of languages with grammatical gender to take masculine as default. Evidence from

---

[12] By a regular phonological rule, the sequence *ti-* becomes a nasally released [tⁿ] before *n*V. This is represented as *hn-* in the first person possessives (cf. *tì-kau* 'he ate' vs. /tìnà/ ➔ *hn-nà* 'eating'). I describe this phenomenon in "Shooting through the nose in Karekare," available for download at the website mentioned in footnote 1. Probably as part of the same phenomenon, the **ti-* of the possessive pronoun has disappeared when the consonant of the pronoun is a coronal—cf. *kaa tìko* 'your (m) head', *kàa cì* 'your (f) head', *kàa to* 'her head', *kàa tǐmu* 'our head', *kàa su* 'their head'.

languages that have lost grammatical gender as an active idiosyncratic lexical parameter shows that an absolute claim for feminine being default is probably too strong, but these languages have generalized originally feminine morphology in ways that clearly show that masculine is NOT the default.

3.   Syntax
3.1.   Marking of Conditional Clauses
The Potiskum area Yobe languages mark conditional clauses in the following ways (... = the position of the clausal proposition).

(10)   Marking of "if/when" conditional clauses in Potiskum area languages
    Bole: bàa...(ye)
    Ngamo (G): na...(-i)
    Karekare: ...ya/ye
    Ngizim: ...-n/nən

That is, Bole and Ngamo have clause-initial markers translatable as 'if/when' and optional clause final markers. Karekare and Ngizim have only clause-final markers. The examples in (11) illustrate most of the options.[13]

(11)   Conditional clauses in Potiskum area languages
Bole:

Bàa ka ngora sootà   ye,   ngòrii gà bòo pàtà zònge.
if   you tie.FUT lie   "if"   tie on end tail hyena
    'If you are going to tie up a lie, tie on the end of a hyena's tail.' (because the hyena will run off to the bush with it)

Ngamo (G):

Ngoi na an-ko ta a go-nni ɗala dala-i ngap-ni hata-n-ni.
person if tell you quote AUX going-he swallow axe-"if" hold-for him handle-of-it
    'If a person says he is going to swallow an axe, hold its handle for him.'

Karekare:

Amu taɓka ya, eko kuma bai.
water spill.CPL "if" do.CPL collecting not
    'If water spills, it can't be collected back together.'

Ngizim:

Nən ma a ntana gawa-n, a-gaf-ici gəji-u.
person QUOTE AUX swallow.INCPL axe-"if" hold.IMPER-for him handle-the
    'If a person says he is going to swallow an axe, hold its handle for him.'

---

[13] For additional examples, see an extended version of this paper at the website mentioned in footnote 1.

Russell G. Schuh

The conditional clauses in (11) exemplify what I call *imperfective* discourse in Schuh (1998:165), i.e. discourse that refers either to events that have not yet taken place or that apply generically. Conditional clauses in imperfective discourse in Yobe languages, and Chadic languages in general, can be translated as English 'if' or 'when' depending on the certainty of the event's taking place. There is no formal difference corresponding to the English translations.

"When" clauses also appear in *perfective discourse*, i.e. discourse which relates discrete completed events, as in a historical text, a story, or a report. These clauses use the markers in (12), illustrated in (13).

(12)   "When" clauses in perfective discourse in Potiskum area languages
    Bole: ...(ye)
    Ngamo (G): No external marking of "when" clauses in perfective discourse.
    Karekare: ...(ma)
    Ngizim: ...(tənu/ngum)

(13)   "When" clauses in perfective discourse
Bole

Kòɓam	màala	ye,	ita	bòli	adà à	gàa	mòcci.
head to.CPL	bush	"when"	she	find.SJN	dog at	inside	locust bean

'When she headed to the bush, she found the dog up in a locust bean tree.'

Ngamo (G)

Turum	ha'ako,	sai	a-goptu	bo gaba	sot!
lion	eat.CPL	then	beat.SJN-VENT	cough	*sot*

'When the lion had eaten (the medicine), then he emitted a cough *sot*!

Karekare

Fati	ngataka	min	tan	waɗa	ma,	sai	rasu	a	benu	tiɗsu.
sun	fall.CPL	people	eat.CPL	food	"when"	then	enter.SJN	in	house	lie.SJN

'When the sun had gone and the people had eaten, then they entered the house and lay down.'

Ngizim

Ɗa	tlanu	tənu,	ja	da	nai.
town	dawn.CPL	"when"	dog	AUX	come.SJN

'When the area dawned, the dog came.'

That is, "when" clauses in perfective discourse either have no overt marking at all or have a clause-final marker. The principal marking of clausal relations is in the verb aspect choice. Sequential clauses in all these languages use the subjunctive (SJN) in both perfective and imperfective discourse. The completive (CPL) conveys posteriority. Note in the examples in (13), the "when" clauses all have a completive verb whereas the "main" clauses have a subjunctive. The effect of the completive is to take a step back in the sequence, often repeating the immediately preceding action, while the subjunctive in the main clause is the next

event in the story sequence. A translation closer to the Yobe structure would thus be (illustrating with the Bole example), "She had headed to the bush (and) she found the dog up in a locust bean tree," where the pluperfect ("when") clause sets up a context for the next event in the sequence in the same way that a conditional clause in imperfective discourse sets up a context for the consequent clause.

Summarizing the situation for "conditional" clauses in perfective and imperfective discourse, all the Yobe languages utilize at least a clause-final marker (with the apparent exception of Ngamo in perfective discourse "when" clauses, where sequence of tenses is enough),[14] some require a clause-initial marker in "if/when" clauses in imperfective discourse, and none use a clause-initial marker in "when" sentences in perfective discourse (again, with the exception of Ngamo, perhaps dialectally—see footnote 14).

Other types of subordinate clauses in these languages all use clause-initial conjunctions, e.g. "before" and "after" clauses, "purpose" and "reason" clauses, and relative clauses. One thus wonders what the source of the POST-clausal markers in conditional clauses is. The answer is that all the post-clausal markers come from definite determiners, some of which are still in use in the respective languages, and some of which are no longer used as definite determiners, but which comparative evidence, even at a shallow level, reveals the source to be definite determiners.

(14)   Definite determiners as the source of clause final conditional markers

	Clause type	Clause marking	Determiner source
Bole	'if/when' 'when'	bàa...(ye) ...(ye)	Definite article: *tèmshi yê* 'the sheep' (same)
Ngamo (G)	'if/when' 'when'	na...(-i) (no marking)	Definite article: *tèmshis'è* 'the sheep' (cf. *-i* in Yaya *tèmshì'i* 'the sheep')
Karekare	'if/when' 'when'	...ye/ya ...(ma)	Definite article: *lo-yi* 'the meat' Demonstrative: *kwàrà 'âm* 'that house' (cf. Kanakuru *gamii mè* 'this ram')
Ngizim	'if/when' 'when'	...-n/nən ...(tənu/ngum)	< Masculine proximal demonstrative: cf. Bade (G): *kwàm-áàni* 'that bull' "Known" dem.: *tlà tənu* 'that cow' cf. definite article: *soonòo-gu* 'the shoe'

The Yobe languages have grammatically instantiated a semantic connection discussed in Schlenker (2004) between definite descriptions and conditionals. Schlenker says (p. 448), "*If* [and *when* as a marker of conditionals—RGS] is

---

[14] This is the case in several dozen examples drawn from folktale texts in the Gudi dialect of Ngamo. The Yaya dialect consistently uses a clause initial maker *ido*, e.g. Ngamo (Y) *Ido Ba Zenge nduno'o, nzuba sai fani-ni*. 'When Brother Hyena came(PF), fear then overcame-him(SJN).' Elicited data that I collected in the mid-1970's, and which I believe was from the Gudi dialect (I knew nothing about Ngamo dialect distinctions at the time), also shows clause-initial *ido*, e.g. *ido ndano, yak nzi nam ula-su* 'when (they) came(PF), then they collected(SJN) their-goods'.

simply the form taken by *the* when it applies to a description of worlds." Schlenker's paper provides a formal semantic comparison of definite descriptions and conditionals that goes beyond the scope of the present discussion. Suffice it to say that an expression marked by a definite article (or demonstrative) refers to an *entity* that is most salient to the domain of discourse whereas a conditional clause refers to a *world* that is most salient to the domain of the consequent clause or, in perfective discourse, to the domain of the next event in the narrative.

The source of conditional clause marking in definite determiners also explains another fact about conditional marking in these languages, viz. the fact that the clause-final markers are not obligatorily present (with the exception of Karekare and Ngizim "if/when" clauses). In Yobe languages, and Chadic languages in general, determiners are never required by grammatical rules. It is a speaker's choice to overtly mark a noun as "known" or "new." The grammatical optionality is reflected in conditional clause marking. Karekare and Ngizim do require clause-final markers in "if/when" clauses. This seems to be a case of grammaticalization, probably as a way to assure that imperfective and perfective conditionals be kept separate. Comparative evidence from languages like Kanakuru and Tangale suggests that at least the Bole-Tangale group of languages originally used a clause-initial conjunction meaning "if/when," as is still the case in Bole and Ngamo. Loss of clause-initial marking in such clauses has either been facilitated or has been compensated for by requiring clause-final marking.

3.2. WH Question Formation and Marking of Questioned Subjects

A characteristic feature of all the Yobe languages is the following positioning of WH question words:

(15)   Word order in questions in Yobe languages
        Questioned non-subjects: in situ
        Questioned subjects: post-VP

In (16) are two questions for each language, the first with a questioned object, the second with a questioned subject. "Neutral" word order for all Yobe languages is rigid SVO. Like direct objects, other questioned postverbal constituents such as indirect objects or instrumentals occupy the position they would take in statements. Focused constituents, such as constituents that answer WH question words, have the same positioning and additional morphological marking, if any, as the corresponding WH words.

(16)   Questioned objects and questioned subjects in Potiskum area languages
Bole

Sauna	gojjuu	le?				'What did Sauna buy?'
Sauna	buy.CPL	what				
Gojjitu	sooto	in	Sauna	ye	lo?	'Who sold bean cakes to Sauna?'
sell.CPL	bean cake	to	Sauna	Q-SBJ	who	

Ngamo (G)

Sauna	kaja	miya?			'What did Sauna buy?'
Sauna	buy.CPL	what			

Bo'ota	soto	ki	Sauna-i	lo?	'Who sold bean cakes to Sauna?'
sell.CPL	bean cake	to	Sauna-Q-SBJ	who	

Karekare

Na	'yu	miya?		'What have I done?'
I	do.CPL	what		

Tukà	wàɗa-yì	nà	láà?	'Who ate the food?'
eat.CPL	food-the	Q-SBJ	who	

Ngizim

Sauna	mase	tam?			'What did Sauna buy?'
Sauna	buy.CPL	what			

Dəbɗə	are	i	Sauna-n	tai?	'Who sold bean cakes to Sauna?'
sell.CPL	bean cakes	to	Sauna-Q-SBJ	who	

This syntactic pattern, with in situ questioned non-subjects but post-VP questioned subjects, is not unique to the Yobe languages. It extends along the entire eastern edge of the West Chadic speaking region, from Bade and Duwai in the north to Tangale and Kanakuru in the south. Within Yobe State, the language to the immediate west of the Yobe languages is Hausa, which fronts all WH question words. Further south, the Chadic languages west of "post-VP subject" languages, such as Kirfi and Galambu of the Bole-Tangale group (Schuh 1978b), have in situ order for all WH questions including subjects.

Although the word order pattern illustrated above is not unique to the languages of Yobe State, there is a feature that IS unique to the Potiskum area languages, namely morphological markers (boxed in the examples) of postverbal questioned subjects. Kanakuru (Newman 1974) and Tangale (Jungraithmayr 2002), for example, have postverbal questioned subjects but no marking other than position, e.g. Tangale *poduk sóbók-no nóŋ*? 'who has taken my slaves out?', lit. "removed slaves-my who" (Jungraithmayr 2002:14).

