

RENEWAL NOTICE

Dear Subscribers,

This is the last number of *Studies in African Linguistics* Volume 8. The editors hope that you will take the time now to renew your subscription so that you will continue to receive the journal without interruption.

As usual the annual subscription rate includes not only the three regular issues of the journal, but also supplements as they appear.

Please remove this form and send it with your remittance to

Studies in African Linguistics
Department of Linguistics
UCLA
Los Angeles, CA 90024
USA

NAME: _____

CURRENT ADDRESS: _____

Please check here if this is a new address.

Please renew my subscription to SAL for Volume 9 (1978). Enclosed is

_____ \$8.00 (individual domestic and overseas surface mail)

_____ \$15.00 (individual overseas air mail)

_____ \$12.00 (institutional subscription)

Make checks payable to African Studies Center.

TABLE OF CONTENTS

Articles

Malillo Morolong and Larry M. Hyman, ANIMACY, OBJECTS AND CLITICS IN SESOTHO	199
Gerard M. Dalgish and Gloria Sheintuch, ON THE JUSTIFICATION FOR LANGUAGE--SPECIFIC SUB--GRAMMATICAL RELATIONS	219
Patrick R. Bennett and Jan P. Sterk, SOUTH CENTRAL NIGER--CONGO: A RECLASSIFICATION	240
Paul Newman, CHADIC EXTENSIONS AND PRE--DATIVE VERB FORMS IN HAUSA . .	275
Jilali Saib, THE TREATMENT OF GEMINATES: EVIDENCE FROM BERBER . . .	299

ANIMACY, OBJECTS AND CLITICS IN SESOTHO¹

Malillo Morolong
Simon Fraser University

and

Larry M. Hyman
University of Southern California

In this paper a close look is taken at the object relation in Sesotho. The properties of word order, subjectivization, and cliticization are considered carefully as possible criteria for direct object status. The extreme importance of *animacy* in determining which argument(s) of a verb will receive direct object properties is demonstrated with respect to utterances with a benefactive and an accusative object, as well as sentences with an "affected possessor" and a possessed (body) part. A refinement in the use of cliticization as a direct object criterion is proposed on the basis of the need to distinguish between different *means* by which a clitic object marker can come into being (pronominalization, left-dislocation, relativization, object-agreement). It is concluded that the direct object is as much a *discourse* notion as it is a grammatical notion in Sesotho, as it probably is in (Eastern) Bantu as a whole.

1. Introduction

Over the past few years renewed interest in the nature of grammatical relations has resulted in a number of studies devoted to the nature of the (direct) object in various Central Bantu languages. While the approach has

¹An earlier version of this paper was presented on May 5, 1977 to the U.S.C. Seminar in Bantu Syntax. We would like to especially thank Annie Hawkinson and Linda Arvanites for their comments on this earlier version. The first author is grateful to Prof. Alfredo Hurtado, Chairman of Graduate Studies, Department of Modern Languages, Simon Fraser University, for initiating the visit to Los Angeles which resulted in this collaboration and to the Dean of Arts and the Vice-President of University Services for financing this trip.

sometimes varied, each study attempts to provide a structure of relations (grammatical or otherwise) into which the various arguments of the Bantu verb can be hierarchized. These attempts have at times met with certain difficulties, not the least of which is the problem of applying such traditional notions as "direct" vs. "indirect" object to the Bantu structures. A typical Bantu sentence consists of a subject, an agglutinated verbal complex, and a potential sequence of one to three nouns (or their corresponding pronouns, incorporated into the verbal complex). Since these nouns (and pronouns) are not case-marked, nor are they typically preceded by prepositions except in recent developments, the surface structure of such utterances as that from Sesotho in (1)

- (1) ke-bótselítsé baná ngoanána lipótso 'I asked the girl a question
I-asked/APP children girl question for the children'

deceptively suggests double (and even triple) accusatives, or direct objects. Recognizing, however, that such nouns display different grammatical properties (to be discussed below), Bantuists have developed terminology distinguishing a number of kinds of objects: principal, subsidiary, proximate, applied, directional, partitive, prepositional, etc.² These notions have in common the desire to develop appropriate terminology for Bantu, rather than the desire to fit Bantu within some universal framework. It is clear that Bantu has some of its own biases in the way it treats objects, but unless language is to have infinite ways of assigning hierarchical structure to the various arguments of a verb, we must argue that there is something basically non-arbitrary about the way Bantu accommodates its many objects.

In any study of the object relation in Bantu (whether comparative or language-specific), two separate steps must be taken. First, one must determine what the general properties of an object are. Second, one must determine *which* arguments (nouns and pronouns) can have access to these properties. Only after these two steps have been taken, and the results compared from languages representative of the Bantu zone, can we approach the most important question: *why* these properties and *why* these arguments?

²For Sesotho, see in particular Jacottet [1936:160-163] and Doke [1957:399-401].

It is these considerations which motivate the present study. In the following paragraphs, a close look is taken at the object relation in Sesotho, a Southern Bantu language (Guthrie S.33) spoken in Lesotho and parts of South Africa. As implied in the title, our concern will be, first, to demonstrate the role of animacy in determining object status, and second, to evaluate the use of cliticization as a criterion for establishing objecthood. It will be shown that non-grammatical considerations play a critical role in determining grammatical relations, and that the only adequate approach to the Sesotho (and Bantu) object relation is one which derives grammatical relations directly from the semantics and pragmatics of communication.

2. Animacy

While Kimenyi [1976], Gary and Keenan [1977], Duranti and Byarushengo [1977], and others, have used a number of additional tests, we shall be concerned here only with the following three general criteria for hierarchizing or "ranking" the arguments of a verb:

(i) *Word order*: Greater proximity to the verb may mean that a given noun has higher ranking (is more object-like) than a noun with lesser proximity to the verb.

(ii) *Subjectivization*: Greater accessibility to the subject slot of a passivized verb may mean that a non-subject noun in the corresponding active sentence is more object-like than one with lesser accessibility to the subject slot of the passive.

(iii) *Cliticization*: Greater accessibility to the clitic infix position may mean that the corresponding noun is more object-like than a noun with lesser accessibility to the clitic infix position.

A noun which can occur directly after the verb, which can be subjectivized in the corresponding passive, and which can be cliticized in the infix position is generally said to be a principal, proximate, or *direct* object (henceforth, DO). We shall have more to say about these notions as criteria for DO status, but first consider the following factors influencing the ranking of non-subject arguments (cf. Hawkinson and Hyman [1974]):

(i) *Case relations*: An argument whose semantic case is higher in the following hierarchy, BEN > DAT > ACC > INSTR [incomplete], will tend to have more DO properties than an argument whose semantic case is lower in the hierarchy.