The resemblance of the post-VP subject makers to the markers of conditional clauses discussed in §4.1 is immediately obvious. Only Karekare departs from the pattern. Assuming that conditional markers and post-VP subject markers all have their source in definite determiners, the semantic motivation for using these markers with questioned and focused subjects emerges. The post-VP subject marker is actually in constituency with the clausal material that precedes it, and that material refers to a *presupposed proposition* that is relevant to the variable expressed by the WH word or to the value assigned in place of the variable, i.e. the referent that answers the WH question or a subject that takes pragmatic focus.

The source of the Karekare post-VP subject marker *na* is not clear at the moment. A definite determiner of the form "*n*" is widespread in West Chadic, but not in the Bole-Tangale languages. It is possible that the Karekare *na* derives from *ma* (see the data on "when" clauses above) or, alternatively, it may be borrowed

from the formative *-n/nən* of Ngizim, which has heavily influenced Karekare in other ways.

References

Bargery, G.P. 1934. *A Hausa-English Dictionary and English-Hausa Vocabulary*. London: Oxford University Press.

Gimba, Alhaji Maina. 2005. Gender in Bole. Ms., University of Maiduguri.

Jungraithmayr, Hermann. 2002. *Síndi: Tangale Folktales (Kaltungo, Northeastern Nigeria)*. Westafrikanische Studien Band 23. Köln: Rüdiger Köppe Verlag.

Newman, Paul. 1969/70. Linguistic relationship, language shifting, and historical inference. *Afrika und Übersee* 53:217-223.

Newman, Paul. 1974. *The Kanakuru Language*. West African Language Monographs 9. Leeds: Institute of Modern English Language Studies, University of Leeds and West African Linguistic Society.

Newman, Paul. 1977. Chadic classification and reconstructions. *Afroasiatic Linguistics* 5:1-42.

Schlenker, Philippe. 2004. Conditionals as definite descriptions (a referential analysis). *Research on Language and Computation* 2:417-462.

Schuh, Russell G. 1977. Bade/Ngizim determiner system. *Afroasiatic Linguistics* 4:101-174.

Schuh, Russell G. 1978a. Bade/Ngizim vowels and syllable structure. *Studies in African Linguistics* 9:247-283.

Schuh, Russell G. 1978b. *Bole-Tangale Languages of the Bauchi Area (Northern Nigeria)*. Marburger Studien zur Afrika- und Asienkunde Serie A:Afrika Band 13. Berlin: Dietrich Reimer Verlag.

Schuh, Russell G. 1981. Using dialect geography to determine pre-history: a Chadic case study. *Sprache und Geschichte in Afrika* 3:201-250.

Schuh, Russell G. 1998. *A Grammar of Miya*. University of California Publications in Linguistics Volume 130. Berkeley, Los Angeles and London: University of California Press.

Schuh, R.G. 2003. The Linguistic influence of Kanuri on Bade and Ngizim. *Maiduguri Journal of Linguistic and Literary Studies (MAJOLLS)* 5:55-89.

Russell G. Schuh
Department of Linguistics
UCLA
Los Angeles, CA 90095-1543
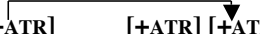
schuh@humnet.ucla.edu

Ultra Long-Distance ATR Agreement in Wolof

MARIAME SY
University of California, Los Angeles

0.   Introduction

Wolof (West Atlantic) exhibits a pattern of vowel harmony which involves the categorization of all the vowels in the language into two harmonic sets with regard to the feature [ATR] (advanced tongue root). Though the domain of harmony is usually described as being the word, in Wolof this domain extends to syntactic phrases. This fact, in and of itself, is neither unusual nor particularly exciting. Indeed, there are many African languages in which phonological processes apply at the phrasal level. What is exceptionally interesting about Wolof is that in some cases the process of vowel harmony appears to apply across syntactic phrases, showing patterns to which we refer as "ultra-long distance [ATR] agreement" (henceforth ULDAA), as shown in (1).

(1)   a. **[-ATR]   [-ATR]   [-ATR]**   b. **[+ATR]   [+ATR] [+ATR]**
       xaj b- u   weex   **b- ale**   **béy** w- u   réy   **w- ëlé**
       dog CL-REL be.white CL-DEM.DIST   goat CL-REL be.big CL-DEM.DIST
       'that white dog'   'that big goat'

Although it seems reasonable to assume that the final demonstratives in the examples in (1) harmonize with the head nouns which they modify, nothing tells us that the trigger of harmony is not the immediately preceding adjectival verbs. In fact, since these verbs are specified for the [ATR] feature and thus cannot incur any harmony effects from the noun, it seems more reasonable, from a phonological point of view, to assume that the verbs are the actual triggers of harmony on the demonstratives. Then, the question immediately arises as to which of these assumptions is in actual fact correct. The first part of this article reports on the findings of an experimental study designed to answer this question.

   Given the fact that nouns and adjectival lexemes carrying opposite [ATR] values may combine within the same structure, such disharmonic strings naturally come to mind as the perfect testing ground. Examples of this type of disharmonic DP are given in (2) below.

(2)[1] a. **[-ATR]**   [+ATR] **[-ATR]**   b. **[+ATR]**   [-ATR]   **[+ATR]**
       xaj b- u   réy **b- ale**   **béy** w- u   weex **w- ëlé**
       dog CL-REL be.big CL-DEM.DIST   goat CL-REL be.white CL-DEM.DIST
       'that big dog'   'that white goat'

The results of an acoustic analysis comparing harmonizing determiners in DPs such as those in (2) to the same determiners in fully harmonic DPs such as those in (1) are presented. These results support the claim that ULDAA occurs in relative clauses.

    Accordingly, the second part of this study considers various approaches to accounting for such phenomena. We argue that neither a purely phonological account nor a purely syntactic one adequately handles the data, and propose an analysis within the framework of Optimality Theory.

1.   Background
1.1.   The Wolof Vowel System

Wolof vowels can be divided into two mutually exclusive harmonic sets with regard to the feature [ATR], as shown in (3).

(3)[2]   [+ATR]   [-ATR]
            i   ii   u uu
            é éé   ë[3]   ó óó   e   a   o oo
                                   à/aa

As apparent from the data in (3), high vowels lack [-ATR] counterparts, and low vowels lack [+ATR] counterparts, restricting the set of alternating vowels to mid vowels. Tongue root harmony in Wolof is generally described as a progressive (left-to-right) process (Ka 1988, Archangeli and Pulleyblank 1994, and others) as evidenced in (4) below.

(4)   a. liit   -ël -ënté -ëndóó -wóón b. jàng -al   -ante -andoo -oon
        play.flute -BEN -REC -COM -PAST   read- BEN -REC -COM -PAST
        'played the flute together for each other'   'read together for each other'

    Although high and low vowels equally contribute to producing disharmonic forms, they do so in rather opposite ways. In word-initial position, both types of vowel induce harmony on the vowels located to their right. However, in word-

---

[1] Abbreviations: BEN=benefactive; CL=class; COM=commutative; DEM=demonstrative; DIST=distal; GEN=genitive; INDEF=indefinite; PREP=Preposition; PROX=proximate; REC=reciprocal; REL=relativizer.

[2] For convenience we use Wolof orthography to represent these sounds rather than the conventional diacritics for the [ATR] feature.

[3] This vowel has no long counterpart in the native vocabulary; the few forms in which it appears are all loanwords.

medial position, the two types vowels behave differently. High vowels in this position display typical transparent behavior, i.e. they neither incur harmony from vowels located to their left, nor do they trigger harmony on those located to their right.

(5)   a. bind-óón-në 'wrote-3sg.'   b. soppi-woon-na 'changed-3sg.'
        tuur-óón-në 'spilled-3sg.'   teeru-woon-na 'welcomed-3sg.'

The examples in (5a) show that the past and perfective affixes surface respectively as [+ATR] *-óón* and *-në*, harmonizing with the initial high vowel stems with which they occur. Yet in (5b), where the stems contain medial high vowels and [-ATR] initial vowels, these same affixes surface as [-ATR] *-oon* and *-na*, incurring no harmony effects from the immediately preceding high vowels.

    Medial low vowels, on the other hand, never allow spreading of the opposite tongue root feature across their own, but instead initiate their own harmony domain to their right, thus showing typical opaque behavior, as evidenced in (6); the past and perfective affixes surface as [-ATR] in both (6a) and (6b), although the stems in (6b) contain [+ATR] initial vowels.

(6)   a. tàkk-oon-na   'caught fire' b. liyaar-oon-na   'hand-gestured'
        laal-oon-na   'touched'   nguufaan-oon-na 'carried in arms'
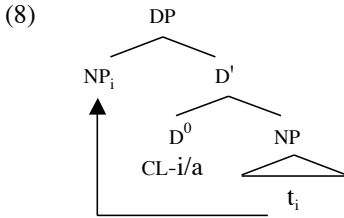   sang-oon-na   'covered'   misaal-oon-na   'gave examples'

1.2.   DP Structure

Wolof is, in the general case, a head-initial (SVO) language. However, within DP heads can be either initial or final; that is, indefinite determiners occur pre-nominally while definite determiners and demonstratives appear in post-nominal position. All of these elements agree with the class and/or [ATR] feature of the noun, as illustrated in (7).[4]

(7)   a. (a-   w) fas   b. fas   w- i   c. fas   w- a
        INDEF-CL horse   horse CL-DEF.PROX   horse CL-DEF.DIST
        'a horse'   'the horse'(proximal)   'the horse' (distal)
      d. (w-enn)   fas   e. fas   w- ii   f. fas   w- ale
        CL-one   horse   horse CL-DEM.PROX   horse CL-DEF.DIST
        'a/one horse'   'this horse'   'that horse'

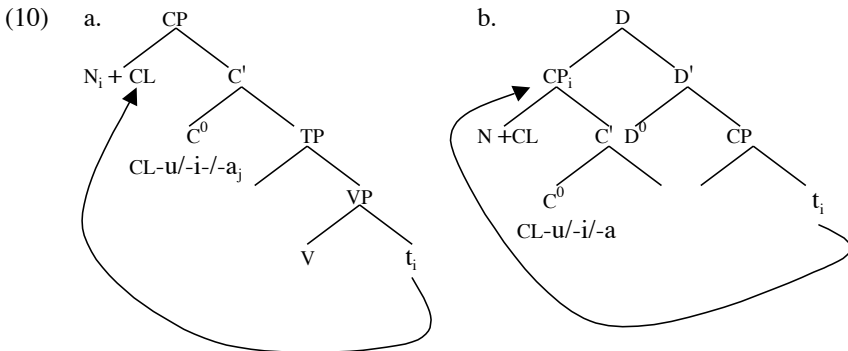   Given these facts, we assume that definite DPs are derived as in (8).

---

[4] Wolof nouns are invariant; class and number are marked on the dependents of the noun. Interestingly, the noun class marker, which appears as the initial consonant of the nominal modifier, is often a copy of the initial consonant of the noun. Though other criteria may determine class agreement, it is worth noting that elements that show phonological class agreement (including less-than-perfect copying) are the same ones that show ATR agreement. For more details on noun classification, see Sy (2003, 2004).

(8)

```
              DP
          /        \
      NP_i          D'
       ↑         /      \
       |       D^0       NP
       |      CL-i/a    / \
       |_____  t_i
```

Other nominal modifiers include relative clauses and genitive phrases. Relative clauses comprise two types which have very similar structures: Wolof adjectives are verbs which are introduced by the same type of relativizer used in non-adjectival clauses. These relativizers, as well as the determiners that may occur in such DPs, agree in noun class, [ATR] feature, and other deictic features with the head noun, as shown in (9).