(ii) *Animacy*: An argument whose referent is higher in the following *personal* hierarchy, 1st > 2nd > 3rd human > 3rd animal > 3rd inanimate [incomplete], will tend to have more DO properties than an argument whose referent is lower in this hierarchy (henceforth referred to as "animacy").

(iii) *Determinedness*: An argument whose referent is more determined (given, old, definite, specific) will tend to have more DO properties than an argument whose referent is less determined (new, indefinite, non-specific).

In this section we shall be concerned primarily with the role of animacy in determining DO status. In order to isolate the animacy factor the following benefactive paradigms will serve as test sentences:

(2) Test sentences

- | | | | |
|----|-------------------------------------|--------------|--------------|
| a. | I cooked food for the child | BEN=human | ACC=nonhuman |
| b. | I cooked food for the feast | BEN=nonhuman | ACC=nonhuman |
| c. | I called the children for the feast | BEN=nonhuman | ACC=human |
| d. | I called the children for the chief | BEN=human | ACC=human |

As can be seen to the right, all four possible combinations of human/non-human BEN/ACC are considered. The presence of a benefactive (whether human or non-human) is made possible by the *-il/el-* applicative extension (APP), which in the examples cited in this paper fuses with the *-il-* past tense suffix. We shall gloss verbs which carry this applicative extension as 'cook/APP', 'call/APP', etc., rather than attempting to indicate the morpheme breaks. We also will not normally provide an English translation, since the meaning will be clear from the word by word gloss which accompanies each example.

We begin, then, by considering the word order properties of test sentences a-d, in (3).

- (3) a. ke-phehétsé ngoaná lijó I-cooked/APP child food
 *ke-phehétsé lijó ngoaná I-cooked/APP food child

(3) cont.

- | | | |
|----|-----------------------------|-----------------------------|
| b. | ke-phehétsé mokété lijó | I-cooked/APP feast food |
| | ke-phehétsé lijó mokéte | I-cooked/APP food feast |
| c. | ke-bítselfítsé baná mokéte | I-called/APP children feast |
| | *ke-bítselfítsé mokété baná | I-called/APP feast children |
| d. | ke-bítselfítsé morena baná | I-called chief children [A] |
| | ke-bítselfítsé baná morena | I-called children chief [A] |

As seen from the second example in (3a) and (3c), when two nouns follow the verb, one of which is human, the other of which is nonhuman, the human noun *must*, independent of its semantic case (BEN or ACC), directly follow the verb. When both nouns are nonhuman (3b) or both nouns human (3d), both word orders are possible, resulting in ambiguity if both interpretations make sense. Both sentences in (3d) have the two meanings 'I called the children for the chief' and 'I called the chief for the children', and are therefore marked [A] for "ambiguous". The two sentences in (3b) mean 'I cooked the food for the feast', since the other interpretation 'I cooked the feast for the food' is not likely.³

In (4) we attempt to subjectivize each noun in the four test sentences.

(4) Subjectivization

- | | | |
|----|-----------------------------|---|
| a. | ngoaná ó-phehétsóé lijó | child he-was-cooked/APP food |
| | lijó lí-phehétsóé ngoaná | food it-was-cooked/APP child |
| b. | mokété ó-phehétsóé lijó | feast it-was-cooked/APP food |
| | lijó lí-phehétsóé mokéte | food it-was-cooked/APP feast |
| c. | baná bá-bítselfítsoé mokéte | children they-were-called/APP feast |
| | *mokété ó-bítselfítsoé baná | feast it-was-called/APP children |
| d. | morena ó-bítselfítsoé baná | chief he-was-called/APP children [A] |
| | baná bá-bítselfítsoé morena | children they-were-called/APP chief [A] |

³Compare, however, the following two sentences which are ambiguous:

- | | | |
|------|--------------------------|--------------------------|
| (i) | ke-rékétsé táfolé leséla | I-bought/APP table cloth |
| (ii) | ke-rékétsé leséla táfole | I-bought/APP cloth table |

Although it is more likely that someone would buy a cloth for a table than a table for a cloth, both sentences potentially carry both meanings.

Here we observe only one case where subjectivization is not possible: the second sentence of (4c), where there is a nonhuman BEN and a human ACC. Even the nonhuman ACC noun *lijó* 'food', which could not occur directly after the verb in (3a), can subjectivize, as seen in (4a).

Finally, turning to the third criterion, we attempt to cliticize each noun in the four test sentences in (5).

(5) Cliticization

- | | | |
|----|---------------------------------|-------------------------------------|
| a. | <i>ke-mó-phehétsé lijó</i> | I-him-cooked/APP food |
| | <i>ke-lí-phehétsé ngoaná</i> | I-it-cooked/APP child |
| b. | <i>ke-ó-phehétsé lijó</i> | I-it-cooked/APP food [it for feast] |
| | <i>ke-lí-phehétsé mokéte</i> | I-it-cooked/APP feast [it food] |
| c. | <i>ke-ba-bítselfítsé mokéte</i> | I-them-called/APP feast |
| | * <i>ke-o-bítselfítsé baná</i> | I-it-called children [it for feast] |
| d. | <i>ke-mo-bítselfítsé baná</i> | I-him-called/APP children [A] |
| | <i>ke-ba-bítselfítsé morena</i> | I-them-called/APP chief [A] |

Again, it is only the second sentence in (5c) which fails our test, and the two sentences in (5d) are ambiguous as to whether the noun or pronoun is BEN or ACC.

The following conclusions can be drawn from the data in (3), (4), and (5).

- (i) Human nouns precede nonhuman nouns.⁴
- (ii) When a nonhuman BEN combines with a human ACC (the opposite of one's expectancies), the former ceases to bear a direct relationship to the verb, and thus cannot undergo subjectivization and cliticization.
- (iii) When both nouns are equal in animacy, there is potential ambiguity, with both word orders being possible. Note that the choice of word order in such cases will depend on the givenness/newness of each argument. Thus, the first sentence in (3b) answers the question 'what did you cook for the feast?', while the second sentence answers the question 'what did you cook the food for?'

⁴Cf. Jacottet [1936:161]: "If the *two objects are nouns*, both are of course placed after the verb, the nearest object (of the person) coming first, the other object (of the thing) following it" [his emphasis].

We tentatively conclude, then, that except for *mokéte* 'feast' in the (c) sentences, both nouns are DO's in all of the above examples. This includes *lijó* 'food' in (3a), even though it cannot occur directly after the verb. We maintain that it is a DO whose position, however, is restricted by a surface constraint disallowing the postverbal sequence *nonhuman - human. What will be interesting to us is explaining why (4a) and (5a) have different grammatical properties from (4c) and (5c), since both the (a) and (c) sentences involve a human and a nonhuman. We note, finally, that unlike other Bantu languages, there is no evidence that the BEN should optimally ("in the unmarked case") precede the ACC. Instead, the only relevant consideration for determining the ordering of nonsubject nouns is animacy.⁵

3. Cliticization

As seen in (4) above, one of the most widely cited criteria for establishing DO status is the ability of a noun to be "replaced" by a corresponding clitic pronoun in the object marker (OM) infix slot in the verbal complex. In Bantu, there are potentially four different ways in which an OM clitic can arise in discourse, all of them attested in Sesotho:

(i) *Pronominalization*: When a noun is "replaced" by a pronoun, the latter typically surfaces in the OM position immediately preceding the verb stem, e.g. *ke-a-mo-bóna* 'I see him' (lit. I-PRES-him-see).

(ii) *Left-dislocation*: When a noun is left-dislocated, Bantu languages usually require a "copy pronoun" in the verbal complex, e.g. *ngoaná ke-a-mo-bóna* 'the child, I see him' (lit. child I-PRES-him-see).

(iii) *Relativization*: Many Bantu languages require a "resumptive pronoun" when relativizing on a noun, e.g. *ngoaná éò ké-mo-bóná-ng* 'the child that I see' (lit. child that I-him-see-REL).

⁵The situation is nearly the opposite in other Bantu languages, e.g. Shona [Hawkinson and Hyman 1974], Logooli [Rachel Angogo, personal communication]. In these languages the order of the two nouns is flexible *only* if no ambiguity results. Thus, if both nouns were human, the BEN would have to precede the ACC. In Logooli, if the BEN is human and the ACC nonhuman, both orders are possible; if, on the other hand, the BEN is nonhuman and the ACC human, only the order ACC-BEN is possible.

(iv) *Object agreement*: Some Bantu languages allow an OM clitic to co-occur with its coreferential noun after the verb.⁶ In Sesotho only the "long" form of the present tense allows this object agreement, e.g. *ke-a-mo-bóna ngoaná* 'I see the child' (lit. I-PRES-him-see child).

Except for the northwestern extremity of the zone, all Bantu languages use the OM clitic for the purpose of pronominalization. Most, if not all, of the same languages allow or require the OM clitic when its coreferential noun is left-dislocated. Bantu languages split on the question of whether an OM clitic is required when an object is relativized, while relatively fewer Bantu languages make use of true object agreement. The reason for distinguishing the above four *functions* of the OM clitic is that, in Sesotho, the ability to obtain cliticization depends upon the way in which the clitic comes into being. Since it is restricted to one tense only, we shall not discuss object-agreement any further in this study.

We begin by left-dislocating in (6) and relativizing in (7) each noun in our original test sentences.

(6) Left dislocation

a.	<i>ngoaná ke-mó-phehétsé lijó</i>	child I-him-cooked/APP food
	<i>lijó ke-lí-phehétsé ngoaná</i>	food I-it-cooked/APP child
b.	<i>mokété ke-ó-phehétsé lijó</i>	feast I-it-cooked/APP food
	<i>lijó ke-lí-phehétsé mokéte</i>	food I-it-cooked/APP feast
c.	<i>baná ke-ba-bítselítse mokéte</i>	children I-them-called/APP feast
	<i>*mokété ke-o-bítselítse baná</i>	feast I-it-called/APP children

⁶Object-agreement usually marks the object noun as given or definite, e.g. in Sukuma, where there is a curious interplay between object-agreement and the presence vs. absence of the preprefix [Herman Batibo, personal communication]. In some Bantu languages, e.g. Swahili, Nyakyusa [Duranti 1977], object-agreement is obligatory with human object nouns. It is these object-agreeing languages which tend also to restrict DO properties to a single argument in the sentence. Note, also, that many Bantu languages have a process of right-dislocation, e.g. Haya [Byarushengo, Hyman, and Tenenbaum 1976, Tenenbaum 1977a,b]. In such utterances the clitic OM is to be identified as a pronoun, rather than as an object-agreement marker (cf. Byarushengo and Tenenbaum [1976]).

(6) cont.

- d. morena ke-mo-bítseítsé baná chief I-him-called/APP children [A]
baná ke-ba-bítseítsé morena children I-them-called/APP chief [A]

(7) Relativization

- a. ngoaná éò ké-mó-phehétséng lijó child that I-him-cooked/APP food
lijó tsèò ké-lí-phehétséng ngoaná food that I-it-cooked/APP child
- b. mokété óò ké-ó-phehétséng lijó feast that I-it-cooked/APP food
lijó tsèò ké-lí-phehétséng mokéte food that I-it-cooked/APP feast
- c. baná báò ké-ba-bítseítséng mokéte children that I-them-called/APP feast
*mokété óò ké-o-bítseítséng baná feast that I-it-called/APP children
- d. morena éò ké-mo-bítseítséng baná chief that I-him-called/APP children [A]
baná báò ké-ba-bítseítséng morena children that I-them-called/APP chief [A]

The sentences in (6) and (7) confirm our earlier observation that only *mokéte* 'feast' in the (c) sentences is not a DO. Thus, cliticization by left-dislocation and by relativization yields the same results as cliticization by pronominalization in (5).

A complicated situation arises, however, when we attempt to subjectivize one noun and pronominalize the other, as seen in (8).

(8) Subjectivization of one noun, pronominalization of the other

- a. ngoaná ó-li-phéhetsoe child he-it-was-cooked/APP
*lijó lí-mo-phéhetsoe food it-him-was-cooked/APP
- b. *mokété ó-li-phéhetsoe feast it-it[food]-was-cooked/APP [R]⁷
*lijó lí-o-phéhetsoe food it-it[feast]-was-cooked/APP
- c. *baná bá-o-bítseílitsoe children they-it-were-called/APP [R]
*mokété ó-ba-bítseílitsoe feast it-them-was-called/APP
- d. ?morena ó-ba-bítseílitsoe chief he-them-was-called/APP [A]
?baná bá-mo-bítseílitsoe children they-him-were-called/APP [A]

⁷The symbol [R] indicates that a reversed meaning is obtained, e.g. 'the feast was cooked for it [the food]' in (8b).

In (8) we observe that when one noun is subjectivized and the other pronominalized (as a clitic), only the first sentence in (8a) is completely acceptable. In order for such a structure to evolve it must be the case that the subject of the passivized verb is human. The first sentence in (8a) is distinctively less acceptable if 'goat' is substituted for 'child'. This explains the unacceptability of the second sentence in (8a) and the two sentences in (8b). The unacceptability of the two sentences in (8c) can be attributed to the fact that *mokéte* 'feast' is not a DO and therefore cannot appear in either the subject or OM position in general. The two sentences in (8d) are marginally acceptable. Bantu languages differ in their treatment of these: for some what is important is that the subject be human, i.e. highest in the personal hierarchy; for others what is important is that there be an animacy *differential* between the subject and the OM clitic pronoun, where the subject is higher in animacy than the object. The questionable status of these sentences seems to indicate an awareness of both strategies: the fact that the subject is human argues in favor, while the fact that the OM clitics are *equally* human argues against the acceptability of (8d). If we substitute first or second person subject pronouns for 'chief' and 'child', (8d) becomes more acceptable, since a new differential along the "animacy" scale is created. Conversely, if we substitute first or second person *object* pronouns for 'them' and 'him', both sentences in (8d) become entirely ungrammatical.