(9) a. **góór g**- u   njool **gële**   b. **xale b**- i   toog **bale**
       man CL-REL.INDEF tall   DEM.DIST   child CL-REL.DEF sit   DEM.DIST
       'that (distal) man who is tall '   'that (distal) child who is seated'
     c. góór **g- ë**   wex- oon **g- ë**   ca   ndaw- am
       man CL-REL.DIST bitter-PAST CL-DEF.DIST PREP youth-his
       'the (distal) man who was disagreeable in his youth'

Following Torrence (2005), who, based on the dependencies discussed above, argues for a raising promotion analysis (as in Kayne 1994) of Wolof relatives, we assume the following structure.

(10)   a.   b.

```
          CP                                           D
       /      \                                     /      \
   N_i + CL    C'                               CP_i        D'
     ↑       /    \                           /     \     /    \
     |     C^0     TP                     N +CL     C'  D^0    CP
     |   CL-u/-i-/-a_j  \                         /   \         \
     |                   VP                      C^0          t_i
     |                  /  \                   CL-u/-i/-a
     |                 V    t_i
```

2.   The Experiment
2.1.   Method
The data for this experiment was recorded in the UCLA phonetics lab sound booth. The speakers are two male Senegalese natives in their thirties: one a speaker of the dialect of Dakar, the other a speaker of the dialect of St.-Louis. The

collected data consist of 108 tokens distributed among three different sets of 36 tokens as follows.

Set 1 contains 36 minimal or near-minimal pairs in single root form in the carrier sentence 'say X also', where X stands for the target token. These are exemplified in (11).

(11)   a. [-ATR] root   b. [+ATR] root
            waxal *kër*   itam   waxal *ker*   itam
            'Say   house also.'   'Say   shade also.'

Set 2 consists of nouns which minimally contrast with respect to tongue root value only, combined with different harmonizing determiners that contain front, central, and back mid vowels as shown in (12). The nouns were put in a carrier sentence of the type 'I saw X yesterday', where X stands for the target structure.

(12)   a. [-ATR] noun   b. [+ATR] noun
            gis naa *ker*   *ga* démb   gis naa *kër*   *gë* démb
            see 1sg. shade the yesterday   see 1sg.   house the yesterday
            'I saw that shade yesterday.'   'I saw that house yesterday.'

Set 3 contains the same pair of nouns as Set 2, but here the nouns are combined with two types of adjectival relative clauses, one [+ATR], the other [-ATR], so as to produce both harmonic and disharmonic DPs, as illustrated in (13).

(13)   a. [-ATR] N, [+ATR] ADJ CP   b.   [+ATR] N, [-ATR] ADJ CP
            gis naa *[ker*   *gu réy ga]* démb   … *[kër*   *gu weex gë]*…
            see 1sg. [shade REL big the] yesterday   … [house REL white the]…
            'I saw that big shade yesterday.'   '…that white house…'
       c. [-ATR] N, [-ATR] ADJ CP   d.   [+ATR] N, [+ATR] ADJ CP
            gis naa *[ker*   *gu   weex ga]* démb   …*[kër*   *gu   réy gë]*…
            see 1sg. [shade REL white the] yesterday …[house REL big the]…
            'I saw that white shade yesterday.'   '…that big house…'

The recording was digitized using PCQuirer with a sample rate of 11 KHz. Waveforms and spectrograms were generated from the same program and combined to measure the target vowels' first and second formant frequency.[5] These formant values were then used to calculate mean and standard deviation, as well as to perform statistical comparisons pairwise (t-tests, two sample, assuming equal variance).

The formant values of the target vowels in fully harmonic simplex DPs (Set 2) were compared to those of the vowels in the disharmonic complex DPs (Set 3).

---

[5] In the interest of space, some details about vowel measurement as well as statistical data will be left out.

The range of formant values in these two sets was also compared to that of the base forms (Set 1).

3.   The Results
3.1.   Base Forms
The results show a two-way split between the front and central vowels on the one hand, and the back vowels on the other. As apparent in (14) below, the average F_1 frequency value for [+ATR] mid front vowels (360.1 Hz) is lower than that of their [-ATR] counterparts (553.8 Hz). Similar facts can be observed in mid central vowels; [+ATR] central vowels have lower average F_1 frequency values at 486.8 Hz, as opposed to 643.6 Hz for their [-ATR] counterparts. The results of the two-way analysis performed on this data set show a highly significant difference with regard to F_1 frequency ($p<0.006$). F_2 frequency values, on the other hand, though consistently higher in the [+ATR] set for both front and central vowels, showed no significant difference between the two opposite sets ($p>0.5$).

(14)

	FRONT		CENTRAL		BACK	
	[+ATR]	**[-ATR]**	**[+ ATR]**	**[-ATR]**	**[+ATR]**	**[-ATR]**
Avg. F_1 value (Hz)	360.1	553.8	486.83	643.6	357.5	541.5
Avg. F_2 value (Hz)	1872.8	1854.3	1403.3	1344.3	944	1055
Std Dev F_1	35.8	17.8	22.1	79.7	9.57	46.06
Std Dev F_2	46.8	77.8	15.77	51.2	44.8	69.89

Turning to back vowels, the data in (14) shows the same correlation between low first formant frequency and tongue root advancement found in front and central vowels. The average F_1 frequency value for [+ATR] vowels is 357.5 Hz while that of [-ATR] vowels is 541.5 Hz. As far as F_2 values are concerned, back vowels indicate the reverse correlation of that found in front and central vowels: the average F_2 frequency value for mid back vowels is higher in the [-ATR] set. This reverse effect may be the result of different degrees of lip rounding; the advanced vowel, which is produced with a noticeably higher degree of lip round-ing, incurs a higher degree of second formant lowering. The results of the t-test performed on back vowels show that there is a significant difference between the [+ATR] set and the [-ATR] set for both F_1 and F_2.

In summary, we may conclude from these results that in Wolof, [+ATR] front and central vowels have higher and fronter tongue position than their [-ATR] counterparts, and [+ATR] back vowels have a higher and backer tongue position. Overall, the results also indicate that F1 is the most important acoustic correlate of [ATR] vowels, a fact which confirms the findings of earlier studies on the acoustic properties of [ATR] vowels (Lindau 1974, 1979).

3.2. DPs

The vowels of the determiners in all DP types display the same acoustic properties as those in base forms.

(15) Average [+ATR] formant values in all DP types (Hz)

Type	Front		Central		Back	
	F_1	F_2	F_1	F_2	F_1	F_2
Base forms	360.1	1872	486.8	1403.3	357.5	944
Simplex DPs	370.3	1865	432	1399.6	474	962
Harmonic complex DPs	369.66	1793	422.1	1463.1	478	1059
Disharm. complex DPs	380.5	1844	442	1419.8	492	960.5

(16) Average [-ATR] formant values in all DP types (Hz)

Type	Front		Central		Back	
	F_1	F_2	F_1	F_2	F_1	F_2
Base forms	553.8	1854.3	643.6	1344.3	541.5	1055
Simplex DPs	465	1858.5	506	1447	525.6	995.8
Harmonic complex DPs	451	1793	557.6	1530	537	1059
Disharm. complex DPs	480	837.3	523.6	1511	551.5	1019

As can be noted from (15) and (16) above, in front and central vowels the average F_1 frequency value is lower in the [+ATR] set, ranging between 360 and 380 Hz in front vowels, and 422 and 486 Hz in central vowels in all DP types, as compared to their [-ATR] counterparts, whose average F_1 values range between 450 and 550 Hz in front vowels, and 500 and 650 Hz in central vowels. Just as in base forms, F_2 frequency values in both classes of vowels are higher in the [+ATR] set, but again not significantly different.

Back vowels in all DP types also exhibit the same properties as their analogs in root forms; both F_1 and F_2 frequency values are lower in the [+ATR] set, though in this case the t-test results indicate that there is a significant difference with regard to F_1 as in root forms, but not F_2, unlike in root forms.

The results of the two-way analysis comparing all DP types show that there is no significant difference between any two sets of vowels in the three types of DP (simplex, harmonic, and disharmonic) with the same head noun. Importantly, the t-test performed on sets of DPs with heads of opposite values shows that there is a significant difference between the determiner vowels occurring with [+ATR] head nouns and those occurring with [-ATR] head nouns.

Considering the data in (15) and (16) more closely, one may find slight differences in average formant values between vowels in root forms and alternating vowels in DPs. These differences seem to represent a stronger tendency in root vowels toward the observed correlations: average first formant frequencies are lowest in root vowels in the [+ATR] and highest in root vowels in the [-ATR] set, as compared to DP vowels. However, the differences are quite minimal, as the range

of variability appears limited; for example, frequency values in front vowels remain characteristically under 400 Hz in the [+ATR] set, while they are consistently above 450 Hz in the [-ATR] set with low standard deviation numbers.

In conclusion, the overall result of this experiment reveals that the relationship between tongue root advancement/retraction and first and second formant frequency values observed in root forms holds in alternating determiner vowels as well. This relationship holds whether or not there is intervening material between the head noun and the determiner.

4. The Quest for an Account

Given the evidence for the occurrence of ULDAA in Wolof, we must now confront the problem of accounting for this phenomenon. Considering the fact that we are dealing with a phonological process, it is only natural to seek an answer in the phonology. In the following section we discuss various ways of approaching the problem from a phonological point of view.

4.1. The Phonology Can Do Half the Job

Setting aside the evidence presented above, let us assume for the sake of argument that phrase level harmony proceeds in exactly the same way as word-internal harmony (i.e. locally from left to right), and consider the data in (17) below.

(17)　a. kër　gu réy g**ë**　　　b. ker　gu réy g**a**
　　　　 house REL big　the　　　　 shade REL big　the
　　　　 'the big house'　　　　　　 'the big shade'

　　　c. kër　gu réy g**A**　　　d. ker　gu réy　gA
　　　　　　　[+ATR]　　　　　　　　　　[+ATR]
　　　　　　　　 |　　　　　　　　　　　　　 |
　　　　　　[**réy** | **gA**]　　　　　　[réy | g**A**]
　　　　　　　[+A̸TR]　　　　　　　　　[+A̸TR]
　　　　 kër　gu réy gë　　　　 * ker　gu réy　gë

　The example in (17a) is not a problem because the two potential triggers in the structure bear the same [+ATR] feature, but the one in (17b) is. In this case, letting harmony apply in the usual manner forces us to choose the trigger which is most linearly adjacent to the target, namely the [+ATR] adjectival verb *réy*, as schematized in (17c-d). As evidenced by the ungrammaticality of (17d), such an approach only works if the head noun and the verb bear the same harmonic feature as in (17a).

Alternatively, one may take an approach akin to Indirect Reference Theory[6] (Selkirk 1996, Truckenbrodt 1998, Hayes 1989) and argue that the domain of word-external harmony is the phonological phrase (P-phrase). Such an approach fails in exactly the same way as the previous one, for the simple reason that we must select one of the potential triggers as the head of the P-phrase; if we choose the head noun, then we must explain how we prevent the verb from triggering harmony on the determiner in the correct output in (17b), and if we choose the adjective, then once again we derive the incorrect form in (17d). Furthermore, other patterns of phrasal harmony such as those in (18) below resist the idea of P-phrasing as well.