What is important for our discussion of cliticization is that the acceptability of some of the sentences in (8d) is enhanced when the OM clitics arise through left-dislocation, in (9), or relativization in (10).

(9) Subjectivization of one noun, left-dislocation of the other

- | | | |
|----|-------------------------------|--|
| a. | lijó ngoaná ó-li-phéhetsoe | food child he-it-was-cooked/APP |
| | *ngoaná lijó lí-mo-phéhetsoe | child food it-him-was-cooked/APP |
| b. | *lijó mokété ó-li-phéhetsoe | food _i feast _j it _j -it _i -was-cooked/APP |
| | *mokété lijó lí-o-phéhetsoe | feast _i feast _j it _j -it _i -was-cooked/APP |
| c. | *mokété baná bá-o-bítsélitsoe | feast children they-it-were-called/APP [R] |
| | *baná mokété ó-ba-bítsélitsoe | children feast it-them-was-called/APP |
| d. | baná morena ó-ba-bítsélitsoe | children chief he-them-was-called/APP [A] |
| | morena baná bá-mo-bítsélitsoe | chief children they-him-were-called/APP [A] |

(10) Subjectivization of one noun, relativization of the other

- a. lijó tséò ngoaná á-lí-phehétsoèng food that child he-it-was-cooked/APP
 *ngoaná éò lijó lí-mó-phehétsoèng child that food it-him-was-cooked/APP
- b. *lijó tséò mokété ó-lí-phehétsoèng food_i that feast_j it_j-it_i-was-cooked/APP
 *mokété óò lijó lí-ó-phehétsoèng feast_i that food_j it_j-it_i-was-cooked/APP
- c. mokété óò baná bá-ó-bítseíítsoèng feast that children they-it-were-called/APP
 *baná bàò mokété ó-bá-bítseíítsoèng children that feast it-them-was-called/APP
- d. baná bàò morena á-bá-bítseíítsoèng children that chief he-them-was-called/APP [A]
 morena éò baná bá-mó-bítseíítsoèng chief that children they-him-were-called/APP [A]

Comparing (8) vs. (9) and (10) we note the same ungrammaticality of the second sentence in (a) and (c), as well as the two sentences in (b). The two differences observed are: (i) the sentences in (d), which were questionable in (8), are acceptable in both (9) and (10); and (ii) the first sentence of (c), which was unacceptable in (8), and is unacceptable in (9), is acceptable in (10). In other words, when the coreferential nouns are present, and the OM clitic is created through left-dislocation or relativization, the animacy constraint is relaxed in the (d) sentences. Also, *mokéte* 'feast', which could not cliticize through pronominalization in (8c), nor through left-dislocation in (9c), *can* cliticize through relativization in (10c). There appear to be two ways to interpret this last fact: (i) cliticization remains a DO property, in which case (under certain circumstances not fully understood) a non-DO *becomes* a DO when it is relativized; or (ii) cliticization is not a DO property *except* in so far as the clitic is a *true* pronoun. While it would be instructive to compare the strategies of a number of Sesotho speakers, it appears that one important variable in determining the acceptability of an utterance is whether an antecedent noun cooccurs or not with its corresponding clitic. This would predict the difference between (8) vs. (9) and (10), but *not* the difference between the first sentences of (9c) and (10c). Thus, in addition to the presence vs. absence of the antecedent noun, we may have to establish a hierarchy of cliticizability, depending upon the means by which the clitic is produced.⁸

⁸A most convincing case involving object-agreement was pointed out to

Before leaving this section we should note, however, that the ability of a noun to be pronominalized does not in itself imply that that argument is a DO. This is because in addition to the OM clitics, there exists a series of "absolute pronouns" [Doke 1957:107ff], one of whose functions is to provide a pronoun when a clitic OM is not possible. Since, unlike certain other Bantu languages, there is only one OM clitic slot, an absolute pronoun must be used if both nonsubject nouns in the test sentences are pronominalized, as seen in (11).

(11) Pronominalization of both nouns

- | | | |
|----|------------------------|---------------------------------|
| a. | ke-mó-phehétsé tsoná | I-him-cooked/APP it |
| | ke-lí-phehétsé eená | I-it-cooked/APP him [A] |
| b. | ke-lí-phehétsé ooná | I-it[food]-cooked/APP it[feast] |
| | ke-ó-phehétsé tsoná | I-it[feast]-cooked/APP it[food] |
| c. | ke-ba-bítselfítsé ooná | I-them-called/APP it[feast] |
| | *ke-o-bítselfítsé boná | I-it[feast]-called/APP them |
| d. | ke-mo-bítselfítsé boná | I-him-called/APP them [A] |
| | ke-ba-bítselfítsé eená | I-them-called/APP him [A] |

In (11) we see that it is only the second sentence of (11c) which cannot allow two pronouns (since *mokéte* 'feast' is not a DO). A second environment where absolute pronouns substitute for ungrammatical clitic pronouns is in the following sentences corresponding to those in (8):

us by Herman Batibo [personal communication]. In his language, Sukuma, the following sentence is ungrammatical:

- | | | |
|-----|------------------------|-----------------------------|
| (i) | *a-ka-bii-nhwá | 'he was given them [human]' |
| | he-past-them-was given | |

This shows that a passive construction will not tolerate a human pronoun in the OM slot. However, as seen in (ii),

- | | | |
|------|----------------------|-----------------------------|
| (ii) | a-ka-bii-nhwá abaana | 'he was given the children' |
|------|----------------------|-----------------------------|

the sentence is acceptable if the preprefixed (= [+determined]) object noun cooccurs with the OM clitic. What could not cliticize as a pronoun *can* cliticize as an agreement marker!