(18) a. [$_{DP}$ tank-u [$_{NP}$ béy [$_{D}$ b -àle]]]$_{DP}$ b. [$_{NP}$ tank-u [$_{DP}$ béy [$_{D}$ w -ëlé]]$_{DP}$
          leg -GEN goat   CL-DEM   leg -GEN goat   CL-DEM
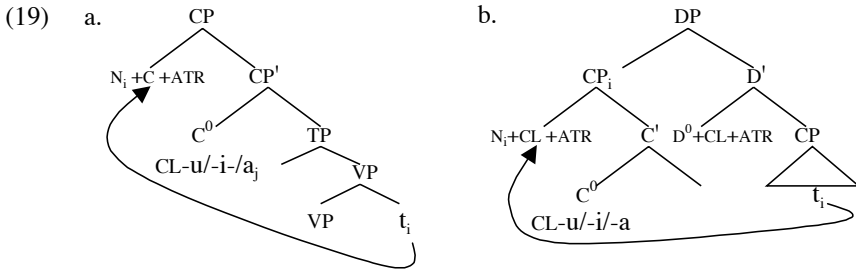   'that leg of goat'   'the leg of that goat'

Confronted with the data in (18), the phonological phrase approach would have to either assume that these two phonological phrases have two different heads (*tànk* in (18a) and *béy* in (18b)), or that the whole genitive in (18a) forms a P-phrase, but the one in (18b) *béy wëlé* forms a P-phrase to the exclusion of *tànk-u*. Both of these accounts amount to treating the same phonological structure in different ways.

To sum up, then, a purely phonological account of Wolof phrasal harmony is possible only in structures which contain lexical items with the same harmonic feature. In cases in which lexical items disagree, a phonological account becomes untenable in the face of ULDAA. Such cases are the very ones amenable to a syntactic account, as we will see in the following section.

4.2.   Syntax to the Rescue?

The cases which are problematic for a phonological account have one important common characteristic: they are structures in which a disharmonic syntactic constituent intervenes between the trigger and target of harmony within a bigger constituent, as in (17a-b) and (18a). In (17) the definite determiners agree with the nouns within their own projections, but the two are separated by the CP containing the adjectival verbs on the surface. Similarly, in (18a) the demonstrative *bale* is the head of the DP containing *tànk*, and therefore agrees with that noun, though there is an intervening NP complement of the noun between the two. Obviously, whatever local relationship exists between these nouns and the determiner must be accessible by the phonology, though it has been destroyed by syntactic operations. If we assume that [ATR] agreement can be instantiated through Spec-head agreement, if such a configuration is present at some point in the syntax, it is possible to check the [ATR] feature between N and D, as shown in (19).

---

[6] Such an approach has been proposed by Ka (1988) for Wolof. However, no cases of ULDAA were considered in his study.

(19) a.

```
                CP
       Nᵢ +C +ATR      CP'
                    C⁰        TP
                CL-u/-i-/-aⱼ      VP
                              VP      tᵢ
```

b.

```
                        DP
               CPᵢ              D'
        Nᵢ+CL +ATR    C'   D⁰+CL+ATR   CP
                    C⁰                       tᵢ
                CL-u/-i/-a
```

This approach straightforwardly captures at once the dependencies between the relativizing element, the determiner, and the head noun with regards to noun class agreement, vowel harmony, and other feature dependencies such as deixis.

Be that as it may, syntactic locality runs into problems in other cases of ULDAA. Consider (20) below.

(20) a. **ba   ma** góór gë **joxee** tééré ba   b. **bë   më** xale ba **yónné** tééré ba
         when 1sg. man the give book the   when 1sg child the send book the
         'when the man gave me the book'   'when the child sent me the book'

In the temporal clause in (20), the trigger (the verb) and the targets (the complementizer and the object clitic) are separated by the DP subject of the clause. Although it can be argued that the object pronoun is generated in post-verbal position, and thus has been in a local relationship with the verb at some point in the derivation before moving to its surface preverbal position, in the case of the complementizer, it would be hard to make such a case using only the usual syntactic mechanisms available in the derivation of such structures. What seems to be relevant in these cases and the ones previously discussed is the relation between trigger and target. In all these cases, harmony applies between a head and its complement within the XP that contains them. Any adequate account of these patterns must therefore allow the phonology to refer to syntactic structure and syntactic relations.

5.   Optimality Theory

Because it is an output oriented theory, OT seems suitable to resolve the problem. Archangeli and Pulleyblank (2002), in a study of Kinande [ATR] harmony, found that some vowel sequences behave differently in different morphological structures. In such cases, a purely phonological analysis leads to a paradox. The authors argue that this paradox points to the necessity of incorporating morphosyntactic domains distinguishing between root, stem and macro-stem. The general idea behind this proposal is that morphological domains determine the range of prosodic constituents (the Domain Qualification Hypothesis).

Extending this idea to the Wolof data, we might construe the role of syntax in Wolof phrasal harmony to be the same as that of morphology in DQH; that is,

syntactic bracketing may determine the range of harmony domains. In addition, syntactic relations determine trigger-target mapping.

Given the patterns observed in the data under consideration and the basic facts of Wolof vowel harmony, we can isolate the relevant conditions. Since lexical items always trigger harmony and never undergo harmony, there must be a faithfulness condition preserving their harmonic feature, and one ensuring that they trigger harmony on functional elements with which they co-occur. These two conditions are formalized in (21) and (22).

(21)   IDENT [ATR] LEX: The [ATR] feature specification of a lexical item is preserved in the output.

(22)   AGREE [ATR] LEX: Only lexical items may trigger harmony.

Furthermore, harmonizing elements agree with the lexical items in head position, not those in specifier position. This constraint is notated as in (23).

(23)   AGREE [ATR] HD: functional heads harmonize with lexical heads, not lexical specifiers.

Finally, lexical heads do not induce harmony on dependents of other heads: there is a requirement that every XP be a harmonic domain, as formalized in (24).

(24)   AGREE [ATR] XP: Every XP is a harmonic domain.

These constraints must crucially rank lower than the grounding harmony conditions HI/ATR, AND LO/RTR to ensure that high vowels remain advanced, and low vowels remain retracted. Then IDENT [ATR] LEX must dominate all other phrase level harmony constraints, since it is never violated. The remaining constraints interact with each other as follows.

AGREE [ATR] HD >> AGREE [ATR] LEX: In adjectival clauses the determiner agrees with the lexical head of the phrase which contains it, not the lexical item closest to it; similarly, in temporal clauses the complementizer and the object clitic do not harmonize with the subject DP closest to them, but with the head of the VP.

(25)   a. * [DP ker [CP gu réy] gë]   b. *[CP bë [VP më [DP góór gë[V joxee]]]]

AGREE [ATR] XP >> AGREE [ATR] LEX : Lexical heads trigger harmony on their own dependents; the determiner inside DP cannot agree with any lexical item outside that DP.

(26)   a. * [CP bë [VP më [CP góór ga [V joxee]]]]   b. *[DP tànku [NP béy]D bëlé]]

These rankings generate the following hierarchy.

(27)   HI/ATR, LO/RTR>> IDENT [ATR] LEX >> AGREE [ATR] HD >> AGREE [ATR] XP
        >> AGREE [ATR] LEX

6.   Conclusion

The Wolof data has provided evidence that the domain of word-external harmony is subject to syntactic restrictions. One of the effects of such restrictions is what we have called ultra-long distance ATR agreement (ULDAA). Such non-local phonological agreement representations have been shown to resist a purely phonological as well as a purely syntactic account, arguing for the need to allow freer interaction of the two components. Within an optimality theoretic frame-work, it is possible to allow such interaction in a single ranked hierarchy.

References

Archangeli, Diana, and D. Pulleyblank. 2002. Kinande vowel harmony: Domains, grounded conditions and one-sided alignment. *Phonology* 19:139-188.
Ka, Omar. 1988. Wolof phonology and morphology: A non-linear approach. Ph.D. diss., University of Illinois Urbana Champaign.
Ladefoged, Peter. 1968. *The Phonetic Study of African Languages*. Cambridge: Cambridge University Press.
Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
Lindau, Mona. 1979. The feature expanded. *Journal of Phonetics* 7:163-176.
Stewart, J.M. 1967. Tongue root position in Akan vowel harmony. *Phonetica* 16:185-204.
Pike, K.L. 1967. Tongue root position in practical phonetics. *Phonetica* 17:129-40.
Sy, Mariame I. 2003. Noun formation and noun classification. UCLA Working Papers in Linguistics 9, Papers in African Linguistics 3: 43-85
--. 2004. Wolof noun classification: A constraint based approach. Paper presented at the 35th Annual Conference on African Linguistics, Harvard U., April 2-4.
--. 2004. Is there long-distance ATR harmony in Wolof? ms., UCLA.
Torrence, Harold. 2005. On the distribution of Wolof complementizers. Ph.D. diss., UCLA.

Mariame Sy
Columbia University Institute of African Studies
1103 International Affairs Building
420 West 118th Street
NY, NY 10027-3331

sms2168@columbia.edu

A Promotion Analysis of Wolof Relative Clauses

HAROLD TORRENCE
University of California, Los Angeles

0.   Introduction

This paper presents an analysis of relative clauses in Wolof, an Atlantic language spoken principally in Senegal and The Gambia, exemplified in (1)a-b below:

(1) a.   jigéén   j.i   Dudu gis-óón   direct object relative
           woman CL.*i* Dudu see-PAST
           'the woman that Dudu saw'

   b.   nit   ñ.a dem ca kër   gë   subject relative
           people CL.*a* go   P   house the[1]
           'the people there who went to the house'

   I will present support for a promotion analysis of relative clauses in Wolof (Vergnaud 1974, Kayne 1994) and argue that relative clauses like (2)a are derived as in (2)b:

(2)   a. xale **b.i**   Isaa dóór
             child CL.*i* Isaa hit
             'the child that Isaa hit'
          b.

----
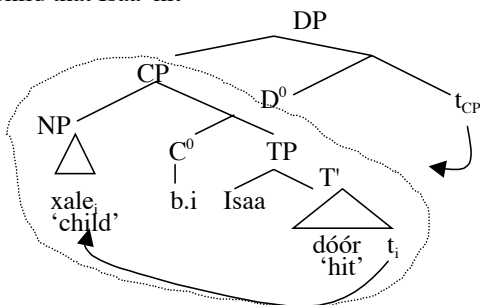
[1] Abbreviations: CL-class, DEF-definite, INDEF-indefinite, P-preposition, XPL-expletive.

Specifically, there are three points which I will argue for here, and which are shown in the tree in (2)b. First, it will be seen that the CL.*i* element is an agreeing complementizer, not a relative pronoun. Second, I will show that the head of the relative clause, *xale* 'child', in (2) raises from inside of TP to Spec,CP. Third, evidence will show that relative clauses in Wolof involve a $[_{DP} D^0 [CP]]$ structure in which the CP raises to Spec,DP (Kayne 1994).

In Section 1, I present the basic facts of Wolof and a description of simple relative clauses. Section 2 contains arguments that the CL.*i* element is an agreeing complementizer. Section 3 presents data showing that the relativized NP is promoted from inside of TP. Section 4 covers CP raising. The final section summarizes the conclusions.

1.   The Basics

Wolof is an SVO language with prepositions, postnominal relative clauses, and twelve basic noun classes:

(3)   xale b.i   lekk-na a.m   mbonaat ca kër   g.ë
       child CL.DEF eat- *na* INDEF.CL turtle   P house CL.DEF
       'the child ate a turtle at the house'

That the nouns *xale* 'child', *mbonaat* 'turtle', and *kër* 'house' belong to different noun classes can be seen from the distinct manifestations of class agreement on the determiners. Thus, with *xale*, the class consonant is *b-*, for *mbonaat* it is *m-*, while for *kër* it is *g-*. In Torrence (2003, 2005), I argue that the element glossed as *-na* is a complementizer that occurs low in the clause.

Wolof relative clauses have *-i*, *-a*, or *-u* on the left edge, preceded by the relativized NP:

(4) a. xale <u>b.**i**</u> jigéén ñi   dàq   -*i*-relative clause
        child CL.i woman the.PL chase
        'the (proximal) child that the women chased'

    b. xale <u>b.**a**</u> jigéén ñi   dàq   -*a*-relative clause
        child CL.a   woman the.PL chase
        'the (distal) child that the women chased'

    c. kéwél <u>**g.a**</u> jigéén ñi   dàq   -*a*-relative clause
        gazelle CL.a woman the.PL chase
        'the (distal) gazelle that the women chased'

    d. xale <u>b.**u**</u> jigéén ñi   dàq   -*u*-relative clause
        child CL.*u*   woman the.PL chase
        'a child that the women chased'

The underlined strings in (4), the "relative markers," consist of *-i/-a/-u* preceded by a class consonant that agrees with the relativized NP. As the translations indicate, the presence of *-u/-i/-a* corresponds to definite versus indefinite interpretations of the relative clause head. When *-i* is present, the head of the relative clause is interpreted as definite and proximal; when *-a* is present, the head NP is definite and distal, while when *-u-* appears, it signals that the relativized NP is indefinite. In this paper, I concentrate on the *-i/-a* relative clauses and exemplify only with *-i-* relatives, although the same facts hold of the *-a-* relatives.

Note first that the vowels *-i/-a* occur as definite determiners, follow the NP, and agree with the NP in class:

(5) a. meew m.**i** 'the milk here' cin l.**i** 'the pot here'[2]   *-i-*
     b. meew m.**a** 'the milk there' cin l.**a** 'the pot there'   *-a-*

The TP that occurs in a relative clause can contain tense, negation, and subject and non-subject clitics. In addition, the definite article optionally appears on the right edge, following the entire TP. The definite article agrees with the relativized NP, even though it may be separated by much material (agreements are indicated by the arrows):

(6) meew m.i [$_{TP}$ leen-fa Bintë jox- al- ul woon xale bi] (m.i)
     milk CL.*i* 3PL- LOC Binta give-BEN-NEG PAST child the   CL.DEF
     'the milk that Binta did not give to the child for them'

The non-subject clitics, *leen* and *fa*, in (6) are quite high in the structure and precede the DP subject, *Bintë*. That the first occurrence of the string *m.i* in (6), immediately following the relativized NP *meew* 'milk', is not the determiner is suggested by the fact that the "DP" variant of (4)d is ungrammatical:[3]

(7) *xale b.u
     child CL.*u*
     'the child'

If the position immediately following the relativized NP were the position of the determiner, it is expected that (7) would be good, contrary to fact. Thus, the

---

[2] The *-u-* is also found in the indefinite determiner, which precedes the noun and agrees in class with it:
 (i) **u.b** xale
       u.CL child
       'a child'
[3] The form in (7) is reported to be good in some dialects.

determiner is what appears on the far right, following TP. Relative clauses can be represented templatically as:

(8)   a. [$_{CP}$ NP CL.*i* [$_{TP}$ Clt$_S$-Clt$_O$-Clt$_{Loc}$ S V O]] (CL.*i*)   *-i*-relative clause
       b. [$_{CP}$ NP CL.*a* [$_{TP}$ Clt$_S$-Clt$_O$-Clt$_{Loc}$ S V O]] (CL.*a*)   *-a*-relative clause

2.   The Status of *-i/-a*

This section addresses the status of *-i/-a* in the relative markers. Two immediate possibilities come to mind based on crosslinguistic considerations: relative pronouns and complementizers. That the relative markers are relative pronouns is plausible because they agree with the head noun and look like determiners. An analysis in which *-i/-a* are complementizers has initial appeal because the relative markers occur on the left edge of the clause, where complementizers typically occur in the language and agreeing complementizers are known from other languages (e.g. French *que/qui* alternation, Kayne 1975). I will argue that the relative markers are complementizers that agree with the NP in their specifier, as below:

(9)
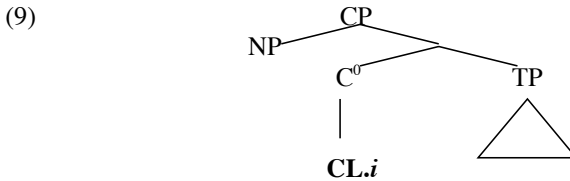
Before analyzing Wolof, consider first relative pronouns in English:

(10)   a. the man who you hit

       b.
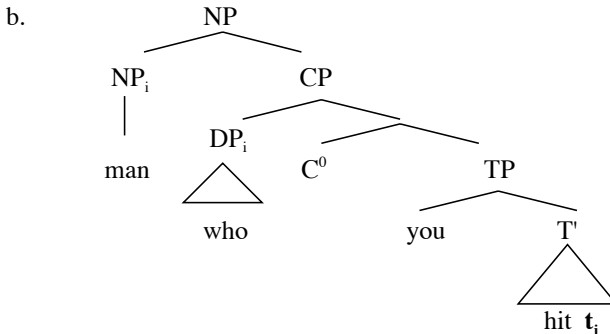
    In the example above, the relative pronoun *who* is merged as the complement of *hit*. It undergoes A´-movement to Spec,CP, where C^0 is silent. The relative CP itself is an adjunct to NP. Roughly, the "head" of the relative clause, *man*, is

coindexed with the relative pronoun, giving rise to the interpretation of the man as the object of *hit*. An important point here is that *man* is never inside of TP in this analysis. Translating this relative pronoun analysis to Wolof, the derivation of (11)a is sketched in (11)b.

(11)　　a. góór **g.i**　nga dóór
　　　　　　 man　CL.*i* 2sg. hit
　　　　　　 'the man who you hit'

　　　 b.
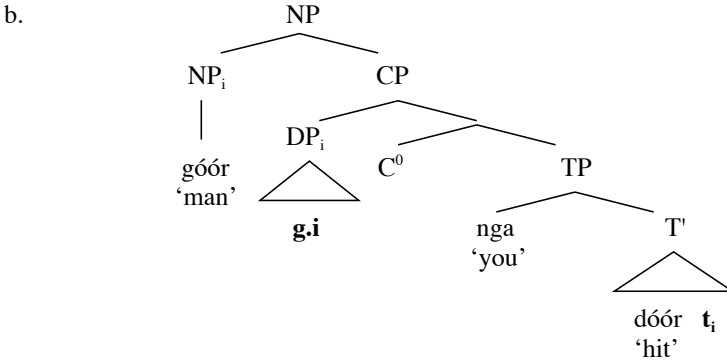
In (11)b, the putative relative pronoun, *g.i*, is merged as the complement of *dóór* 'hit'. It raises to Spec,CP, where C⁰ is silent, as in English. The head of the relative clause, *góór* 'man', is coindexed with the relative pronoun, which agrees overtly in class with the *góór*.

The principal argument against a relative pronoun analysis comes from the fact that it is possible to iterate *-i*:

(12)　　a. ?tééré **b.i**　ñu　foog　[_CP **b.i**　 ma　jàng]⁴
　　　　　　 book　CL.*i* 3pl. think　　CL.*i* 1sg. read
　　　　　　 'the book that they think I read'

　　　 b. ?nit　　 **ñ.i**　nga wax　Bintë [_CP **ñ.i**　Gàllaay dàq　]
　　　　　　 people CL.*i* 2sg. tell　Binta　　CL.*i* Gallaay chase
　　　　　　 'the people who you told Binta that Gallaay chased'

---

⁴ Note that (12)a and b have question marks. The most natural way of doing relativization from an embedded clause is with a cleft:
　(i)　tééré b.i　　ñu　foog　ne [_CP l-　　a-a　　jàng]
　　　 book CL.*i* 3pl. think *ne*　XPL-*a*-1sg. read
　　　 'the book that they think I read'
　(ii)　nit　　 ñ.i　nga wax Bintë　[_CP l-　　a Gàllaay dàq]
　　　　person CL.*i* 2sg. tell Binta　　XPL-*a* Gallaay chase
　　　　'the people who you told Binta that Gallaay chased'

In (12)a, the relative marker *b.i* occurs twice: on the left edge of the clause containing *jàng* 'read', and again on the left edge of the clause containing *foog* 'think'. This is unlike the behavior of relative pronouns in English:

(13)   *the man **who** you think **who** Gary chased

If the Wolof CL.*i* is a relative pronoun, the iteration is unexpected. This is because under the relative pronoun analysis, *b.i* in (12)a is the complement of *dóór* 'hit'. It is thus mysterious how and why it can occur multiple times. However, if the relative markers are complementizers, multiple occurrence is expected. Complementizers introduce CPs. When there are multiple CPs, multiple complementizers are expected. This is like English *I think **that** you said **that** Bill knows **that***…. From these considerations, I conclude that -*i*/-*a* are complementizers, not relative pronouns.

3.   NP Raising in Relative Clauses
In this section, we consider the relation between the relativized NP *taabal* 'table' and the gap (underlined) inside of TP:

(14)   **taabal** j.i [$_{TP}$ xale yi   gis __   démb]
         table   CL.*i*   child the.PL see   yesterday
         'the table that the children saw yesterday'

The main issue is whether *taabal* 'table' in (14) is base generated in its surface position or raised from the position corresponding to the gap.

That movement is involved in Wolof relativization can be seen from the fact that relativization in Wolof is sensitive to both weak and strong islands:[5]

Complex NP
(15)   a. tééré b.i jigéén ji jox xale yi   démb
            book CL.*i* woman the give child the.PL yesterday
            'the book that the woman gave to the children'

         b. *xale$_k$ y.i [tééré$_j$ b.i [$_{TP}$ jigéén ji jox **t$_j$** **t$_k$** démb   ]]
            child CL.*i* book CL.*i*   woman the give   yesterday
            'the children that the book that the woman gave yesterday'

Adjunct Island
(16)   a. gis- na- a   Bintë [laata ñu jox tééré yi   xale bi].
            see- *na*-1sg. Binta before 3pl. give book the.PL child the
            'I saw Binta before they gave the books to the child.'

_____
[5] These islands can be repaired with resumptive pronouns.

b. *tééré_i y.i ma gis Bintë [laata ñu jox **t**_i xale bi]
   book CL.*i* 1sg. see Binta before 3pl. give   child the
   'the books that I saw Binta before they gave the child'

Wh- Island

(17)   a. fàtte- na- a k.u sàcc tééré bi.
           forget- *na*-1sg. CL.*u* steal book the
           'I forgot who stole the book.'

b. *tééré b.i ma fàtte k.u sàcc
   book CL.*i* 1sg. forget CL.*u* steal
   'the book that I forgot who stole'

   While the island facts show that relativization in Wolof involves movement, they do not show *what* has undergone movement. It could be that the relativized NP has moved. Alternatively, it could be that the relativized NP is base generated and a silent operator has undergone A´-movement. It would then be the silent operator which is responsible for the island sensitivity. That idiom chunks (Vergnaud 1974) can be relativized provides strong support for the promotion analysis of relative clauses.