(12) Sentences from (8) with absolute pronouns

- a. ?ngoaná ó-phehétsoé tsoná child he-was-cooked/APP it
lijó lí-phehétsoe eená food it-was-cooked/APP him
- b. ?mokété ó-phehétsoé tsoná feast it-was-cooked/APP it[food] [R]
lijó lí-phehétsoé ooná food it-was-cooked/APP it[feast]
- c. baná bá-bítselfitsoé ooná children they-were-called/APP it
*mokété ó-bítselfitsoé eená feast it-was-called/APP them
- d. morena ó-bítselfitsoé boná chief he-was-called/APP them [A]
baná bá-bítselfitsoé eená children they-were-called/APP him [A]

Except for the second sentence of (12c), those sentences which were ungrammatical (or questionable) with a clitic pronoun in (8) are now acceptable with the absolute pronoun (the questionable status of the first sentence of (12b) apparently is due to the tendency to assign a reversed meaning 'the feast was cooked for it'). Note that the fully grammatical first sentence of (8a) is now questionable when an absolute pronoun is used instead of the correct clitic pronoun. A general feature of Sesotho is that when a clitic pronoun is acceptable, the corresponding sentence with an absolute pronoun in its place is questionable.⁹

Since pronominalization is possible without cliticization, and since cliticization is possible without pronominalization, we are left with the conclusion that if cliticization is to serve as a criterion for DO status in Sesotho, we must further stipulate that *the clitic must be a pronoun*. What this means about the nature of the DO will be discussed in section 5.

4. Affected Possessors

One of the perplexing issues arising from the data presented above concerns the non-object status of *mokéte* 'feast' in the (c) sentences. Recall that the test sentence (2c) has a nonhuman BEN and a human ACC. We know from test sentence (2a), which has a human BEN and a nonhuman ACC, that a verb can in principle have two DO's, one of them human and the other nonhuman. But in order for it to do so, it must be the case that

⁹Doke [1957:108] indicates that the absolute pronouns can at least in some cases be substituted for the clitic pronouns with an "emphatic" function.

the BEN is human and the ACC nonhuman, rather than the reverse. The question is why?

After investigating a number of constructions in the language, we have determined that test sentence (a) aligns itself with such double object verbs as 'to give', while test sentence (c) has its grammatical properties mirrored in constructions involving an "affected" possessor. As shown by Voeltz [1976], these involve cases where there is a part-whole relationship between two arguments and, it is important to add, where the whole is affected by an action performed on the part (cf. Hyman [1977]). Our test sentence here will be 'I broke the child's arm', where my breaking the arm necessarily affects the child as a whole (cf. 'I broke the child's stick', where the stick can be broken without affecting the child). In (13), as seen before, the human noun must precede the nonhuman noun.

- (13) a. ke-robílé ngoaná letsóho I-broke child arm
 b. *ke-robílé letsóhó ngoaná I-broke arm child

(Note that the -ílé ending is the past tense marker and does *not* contain the applicative extension.) The normal possessive construction in (14)

- (14) ke-robílé letsóhó lá ngoaná I-broke arm of child

implies that the child is not affected by the breaking of the arm and thus communicates that the arm which was in his possession was not part of his body, e.g. he found a stray arm, brought it to me, and I broke it-- thus without the action of arm-breaking having any effect on the child. When there is no part-whole relationship (and it is therefore harder to affect the possessor by acting on his possession), only the possessive construction is possible, as seen in (15):

- (15) a. ke-robílé lekala lá ngoaná I-broke branch of child
 b. *ke-robílé ngoaná lekala I-broke child branch

What is important for our study of the object relation is that letsóho 'arm' is, like mokéte 'feast' in test sentence (c), not a DO. This is seen in (16), where 'arm' does not subjectivize, and in (17), where it does not cliticize.

- (16) a. ngoaná ó-robíloé letsóho child he-was-broken arm
 b. *letsóhó lé-robíloé ngoaná arm it-was-broken child
- (17) a. ke-mó-robílé letsóho I-him-broke arm
 b. *ke-lé-robílé ngoaná I-it-broke child

As in (8c), both (18a) and (18b) are ungrammatical when one noun is subjectivized and the other cliticized as a pronoun:¹⁰

- (18) a. *ngoaná ó-le-róbiloe child he-it-was-broken
 b. *letsóhó lé-mo-róbiloe arm it-him-was broken

Corresponding exactly to the sentences in (9c) and (10c) are those in (19), where one noun is left-dislocated and the other cliticized as a pronoun, and (20), where one noun is relativized and the other cliticized as a pronoun.

- (19) a. *letsóhó ngoaná ó-le-róbiloe arm child he-it-was-broken
 b. *ngoaná letsóhó lé-mo-róbiloe child arm it-him-was-broken
- (20) a. letsóhó léò ngoaná á-lé-robíloéng arm that child he-it-was-broken
 b. *ngoaná èò letsóhó lé-mó-robíloéng child that arm it-him-was-broken

We therefore conclude that whatever is responsible for the non-DO status of *mokéte* 'feast' in test sentence (c) must also be responsible for the non-DO status of *letsóho* 'arm' in the above examples.¹¹ Before attempting

¹⁰Unlike the first sentence of (12c) the corresponding affected object constructions with absolute pronouns are for some reason still unacceptable:

- (i) *ngoaná ó-robíloé loná child-he-was-broken it
 (ii) *letsóhó lé-robíloé eená arm it-was-broken him

¹¹Further evidence for a relationship between test sentence (c) and the affected object construction is seen from Logooli [Rachel Angogo, personal communication]. In that language, a dialect of Luhya, both nouns in the corresponding sentences can subjectivize and cliticize (although the human ACC and affected object must precede the nonhuman BEN and possessed part, respectively). Of all of the Bantu languages we have looked at, Logooli seems to be the least affected by animacy considerations. Logooli can even freely cliticize human pronouns when a non-human argument is subjectivized:

- (i) ichú'kúríá cha-ḿ-deekerwâ food it-him-was-cooked/APP
 (ii) isúguukú 'yá-ḿ-'rááŋgirwâ feast it-him-was-cooked/APP

a synthesis of the data presented thus far, let us note in (21) that the "affected" possessor need not be human.

- (21) a. ke-robílé sefáté lekala I-broke tree branch
 b. ke-robílé lekala lá sefáte I-broke branch of tree

In (21a) *sefáté* 'tree' appears as the affected possessor, indicating that the branch-breaking has affected the tree. This sentence is thus most aptly translated 'I broke the branch off the tree'. When the tree is not affected, e.g. the branch had already been severed from the tree before I broke it, the ordinary possessive construction is used. Thus, sentence (21b) is best translated 'I broke the branch of the tree' or even 'I broke the tree-branch' (generic).

5. Summary and Conclusion

To summarize, we have observed the following:

- (i) There is a constraint against having a nonhuman noun precede a human noun after the verbal complex.
- (ii) When the BEN and ACC are of equal animacy, both word orders are permitted, with potential ambiguity in each case.
- (iii) When the humanness/nonhumanness of a BEN and ACC are varied, all nouns can subjectivize and cliticize *except* the BEN, if the BEN is nonhuman and the ACC is human.
- (iv) If one argument is subjectivized and the other cliticized as a pronoun, the optimal situation obtains when the resulting subject is human and the cliticized pronoun nonhuman.
- (v) In certain constructions such as that described in (iv), cliticization is facilitated if arising out of left-dislocation or relativization, rather than pronominalization.
- (vi) An absolute pronoun follows the verbal complex if either (a) there is already a clitic occupying the OM slot, or (b) the antecedent noun does not have access to cliticization.
- (vii) Possessed parts do not subjectivize or cliticize in affected possessor constructions and resemble the nonhuman BEN described in (iii).