(18)   a. def- na- a   tééré Senegaal. *def tééré X* = be dedicated to X
           do- *na*- 1sg. book Senegal
           'I am dedicated to Senegal.'
           (lit. 'I do book Senegal')

b. **tééré** b.i nga **def** Senegaal
   book CL.*i* 2sg. do Senegal
   'the fact that you are dedicated to Senegal'

c. **tééré** b.i nga foog ne l- a- a **def** Senegaal
   book CL.*i* 2sg. think that XPL-*a*- 1sg. do Senegal
   'the fact that you think I am dedicated to Senegal'

   (18)a shows the idiom *def tééré* in a typical matrix clause. (18)b and c show that part of the idiom, *tééré*, can be relativized in a single clause ((18)b) and relativized from an embedded clause ((18)c). Since the idiom itself is a type of lexical item, all of its parts are merged together in TP. Thus, *tééré* attains its surface position in Spec,CP by movement from its merge position inside of TP:

(19)
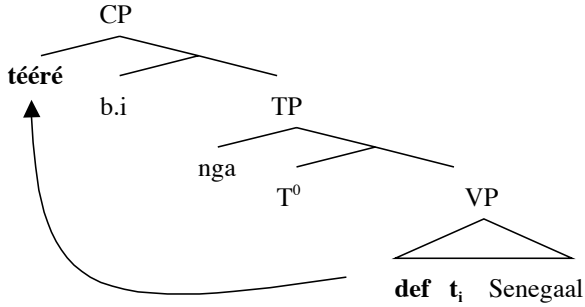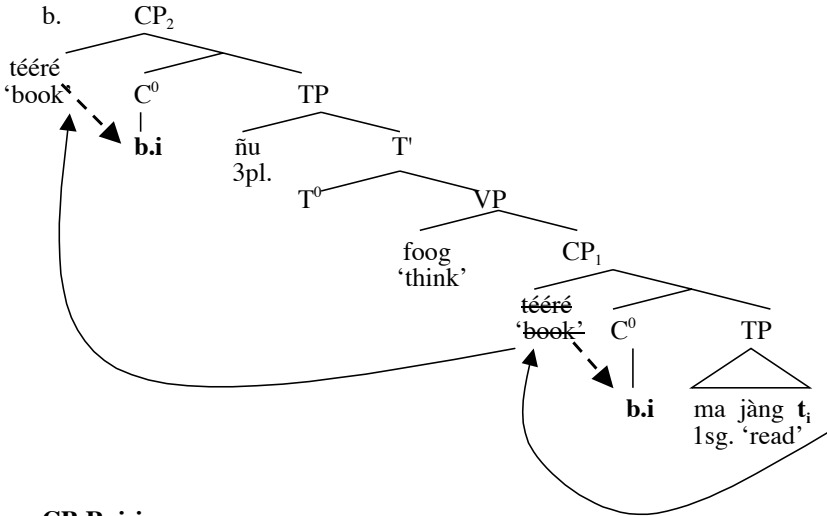

The conclusion that the relativized NP is promoted from inside of TP means that cases where *-i* is iterated (e.g. 20a) arise through successive cyclic movement of the relativized NP through the intermediate Spec,CPs, as in (20)b:

(20)   a. ?tééré **b.i** ñu foog [$_{CP}$ **b.i**   ma jàng]   (=(12)a)
           book CL.*i* 3pl. think   CL.*i* 1sg. read
           'the book that they think I read'

    b.
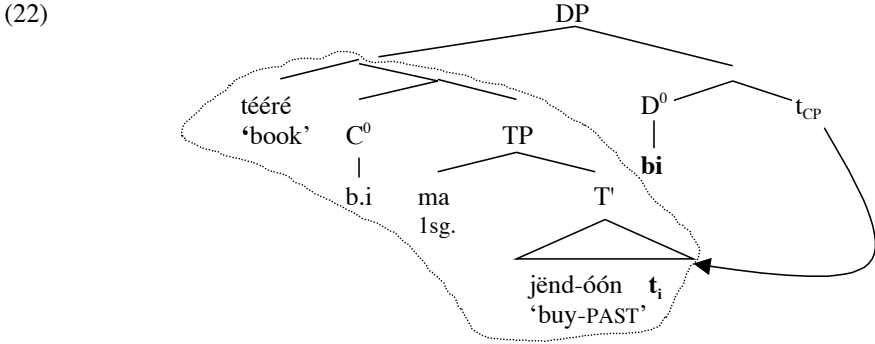

4.   CP-Raising

Wolof provides strong support for an analysis of relative clauses as CP complements to D[0], as proposed in Kayne (1994). This is because Wolof shows this configuration overtly. Recall that the definite determiner may optionally appear on the right edge of a relative clause:

(21)   [$_{CP}$ tééré b.i [$_{TP}$ ma jënd- óón]] **bi**
      book CL.*i*   1sg. buy- PAST   the
     'the book that I bought'

The right peripheral position of the definite determiner, *bi*, follows without stipulation from CP-raising to SpecDP:

(22)

Empirical support for CP raising comes from the distribution of adverbs. Wolof has both run-of-the-mill adverbs like *lool* 'very' ((23)a) and a set of idiomatic adverbs, often called "ideophones," that only occur with particular verbs or classes of verbs ((23)b,c):

(23)   a. Isaa ragal- na **lool** xaj bi.
       Isaa fear- *na* very dog the
       'Isaa is very afraid of the dog.'

      b. dagg- na- a   yapp wi **fàtiit**.   ideophone + verb of cutting
        cut- *na*- 1sg. meat the ADV
        'I cut the meat in one stroke.'

      c. *naan- na- a   meew mi **fàtiit**   ideophone without verb of cutting
        drink- *na*- 1sg. milk   the ADV

The adverb *fàtiit* only occurs with verbs of cutting. Both types of adverbs can occur in relative clauses inside of TP. In that case, they precede the definite determiner, because they have been pied piped with CP:

(24)   a. [$_{DP}$ [$_{CP}$ xaj [$_{C'}$ b.i [$_{TP}$ Isaa ragal **lool**]]] **bi**]
        dog CL.*i*   Isaa fear very   the
       'the dog that Isaa really fears'

b. [$_{DP}$ [$_{CP}$ jën [$_{C'}$ y.i [$_{TP}$ Isaa dàgg **fàtiit**]]] **yi**]
      fish   CL.*i* Isaa cut ADV   the.PL
   'the fish that Isaa cut in one stroke'

Strikingly, both types of adverbs can also be "stranded" to the right of the definite determiner:

(25)   a. [$_{DP}$ [$_{CP}$ xaj bi   Isaa ragal **t$_k$**] [$_{D'}$ **bi** **lool$_k$**]
           dog CL.*i* Isaa fear   the very
      'the dog that Isaa is very afraid of'

    b. [$_{DP}$ [$_{CP}$ xaj [$_{C'}$ y.i [$_{TP}$ Isaa dàgg]]] [$_{D'}$ **yi**   **fàtiit**]

(26)   Adverb positions
       a. Adverb in TP   b. Adverb in ModP
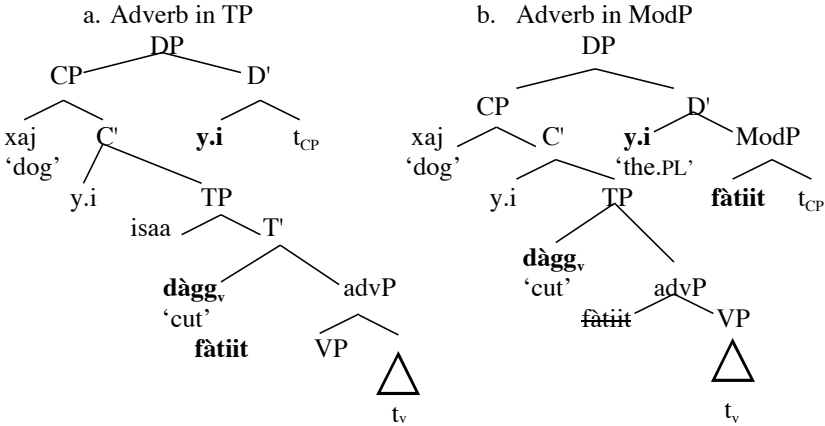
5.   Summary

In this paper, I have argued for several points. A promotion analysis of Wolof relative clauses is supported by idiom chunk relativization. I also concluded that the relative markers are agreeing complementizers, not relative pronouns as in English. This conclusion is based on the possibility of multiple occurrences of the relative markers. Finally, it was argued that Wolof relative clauses involve [D^0 [CP]] structures, where CP is the complement of D^0. In Wolof, this is seen overtly. However, in Wolof, CP raises to Spec,DP. This movement can be detected by the stranding of adverbs.

References

Kayne, Richard. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.
--. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2002. Locality and the left periphery. In Adriana Belletti, ed., *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, 223-251. Oxford: Oxford University Press.

Torrence, Harold. 2003. Verb Movement in Wolof. In Jason Kandybowicz, ed., *Papers in African Linguistics 3. UCLA Working Papers in Linguistics, Number 9*, 85-115. Los Angeles: University of California, Los Angeles Linguistics Department.

--. 2005. On the distribution of complementizers in Wolof. Ph.D. diss., University of California, Los Angeles.

Vergnaud, J-R. 1974. French relative clauses. Ph.D. diss., MIT.

Harold Torrence
University of Kansas
Department of Linguistics
1541 Lilac Lane, Blake Hall – Room 427
Lawrence, KS 66044-3177

torrence@ku.edu

The Alleged Class 2a Prefix bɔ̀ in Eton: A Plural Word

MARK L. O. VAN DE VELDE
Fund for Scientific Research - Flanders / K.U.Leuven

0.   Introduction

Since the beginnings of Bantu linguistics, the extended gender system of the Bantu languages has received a lot of attention.[1] The focus of this attention has been mainly comparative-historical, sometimes to the detriment of careful synchronic analysis. This paper focuses on the morpheme *bɔ̀* in the Cameroonian Bantu language Eton (A71), which should be identified as the nominal prefix of class 2a according to the traditional criteria and terminology. It will be shown that this morpheme is not a prefix but a word, probably a proclitic. Its function is to pluralize a following genderless word. Since the most typical genderless words are proper names and deictically restricted kinship terms (e.g. *tàdá* 'my father'), the result is usually an associative plural. This explains how the succession of the locative preposition *á* and the plural word could grammaticalize into the complex preposition *ábɔ̂* 'chez'. The behaviour of *bɔ̀* in Eton confirms some observations that Dryer (1989) made in his typological study on plural words. The end of this paper presents a brief comparative overview of the class 2a marker in other Bantu languages.

1.   The Noun Class System

Eton is a previously undescribed Narrow Bantu language spoken just north of the Cameroonian capital Yaoundé. It belongs to the Beti part of the Beti-Bulu-Fang dialect cluster (also called *Pahouin*). Its closest relatives are Ewondo and

---

[1] I interpret *noun class* and *gender* as alternative terms for the same phenomenon. In Bantu studies the term *noun class* is usual, but I follow Corbett (1991) in using the term *gender*, except when explicitly referring to the Bantuist tradition. Moreover, I wil use the term *class 1a* for the set of nouns that is traditionally analysed as a subgender of gender 1, but that I analyse as a group of genderless words in Eton (idem for *class 2a*). I wish to thank Dmitry Idiatov and my informants Pie-Claude Ondobo and Désiré Essono for their help. Vertical bars mark morphophonological notation. In this notation accentuated syllables are underlined and all morpheme boundaries are marked as follows: # word boundary, = clitic boundary, - affix boundary. The following glosses are used: AU augment; CON connective; DEM demonstrative; DIM diminutive; IMP imperative; PL plural; PRES present; SF suffix. Arabic numbers are used for nominal gender prefixes in the glosses, Roman numbers for agreement prefixes.

Mengisa. Eton has a typical gender system with ten genders, but lacks locative genders. Table 1 gives an overview. The numbers in the first column refer to the genders reconstructed for Proto-Bantu of which the current Eton genders are the reflex. The second column gives the nominal gender prefix (NPr), and the third column the verbal prefixes (VPr). The gender prefixes have variable forms depending on the form of the following morpheme. In genders 1 and 3 the nominal prefix is a syllabic homorganic nasal. Before a vowel the homorganic nasal is represented by the phoneme /m/, as in the gender 1 noun *m-ìŋgá* 'woman'. The prefix is not syllabic in that case. Genders 9 and 10 historically had a homorganic nasal nominal prefix. This nasal is only preserved before voiced stops, but there are no morphological arguments for treating it as a prefix in present day Eton. The CV-prefixes of genders 2, 4, 6 and 8 are reduced to C- or CG- before morphemes that begin in a vowel. The preconsonantal and prevocalic forms of the gender 5 and gender 7 prefixes cannot be reduced to one morphophoneme and must be described as allomorphs. The forms in the two rightmost columns of Table 1 are suppletive agreement targets. The connective morpheme (Con) is a proclitic that relates a head noun to a modifying noun, as in (1). The choice between the segmental and the tonal form of the connective depends on the syllable structure of the modifying word. The double accent ˝ symbolizes the morphotoneme *D*, which is represented by a high tone, except if preceded by a high tone. In the latter case *D* is represented by a low tone.

(1)   /mèlɔ́ mé béwógô/
         |mè-lɔ́ # mé=bè-wóg-Àl
         6-ear   VI.CON=2-hear-SF
         'the attention of the audience' (lit.: 'the ears of the hearers')

(2) Table 1: The Eton gender prefixes

Gender	NPr	VPr	Con	Dem
1	*Ǹ*	*à*	*à/Ø*	*ɲɔ̌*
2	*bè*	*bé*	*bé/H*	*bǎ*
3	*Ǹ*	*ú*	H	*ɲǔ / vǐ*
	à	*ú*	H	*ɲǔ / vǐ*
	ù	*ú*	H	*ɲǔ / vǐ*
4	*mì*	*mí*	*mí/H*	*mǐ*
5	*ɛ̀/d*	*é/d*	*é/H*	*dǐ*
6	*mè*	*mé*	*mé/H*	*mǎ*
7	*ì/j*	*í*	*í/H*	*yǐ*
8	*bì*	*bí*	*bí/H*	*bǐ*
9	*Ø*	*ì*	*ì/ Ø*	*ɲǐ*
10	*Ø*	*í*	*í/H*	*ɲǐ*

Example (3) illustrates how a connective, a demonstrative and a finite verb agree with a class 6 controller noun.

(3)   /ímébwád mébôd má mé⁺té kwàn/
       lí-mè-búád # mé=b-òd #   mǎ #   mé-Lté # Lkùàn|
       AU-6-thin   VI.CON=2-person VI.DEM   VI-PRES be.ill
       'These slender people are ill.'

   Note that three types of nouns belong to gender 3: nouns with a nasal prefix (3n), nouns with the prefix *ù-* (3u) and nouns that take the prefix *à-* (3a). The plurals of gender 1 nouns belong to gender 2. Gender 9 words can form their plurals freely in gender 10 or 6. Other common gender pairings are 3n/4, 5/6, 7/8, 3a/6 and 3u/5.
   One set of nouns is not represented in Table 1. These are the nouns that have no gender prefix and that trigger the same agreement pattern as gender 1 nouns, as in (4). Their plurals are marked by *bɔ̀* and trigger agreement pattern II, the agreement pattern that is also associated to gender 2, as in (5).

(4)   dwábdɔ̂ à-té   kwàn
       doctor   I-PRES be.ill   'The doctor is ill.'
(5)   bɔ̀ dwábdɔ̂ bé-⁺té kwàn
       PL doctor   II-PRES be.ill   'The doctors are ill.'

These sets of nouns are traditionally called class 1a and class 2a respectively in Bantu studies. Class 1a is treated as a subgender of gender 1, because it has the same agreements, and class 2a as a subgender of gender 2. These subgenders were first signalled by Doke (1927), who observed that they typically contain kinship terms, proper names, the question word "who," and borrowings from European languages. This is also the case in Eton.

(6)   a.   mbán wâmɔ̀ àté zù   'My co-wife is coming.'
       b.   bɔ̀ mbán bâmà bé⁺té zù 'My co-wives are coming.'
(7)   a.   càlà àté zù   'Tsala is coming.'
       b.   bɔ̀ càlà bé⁺té zù   'Tsala and his family are coming.'
(8)   a.   zá àté zù?   'Who is coming?'
       b.   bɔ̀ zá bé⁺té zù?   'Who are coming?'
(9)   a.   lèŋkɔ́d wâmɔ̀   'my raincoat'
       b.   bɔ̀ lèŋkɔ́d bâmà   'my raincoats'

   It appears to be impossible to reconstruct a form for the class 2a marker in Proto-Bantu. Meeussen (1967:100) tried *baa-*, but could not decide on the tone. Some scholars (e.g. Maho 1999:156, Poulos and Louwrens 1994:16) distinguish between class 2a and class 2b based on the form of the marker. Class 2b has a

marker with a back vowel (*ɔ* or *o*) and is found in most languages of zone S (roughly South Africa, Botswana, Lesotho and parts of Zimbabwe and Mozambique), in some zone K languages and in some northwestern Bantu languages, such as Eton. The marker of class 2a has the vowel *a* and is found elsewhere (see the map in Maho 1999:155). I will not make this distinction here for two reasons. First, there are many more lines along which the form of the class 2a/2b marker differs among languages than just vowel quality. Second, the crosslinguistic difference in vowel quality does not reflect any difference in function and morphosyntactic status of the class 2a marker.

2.   bɔ̀ is a Plural Word in Eton

Contrary to what one expects from a Bantu gender marker, *bɔ̀* turns out to be a word in Eton, not a prefix. Its word status can be easily established by means of phonological criteria. This is because the first syllable of every word stem is accentuated. Accentuated syllables have a prominent initial consonant, i.e. a consonant that is longer than the others in intervocalic position and that is not subject to lenition rules. Moreover, they can carry two structural tones, whereas non-accentuated syllables can carry only one. The plural word *bɔ̀* clearly differs from prefixes in this respect. In (10)-(11) two plural nouns are preceded by the locative preposition *á*, which, according to the general tone rules, copies its high tone onto the following word. Since *bɔ̀* is accentuated, it can carry two structural tones, so that the copied high tone can simply add to the original low tone and form a falling tone (10b). The gender 8 prefix *bì-*, on the other hand, is not accentuated. It is not the first syllable of a stem. Therefore it cannot carry both the copied high tone and the original low tone. Consequently, the copied high tone pushes the low tone of the nominal prefix to the right, where it downsteps the high tone of the noun stem *lé* 'tree' (11b).

(10) a. bɔ̀ kálâdà   'books'
      b. á bɔ̂ kálâdà   'in the books'
(11) a. bì-lé   'trees'
      b. á bí-ꜜlé   'in the trees'

The phonetic notation in (12) shows that the initial /b/ of *bɔ̀* is not subject to lenition in intervocalic position, contrary to the initial /b/ of the gender prefix *bì-*.

(12) a. [ábɔ̂kálârà]   'in the books'
      b. [áβíꜜlɔ́]   'in the trees'

The phonotactic generalization that mid back vowels never occur in prefixes provides additional evidence for the word status of *bɔ̀*.

The important formal differences between *bɔ̀* and prefixes such as *bì-* raise the question of whether both have the same function. The answer is no. In Van de Velde (forthcoming) I argue that the nouns commonly referred to as class 1a are

outside of the gender system.[2] Gender and number are intimately linked. Whether a noun is singular or plural is determined by its gender membership. Therefore, genderless nouns need independent number marking. Their singular is zero-marked, and their plural is marked by means of *bɔ̀*, which is not a gender marker, but a plural word.

The argument is most easily illustrated by means of proper names. These can be derived from nouns, noun phrases or even clauses by means of the suffix *-A*. The harmonic vowel |A| is represented by a vowel phoneme identical to that of the preceding stem vowel, except if the latter is close. In that case the harmonic vowel is represented by /a/. If the stem has a structural succession of |u| and |a|, the harmonic vowel is represented by /ɔ/, according to the regular rules of vowel coalescence in Eton.

(13)   a. |ŋgùàg| → /ŋgwàg/ 'stone'
                [9]stone
         b. |ŋgùàg-À| → /ŋgwàgɔ̀/ 'Stone' (proper name)
                stone-SF

Usually the suffix has a low tone, but when it is attached to a monosyllabic noun stem with a rising tone, the high part of the stem tone attaches to the suffix.[3]

(14)   a. |d-ùád| → /dwǎd/ 'furuncle'
                5-furuncle
         b. |dùád-A| → /dwàdɔ́/ 'Furuncle' (proper name; elicited, not attested)
                furuncle-SF

During the derivation, the source noun is also lifted out of the gender system. The original prefix is incorporated into the stem, the original agreement pattern is replaced by agreement pattern I and the original plural formation by change in gender is replaced by the preposition of *bɔ̀*. This is true for all proper names, not only names for persons.

---

[2] Developing the arguments here would lead us too far afield. The claim is made for Eton only, but might prove to be valid in other Bantu languages as well. Note that I do not use the absence of a gender prefix on class 1a words as an argument. As has been said, gender 9 and 10 nouns do not have a prefix either. In order to accept the possibility of genderless nouns, it is important to acknowledge that agreement patterns can have multiple functions. In Van de Velde (forthcoming) I show that agreement pattern I marks agreement either with nouns of gender 1 or with nouns that are less in need of referential disambiguation. Proper names are a prototypical example of the latter, because they are strictly referential.

[3] This peculiar tonal behaviour is historically explainable. Final stem vowels eroded in Eton and the surrounding languages, except in proper names, which tend to be conservative. The final vowel of proper names has been morphologically reinterpreted as a suffix. Rising tones on monosyllabic stems are the result of the historical loss of the final vowel of CVCV-stems with a low-high tone pattern.

(15) a. |m-ùád| → /mwăd/ 'furuncles'
         6-furuncle

      b. |bɔ̀ # dùàd-ʌ| → /bɔ̀ dwàdɔ́/ 'Furuncle and his family/friends/people'
         PL   furuncle-SF

The examples in (16-18) of proper names for plant species are for different reasons morphosyntactically challenging, but they all confirm the analysis of *bɔ̀* as a separate word used to pluralize a genderless element. The name for mimosa in (16a) is a phrase containing an imperative and its object. The plural of this phrasal name is formed by *bɔ̀* in (16b). The high tone on the gender prefix of the object is due to high tone copy from the previous verb form. The following downstep shows that (16a) is not a compound. If there were no word boundary between *wú* and *ùswánɔ̂*, then the copied high tone would have deleted the low tone on the following prefix instead of pushing it to the right.

(16) a. /wú ú⁺swánɔ̂/
         |wú-H #   ù-súán-ʌ̀|
         die-IMP   3-shame-SF
         'plant (sp.), *Mimosa Pudica*' (lit.: 'die of shame')

      b. bɔ̀ wú ú⁺swánɔ̂ 'mimosa plants'

The names in (17-18) are noun phrases. In (17) the gender 5 noun *ɛ̂-sàŋ* 'hand of bananas' is followed by the numeral 'one', which should normally agree in gender, but which here takes a form of agreement pattern I instead of V. Interestingly, the pluralization of the phrasal name involves the pluralization of the head noun *and* the addition of the plural word before the entire phrase. The numeral remains in its singular form. Something similar occurs in the name for a kind of chili pepper (18). This is a connective construction with a head noun expressing a property followed by the connective proclitic (*H=*) plus a dependent noun. In the plural the head noun is pluralized and the connective morpheme still agrees with it. It is not clear whether the second noun is pluralized, since the plural form of *zùd* does not differ from the singular. Anyhow, the pluralization of its constituents does not suffice to derive the plural of the phrasal name. The plural word must again be added.

(17) a. /ɛ̂sàŋ m̀mwágɔ̂/
         |ɛ̂-sàŋ # Ǹ-vúág-ʌ̀|
         5-hand   I-one-SF
         'Horn plantain (sp.)'

      b. /bɔ̀ mèsàŋ m̀mwágɔ̂/
         |bɔ̀ # mè-sàŋ # Ǹ-vúág-ʌ̀|
         PL   6-hand   I-one-SF
         'Horn plantains (sp.)'

(18)   a.   /ńtwáŋ zûdà/
           |H-Ǹ-túáŋ # H=zùd-À|
           ?-3-long   III.CON=[9]buttock-SF
           'chili plant (sp.), *Capsicum frutescens*' (lit.: 'long buttock')

    b.   /bɔ̀ mìntwáŋ mí zûdà/
           |bɔ̀ # mìN-túáŋ # mí=zùd-À|
           PL   4-long   IV.CON=[9]buttock-SF
           'chili plants (sp.)'