From the above observations we conclude that arguments higher in the personal hierarchy, e.g. human vs. nonhuman arguments, are accorded more DO properties than arguments lower in the personal hierarchy. Thus, human

nouns obligatorily precede non-human nouns and have greater access to subjectivization and cliticization.

At this point we would like to suggest that the properties said to characterize DO's are accorded to those non-subject arguments which are the most *prominent* in discourse, either as the syntacticization of universal tendencies, or, if the language permits flexibility, as warranted by a particular discourse situation. Human beings necessarily have greater prominence over nonhumans, since they typically bring about, receive, and are the beneficiaries of actions. Thus, in Sesotho, the prominent position immediately following the verbal complex is reserved for a human noun (if present in the utterance).

Addressing ourselves to (iii), we note that there is a tendency for the BEN to be human and the ACC to be nonhuman. Thus, in terms of both the personal and case hierarchies, the BEN should be accorded grammatical properties commensurate with its relative prominence over the ACC. Thus, it is not surprising that Bantu chooses, via the applicative extension *-il/el-*, to orient the verb towards the BEN (and, in some Bantu languages, away from the ACC, which may receive few DO properties). As a result of this benefactively oriented verb-marking, the human BEN is "secure" enough in its grammatical status in (8a) to allow a nonhuman ACC to cooccur as a clitic pronoun. On the other hand, when the BEN and ACC are equal in animacy, instead of the BEN remaining higher in grammatical status, in Sesotho (and other Bantu languages) a "stand-off" situation arises with both arguments being DO-like, but with neither being too secure in its objecthood. This is seen in (8b), where neither subjectivized argument is secure enough in grammatical status to permit a cooccurring clitic pronoun. Finally, consider the case of the nonhuman BEN and the human ACC. The verb is oriented toward the BEN but on universal grounds, the human argument is more prominent. Because Sesotho is so much more animacy-oriented than case-oriented, it balks the BEN verb-marking and gives all to the ACC. However, because the human ACC is not singled out by the grammar to receive the grammatical correlates of discourse-prominence, (either by verb-marking or in the lexicon), it is necessary to denude the BEN of all grammatical status. The (c) sentences in the data cited throughout this study all point to the non-DO status of this nonhuman BEN.

The same can be said about the nonhuman possessed part in section 3. Here, the fact that the possessor of a body (or tree) part is *affected* results in its being "promoted" to prominent grammatical status. In fact, as argued by Hyman [1977], in the affected possessor construction, the relative statuses of the possessor and the possessed are *inverted*.¹² When we say in Sesotho 'I broke the tree the branch', it is the tree which is prominent. Unlike the (c) sentences, the prominence is not due to the animacy differential between the two arguments, but rather to the *affectedness* of the possessor, i.e. the tree.

To be a DO therefore *means* something in Bantu. Whether it be for its animacy, its case, or its determinedness (cf. section 1), a DO in Bantu is more prominent, affected, or topical than a non-DO. We have seen in a number of examples above the need to address the animacy factor in determining grammatical status. To maintain the DO as a notion devoid of meaning or discourse function makes it difficult to explain why animacy should play any role at all in determining the grammatical status of an argument. We have seen that the various arguments of a verb must be hierarchized on the basis of three scales. We have also seen that the grammatical status of one argument may crucially depend on the nature or presence vs. absence of another argument or arguments. Finally, we have experienced some difficulty in applying the cliticization criterion for DO status. In crucial examples where there are rival claims to the OM clitic position, the resulting sentences are hard to evaluate.¹³ This

¹²In both Kinyarwanda [Kimenyi 1976] and Logooli [Rachel Angogo, personal communication] the affected possessor becomes a DO without the possessed part losing its DO status. An analysis involving inversion of grammatical relations would thus not be appropriate for these languages.

¹³Such sentences involve either (i) one argument pronominalized, the other left-dislocated; (ii) one argument pronominalized, the other relativized; or (iii) one argument left-dislocated and the other relativized. Since each such sentence has *two* arguments vying for the clitic OM position, it would be instructive to submit our test sentences (2a-d) to these operations and determine from a number of speakers a hierarchy for accessibility to clitic object position.

raises the important question of how much of the foregoing is characteristic of all Sesotho, and how much should be attributed to the discourse strategies of the two speakers we were able to consult. Until a systematic study involves the participation of a large number of native speakers, we cannot know whether the variations found in the diverse Bantu languages are characteristic of grammatical differences among languages or strategic differences among speakers. For what we have investigated so far we are only able to say: (i) that cliticization can occur without pronominalization; (ii) that pronominalization can occur without cliticization; and (iii) what occurs as a clitic pronoun can also be the subject of a corresponding passive. Whether subjectivization and clitic pronominalization constitute a test for direct objecthood may simply be a terminological problem. What is clear is that both the properties generally used to establish objecthood and the factors influencing accessibility to these properties find their *raison d'être* in the semantics and pragmatics of communication.

REFERENCES

- Byarushengo, Ernest Rugwa, Larry M. Hyman, and Sarah Tenenbaum. 1976. "Tone, accent, and assertion in Haya." In Larry M. Hyman (ed.), *Studies in Bantu Tonology*, pp. 183-205. Southern California Occasional Papers in Linguistics No. 3. Los Angeles: University of Southern California.
- Byarushengo, Ernest Rugwa and Sarah Tenenbaum. 1976. "Agreement and word order: a case for pragmatics in Haya." *Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society*, pp. 89-99. Berkeley: University of California.
- Doke, C.M. 1947. *Textbook of Southern Sotho Grammar*. London: Longmans, Green and Co.
- Duranti, Alessandro. 1977. "Notes on Nyakyusa syntax." Ms., University of Southern California and University of Rome.
- Duranti, Alessandro, and Ernest Rugwa Byarushengo. 1977. "On the notion of 'direct object'." In Ernest Rugwa Byarushengo, Alessandro Duranti, and Larry M. Hyman (eds.), *Haya Grammatical Structure*, pp. 45-71. Southern California Occasional Papers in Linguistics No. 6. Los Angeles: University of Southern California.

- Gary, Judith Olmsted, and Edward Louis Keenan. 1977. "On collapsing grammatical relations in universal grammar." In Peter Cole and Jerrold M. Sadock (eds.), *Syntax and Semantics 8: Grammatical Relations*, pp. 83-120. New York: Academic Press.
- Hawkinson, Anne K., and Larry M. Hyman. 1974. "Hierarchies of natural topic in Shona." *Studies in African Linguistics* 5:147-70.
- Hyman, Larry M. 1977. "The syntax of body parts." In Ernest Rugwa Byarushengo, Alessandro Duranti, and Larry M. Hyman (eds.), *Haya Grammatical Structure*, pp. 99-117. Southern California Occasional Papers in Linguistics No. 6. Los Angeles: University of Southern California.
- Jacottet, E. 1936. *A Practical Method to Learn Sesuto*. Farnborough Hants, England: Gregg Press Limited. [Reprinted 1968.]
- Kimenyi, Alexandre. 1976. "A relational grammar of Kinyarwanda." Doctoral dissertation, University of California, Los Angeles.
- Tenenbaum, Sarah. 1977a. "On reference in Haya." In Elinor Ochs Keenan and Tina L. Bennett (eds.), *Discourse Across Time and Space*, pp. 283-96. Southern California Occasional Papers in Linguistics No. 5. Los Angeles: University of Southern California.
- Tenenbaum, Sarah. 1977b. "Left- and right-dislocation." In Ernest Rugwa Byarushengo, Alessandro Duranti, and Larry M. Hyman (eds.), *Haya Grammatical Structure*, pp. 161-70. Southern California Occasional Papers in Linguistics No. 6. Los Angeles: University of Southern California.
- Voeltz, Erhard F.K. 1976. "Inalienable possession in Sotho." In Larry M. Hyman, Leon C. Jacobson, and Russell G. Schuh (eds.), *Papers in African Linguistics in Honor of Wm. E. Welmers*, pp. 255-266. *Studies in African Linguistics*, Supplement 6.

ON THE JUSTIFICATION FOR LANGUAGE-SPECIFIC
SUB-GRAMMATICAL RELATIONS*

Gerard M. Dalgish
University of Dar Es Salaam

and

Gloria Sheintuch
University of Illinois

This paper examines the syntactic behavior of a semantically specifiable set of NP's, the locative NP's, in OluTsootso, a Bantu language. We attempt to determine their grammatical relation in sentences, and find that the traditional and generative analysis of these elements as prepositional-phrase non-terms is inadequate. We claim that these locatives constitute a sub-relation of the grammatical relation of direct object. We do this by demonstrating that the set of locatives in question (which we call "locative objects") can be sharply distinguished from prepositional phrase locatives on the basis of behavioral and coding properties. We then show that locative objects and direct objects are the only elements which undergo certain relation-changing processes, trigger similar verbal cliticization processes in pronominalization, and that as targets of relativization and topicalization, have similar but unique morphosyntactic outputs. We show that indirect objects are placed lower on the Accessibility Hierarchy for grammatical relation-changing rules than locative objects, which results in needlessly violating an otherwise universal constraint on subjectivization rules, given the (incorrect) assumption that locative objects are prepositional phrases. We then show that although there is evidence that locative objects and direct objects should be classified together, we find further evidence which suggests that there is a broader grammatical relation of direct object which encompasses the sub-relations of (regular) direct objects and locative objects.

1. Introduction

This paper examines the syntactic behavior of a semantically and syntactically specifiable set of NP's in OluTsootso, a dialect of Luyia spoken

*We wish to convey our special thanks to Charles Kisseberth for helpful discussions and comments. This research was supported in part by NSF grant SOC 75-0024.

in western Kenya, in an attempt to determine the grammatical relation such NP's play in sentences of the language. More specifically, this semantically and syntactically specifiable set of NP's consists of locative NP's taking as prefix one of three locative markers *xu-*, *mu-*, and *ha-*, meaning 'on', 'in', and 'near' respectively. The analyses of such locatives available to date in the literature, be they traditional or transformational in approach, classify them as prepositional phrases, not bearing any grammatical relation to the verb in the sentence. It is shown that such analyses are inadequate, and that these locatives constitute a semantically and syntactically distinct subrelation of the grammatical relation direct object.

Some information about OluTsootso morphology is useful at the beginning of the paper to help clarify the ensuing data. OluTsootso has typical Bantu morphological characteristics including noun classes which govern agreement on verbs for subjects and direct objects and on other grammatical elements such as relative pronouns, demonstratives and possessives. The verbal morphology consists of subject and object agreement prefixes as well as derivational verbal suffixes, among them the "applied" suffix *-il/-el* triggered in general by a benefactive, instrumental, or directional NP in the sentence, the passive suffix *-u/-ibw*, and the locative suffix *-xwo/-mwo/-ho*, each agreeing with one of the three locative markers mentioned in the first paragraph. The distribution of this locative suffix is of immediate relevance to this study, and will be taken up in detail later on in the paper.

2. NP Locatives and Prepositional Locatives

The set of locatives investigated in this study, namely those which take one of the three locative markers *xu-*, *mu-*, and *ha-*, are to be distinguished from the non-term, prepositional phrase type locatives whose prepositions constitute separate words. Contrast, for example, the locative in (1), carrying the prefix *xu-*, with the locative prepositional phrases in (2) with the sequence of prepositions *inyuma ya* or *imbeli ya*, consisting of two separate words each:

- (1) *jon a -tsi -a xu -mu -saala*
 John subject -go -tense(T) locative class -tree
 marker(SM) marker(LM) marker(CM)3
- 'John went on the tree'

(2) jon a -tsi -a {inyuma } ya omu -saala
 John SM -go -T {imbeli } of CM3 -tree
 'John went {behind } the tree'
 {in front of }

2.1. Agreement processes. Evidence (more substantial than that provided by the orthographic conventions of the language) for the non-prepositional status of the locative markers xu- , mu- , and ha- comes from their not being subjected to a general constraint in the language against prepositional stranding. This constraint is illustrated by the Relativization strategies in the language. Relative clauses in OluTsootso appear after the head NP, beginning with a relative pronoun generally agreeing in class with the head NP and in case marking with the target of Relativization. In addition to the relative pronoun, a pronominal copy of the target NP obligatorily appears after the preposition in a relative clause formed on the object of a locative preposition, whereas there is no such copy in a relative clause formed on a NP that is clearly not preceded by a preposition and which cannot be analyzed as an object of one. Example (3) constitutes an example of a relative clause formed on a prepositional phrase type locative:

(3) in -zu e -yi -a jon a -tsi -a inyuma ya yiiyo
 CM9 -house relative -CM9 -RM John SM -go -T behind of class
 clause (nonsubject) (C)9
 marker(RM) demonstrative(D)
 'the tree which John goes behind ...'

Note that without the demonstrative copy yiiyo following the prepositions inyuma ya the relative clause would be ungrammatical. Relativization on NP's that are not objects of prepositions, on the other hand, does not require this additional pronominal copy of the target of Relativization in the relative clause. Examples (4) and (5) illustrate relative clauses formed on subjects and direct objects respectively:

(4) aβa-saatsa a -βa -xol -a emi -limo ...
 CM2-man RM -C2 SM -do -T CM4 -work
 'the men who do work ...'