As seen in the translation of (15b), the plural word usually has an associative meaning when combined with the name of a person. Via this associative meaning *bɔ̀* could combine with the general locative preposition *á* to form the preposition *ábɔ̂*, which means 'chez' (at or to somebody's place). In present-day Eton the locative preposition *á* is incompatible with nouns or pronouns that have human reference, whether genderless or not. If *bɔ̀* were a gender prefix, it could not have combined with a preceding preposition.

(19)   zùgá   èèy   ɲé   ábɔ̂   mǎ, méyâŋ   ɲé
        come   with him to   me   1SG.heal him
        'Bring him to me, so that I heal him.'

Interestingly, the origin of this preposition is reflected in Cameroonian French. When speakers tell me that a certain form is used in the dialect of a person we both know, for instance somebody called Désiré, he would tell me "that is how they say it chez les Désiré(s)," with a plural article before the first name of that person.

3.   A Minor Part-of-Speech Category

This section discusses the grammatical status of the plural word and provides a brief comparison with plural words in other languages. The plural word is always immediately followed by a genderless noun, which may be a phrasal name. It cannot be used in isolation, e.g. deictically or anaphorically. Nothing can be inserted between the plural word and the noun it pluralizes. This is because the order in nominal constituents is strictly head-before-dependent in Eton. From a syntactic point of view the plural word is the head of the noun it pluralizes. It is the plural word that determines gender agreement, not the following noun, which is genderless. Moreover, the augment is prefixed to the plural word, not to the following noun. The augment is a morpheme that is prefixed to the head of a nominal constituent if this head is modified by a demonstrative or a relative clause or if the head is normally a nominal modifier.

(20)   /íbɔ̂ kálâdà bá/
        |í-bɔ̂ # kálâdà # báǀ
        AU-PL book   2.DEM
        'these books'

Since *b-* is the prefix of gender 2 before vowel-initial stems, *bɔ̂* might be analyzable as *b-ɔ̂*.

There is one word in Eton that behaves similarly, viz. the diminutive proclitic *m-ɔ̂H*, plural *b-ɔ̂H*. This word can be put in front of any noun in order to form a diminutive or singulative. A combination of the diminutive word and a noun triggers gender 1 agreement in the singular and gender 2 agreement in the plural.

(21)   a.   /mɔ̂ ɛ́lén àté kù/
             |m-ɔ̂H=è-lén   # à-Lté # Lkùǀ
             1-DIM=5-palm.tree I-PR   fall
             'The small palm tree falls.'

        b.   /bɔ̂ mélén bé⁺té kù/
             |b-ɔ̂H=mè-lén   # bé-Lté # Lkùǀ
             2-DIM=6-palm.tree 2-PR   fall
             'The small palm trees fall.'

        c.   /bɔ̂ méjwág/
             |b-ɔ̂H=mè-júágǀ
             2-DIM=6-wine
             'some wine'

Thus, the plural word forms a minor part-of-speech category with the diminutive word. These words have in common with nouns that they determine concord, i.e. they belong to a gender and that they occupy the initial position in the noun phrase. They differ from nouns in their limited distribution and syntactic dependence and in their grammatical meaning.[4]

The Eton data confirm some observations that Matthew Dryer makes in a typological study on plural words (Dryer 1989). In a sample of 307 languages, Dryer found 48 languages in which the category plural is indicated by means of a separate word. These languages lack number as an inflectional category on the noun. Only two of these 48 languages are spoken in Africa, viz. Gbeya (Adamawa-Ubangi) and Yoruba (Benue Congo). The grammatical category of these plural words varies. In some languages the plural word is a numeral or an article. In others, plural words belong to a minor category or constitute a one-word category of their own.

---

[4] The plural word might be a proclitic as well, but the difference between words and proclitics is only visible in Eton if they carry a final high tone.

Dryer found two languages in his sample in which the plural word belongs to a minor category of noun modifiers: Hixkaryana (a Carib language) and Gbeya. In Hixkaryana, there are four other words in the category: a diminutive *txko* meaning 'small' or 'good'; an augmentative, meaning 'big' or 'bad'; a word indicating 'loss of value', and a word which means either 'dead' or 'set of'. In Gbeya there are three other words in the category that also contains the plural word (which is *ɔ́*): 'even, also, just', 'big, real', and 'a, some, certain, few'. Thus, as in Eton, these minor part-of-speech categories contain a word for 'small' and/or 'big'. Dryer further notes that in a number of other languages in his sample the category to which the plural word belongs (either articles or numerals) also contains a diminutive word.

Dryer also raises the issue of word order. He compared the order in the pair Plural word - Noun to that in the pair Verb - Object and found that in all VO languages the plural word precedes the noun (as in Eton), whereas the plural word follows the noun in the great majority of OV languages. Dryer proposes two possible explanations for this pattern: either plural words usually are the head of the noun phrase in which they occur, or this pattern confirms his claim that there is no general tendency for modifiers to follow the noun in VO languages and to precede in OV languages. The former explanation, i.e. that the plural word is the head of the following noun (in terms of dependency grammar), turns out to be correct in Eton.

4.   The Class 2a Marker in Other Bantu Languages

A superficial glance at descriptions of Bantu languages reveals that the so-called class 2a prefix very often differs from the other nominal gender prefixes. In many cases the differences are formal, e.g. a high tone, as in Lega (Botne 2003:427) or a long vowel, as in Mongo (Hulstaert 1965). In Yao the class 2a marker *atši-* is bisyllabic and in Karanga there are several forms, including *madzi-* and *vadzi-*. In Luganda the class 2a marker is a clitic, rather than a prefix (Hyman and Katamba, to appear). In Myene-Nkomi, finally, the class 2a marker is the only gender prefix that lacks an augment (Rekanga 2000). Often also the class 2a marker has specific meanings such as honorific or associative plural. Examples of the latter can be found in Mongo (22) and Xhosa (23).

(22)   a.   baa Byeka   'Byeka and his family/pupils/followers...'
       b.   baa mésá   'tables and similar things; tables, for instance'
                                           (Hulstaert 1965:145)

(23)   a.   oontlanzi < oo + iintlanzi   'fish and similar things'
       b.   oozingwe < oo + izingwe   'leopards and similar things'
       c.   oomaRhini < oo + amaRhini   'Grahamstown and environment'
       d.   ookulamba < oo + ukulamba   'hunger and similar feelings'
                                           (Hendrikse 1990:391)

Mark L. O. Van de Velde

The honorific meaning can be found for instance in Venda, as in (24), where *vho* is the class 2a prefix.

(24)   Vho-Vele vha khou ḓa
        'Vele is coming.' (Poulos 1990:20)

In some cases the class 2a marker spread to become a general plural marker, replacing or reinforcing markers of plural genders. This happens when the noun class system is disintegrating, as in Kinshasa Lingala (Aikhenvald 2000:388). It can also occur to some extent in language acquisition, as is illustrated by the data in Table 2. These are the results of an experimental study by Kunene (cited via Demuth 1985:311) on SiSwati in which children from 4.5 to 6 years old were given novel word forms and SiSwati nouns out of context and asked to provide the corresponding plural form of the noun. The right column show that young children overgeneralize the class 2a marker *bo* as a plural marker. It replaces the gender 2 and gender 4 prefixes and adds to the prefixes of genders 14 and 15.

(25) Table 2: Gender markers in the acquisition of SiSwati.

Class	Siswati	Experiment
1/2	*umu/ba*	*umu/bo*
1a/2a	*Ø/bo*	*Ø/bo*
3/4	*umu/imi*	*umu/bo*
5/6	*li/ema*	*li/ema*
7/8	*si/ti*	*si/ti*
9/10	*in/tin*	*i/ti*
11/10	*lu/tin*	*li/ema*
14	*bu*	*bu/bo-bu*
15	*ku*	*ku/bo-ku*

The forms and uses of the class 2a marker in the Bantu languages are in need of a thorough comparative study. In the meantime this brief overview suggests that the plural word analysis advanced for Eton might be valid for many other Bantu languages as well. A comparative study should also clarify the origin of the class 2a markers. Denis Creissels (p.c.) suggests that in Tswana the class 2a marker *bó* is the result of the coalescence of *ba* and *ga*, where *ba* is the connective morpheme of gender 2 and *ga* the connective of gender 17. The original meaning of *bó N* was 'the people at N's place'. This is a very attractive hypothesis for Tswana. However, it is not valid for Eton, since the plural word has a low tone, whereas all connectives have a high tone, except that of gender 1.

5.   Conclusion

Based on phonological criteria and the insight that there is a group of genderless nouns in Eton, the morpheme *bɔ̀* could be recognized as a plural word. According

to traditional criteria, however, it would be incorrectly analyzed as the prefix of class 2a, which is said to be a subgender of gender 2. The plural word analysis explains how *bɔ̀* was able to combine with the preposition *á* in order to form a new independent preposition *ábɔ̂* 'chez' and how it can pluralize idioms such as the phrasal names in (16)-(18). It is likely that a plural word analysis of the class 2a marker works for other Bantu languages as well. In that case, plural words are not as rare in Africa as Dryer's typological survey suggests. The fact that the plural word belongs to a minor part-of-speech category together with a diminutive word in Eton confirms some of Dryer's typological observations.

References

Aikhenvald, Alexandra. 2000. *Classifiers. A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

Botne, Robert. 2003. Lega (Beya dialect) (D25). In D. Nurse and G. Philippson, eds., *The Bantu Languages,* 422-449. London: Routledge.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Demuth, Katherine A. 1985. Noun classes and agreement in Sesotho acquisition. In M. Barlow and C. Ferguson, eds., *Agreement in Natural Languages: Approaches, Theories and Descriptions,* 305-321. Stanford: CSLI.

Doke, Clement. 1927. The significance of class 1a of Bantu nouns. In *Festschrift Meinhof*, 196-203. Hamburg: Kommisionsverlag von L. Friederichsen & Co.

Dryer, Matthew. 1989. Plural words. *Linguistics* 27:865-895.

Hendrikse, A.P. 1990. Number as a categorizing parameter in Southern Bantu: an exploration in cognitive grammar. *South-African Journal of African Languages* 10(4):384-400.

Hulstaert, Gustaaf. 1965. *Grammaire du lomongo. Deuxième partie: Morphologie*. Annales série in-8° sciences humaines n° 57. Tervuren: Musée Royale de l'Afrique Centrale.

Hyman, Larry M., and Francis Katamba. To appear. The word in Luganda. (To appear in the Proceedings of Cologne Conference on Typology, June 2001.)

Maho, Jouni. 1999. *A Comparative Study of Bantu Noun Classes*. Göteborg: Acta Universitatis Gothoburgensis.

Meeussen, A.E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3:81-121.

Poulos, George. 1990. *A Linguistic Analysis of Venda*. Pretoria: Via Afrika.

Poulos, George, and Louis J. Louwrens. 1994. *A Linguistic Analysis of Northern Sotho*. Pretoria: Via Afrika.

Rekanga, Jean-Paul. 2000. Le système myene-nkomi (groupe B10). In D.F. Idiata, M. Leitch, P. Ondo-Mebiane and J-P. Rekanga, eds., *Les classes nominales et leur sémantisme dans les langues bantu du nord-ouest,* 61-69. München: LINCOM.

Mark L. O. Van de Velde

Van de Velde, Mark L. O. Forthcoming. Multifunctional agreement patterns in Bantu and the possibility of genderless nouns. *Linguistic Typology*.

Departement Linguïstiek
Blijde-Inkomststraat 21,
3000 Leuven (Belgium)

Mark.Vandevelde@arts.kuleuven.be