- (5) aβa -saatsa a -βa -a en -dol -a ...
 CM2 -man RM -C2 objec- -RM I -see -T
 marker(OM) (object)
 'the men whom I see ...'

The non-prepositional phrase type locatives with the markers xu- , mu- , and ha- behave like regular non-prepositional phrase NP's in their choice of a Relativization strategy. Relativization of such locatives require no additional pronominal copy in the relative clause. The explanation is that these locatives do not include any prepositions so that their Relativization does not involve preposition stranding, thus requiring no additional pronominal copy in the relative clause. Relative clauses formed on non-prepositional phrase NP locatives are illustrated in (6) and (7).

- (6) mu -shi -iro o -mw -a jon a -leer -a eshi -taβo ...
 iM -CM7 -market RM -LM -RM John SM -bring -T CM7 -book
 'in the market in-where John brings the books ...'

- (7) xuu -n -zu o -xw -a jon a -tsi -a ...
 LM -CM9 -house RM -LM -RM John SM -go -T
 'on the house on-where John goes ...'

We have shown that syntactic evidence strongly corroborates the morphological evidence indicating that NP locatives (unlike the other prepositional phrase type locatives) are not prepositional phrases. Exactly what the grammatical relation of these NP locatives is will be discussed later on in the paper.

Locative NP's can be shown to be crucially different from prepositional phrase locatives in other respects as well. The nouns of the prepositional phrase locatives trigger class agreement on demonstrative pronominal forms, whereas the NP locatives govern locative marker agreement. Contrast, for example, the demonstrative constructions (8) and (9), which are formed on a prepositional phrase locative and a NP locative respectively:

- (8) inyuma ya omu -saala yukwo ...
 behind of CM3 -tree demonstrative(D)C3
 'behind that tree there ...'
- (9) xu -mu -saala yu -xwo ...
 on -CM3 -tree D -LM
 'on that tree on-there ...'

In (8) the demonstrative *yukwo* agrees with *omu-saala* 'tree' in noun class, while in (9) the demonstrative *yuxwo* agrees with the locative class 17 marker *xu-*. Similar agreement differences hold for relative pronouns formed on prepositional phrase locatives and NP locatives. The relative pronoun *e-yi-a* in (3), for example, being formed on a prepositional phrase locative, agrees in class with the noun *in-zu* 'house', while the relative pronouns *o-mw-a* in (6) and *o-xw-a* in (7), being formed on NP locatives, agree with the locative classes 17 and 18 whose markers are *mu-* and *xu-* respectively.

It is interesting to note, however, that this difference in coding procedure triggered by the two types of locatives in demonstratives and relative pronouns does not extend to possessives. Possessive pronominal forms for both types of locatives carry class agreement, as examples (10) and (11) indicate.

(10) *inyuma ya tsi -siimba tsi -anje ...*
 behind of CM10 -lion CM10 -my
 'behind my lions ...'

(11) *xu -tsi -siimba tsi -anje /*xw -anje ...*
 on -CM10 -lion CM10 -my on -my
 'on my lions ...'

Returning to the differences between NP locatives and prepositional phrase locatives, another such difference concerns verbal agreement. NP locatives, when in subject position, govern subject agreement, while prepositional phrase locatives never do, (indicating perhaps, that the latter type cannot be promoted to subject position by any syntactic process(es) in the language):

(12) *xuu -n -zu xu -bal -a*
 LM -CM9 -house locative -be warm -T
 (L) SM
 'it was warm on the house'

(13) **inyuma ya inzu yi -bal -a*
 behind of CM9-house CM0 -be warm -T
 ('it was warm behind the house')

Though the locatives in both sentences appear in subject position, the verb *xu-bal-a* in (12) agrees with the locative marker *xuu-*, whereas (13) is

ungrammatical even though the verb *yi-bal-a* agrees with the prepositions *inyuma ya* and their object *in-zu*, both being of class 9.¹

2.2. Syntactic processes. Having established that NP locatives, marked by the prefixes *xu-*, *mu-*, and *ha-*, are distinct from prepositional phrase locatives with respect to the coding properties they trigger in relative pronouns, demonstratives, and on verbs, and with respect to the relative clause formation strategies they employ, we go on to show that their behavior with respect to accessibility to syntactic processes such as Passivization, Tough Movement, Clitic Pronominalization, and Topicalization is also different: locative NP's are accessible to all these processes, while prepositional phrase locatives are not.

Passivization in OluTsootso generally applies to structures such as (14) to produce structures such as (15), thus promoting a direct object to subject status and demoting the underlying subject to *chômeur* status:

(14) *jon a -leer -a eshi -taβo xulua mary*
 John SM -bring -T CM7 -book for Mary
 'John brought the book for Mary'

(15) *eshi -taβo shi -leer -w -a xulua mary neende jon*
 CM7 -book CLSM -bring -passive -T for Mary by John
 marker(PM)
 'the book was brought for Mary by John'

Note that in (15) *eshi-taβo* the derived subject triggers class agreement on the verb. This rule of Passivization can also apply to sentences like (1) to produce passive structures like (16), where the NP locative has been promoted to subject status, triggering locative agreement on the verb:

(16) *xu -mu -saala xu -tsii -bw -a -xwo neende jon*
 LM -CM3 -tree LSM -go -PM -T -locative by John
 clitic(LC)
 'on the tree was gone by John'

¹Though not a noun, *inyuma ya* resembles a class 9 word *inyuma* followed by the class 9 associative marker *y-a*, so that if it would at all trigger subject agreement, it would be in the form of the prefix *yi-*.

Though NP locatives can undergo Passivization, and thus be promoted to subject status regardless of whether or not the verb in the sentence belongs to the class of transitive verbs "traditionally" considered to govern Passivization, prepositional phrase locatives cannot under any circumstances be passivized into subject position. The application of Passivization to a structure like (2), for example, yields an ungrammatical output (17):²

- (17) *inyuma ya omu -saala yi -tsii -bw -a -ho neende jon
 behind of CM3 -tree C9SM -go -PH -T -LC by John
 ('behind the tree is gone by John')

In the same manner locative NP's can be shown to undergo Tough Movement, whereas prepositional phrase locatives cannot. Tough Movement in OluTsootso is generally restricted in domain to direct objects only, the adjectives *angu* 'light, easy' and *tinyu* 'hard' preceded by the copula *ni* serving as Tough Movement triggers. The presumed underlying structure is attested on the surface in sentences such as (18):

- (18) oxu -chama jon ni oxw -aangu
 CM15³ -please John is CM15 -easy
 'to please John is easy'

The Tough Moved version of (18) is (19); Tough Movement having raised the object of the embedded clause to matrix subject position:

- (19) jon ni omw -aangu w -oxu -chama
 John is CM1 -easy CM1 -CM15 -please
 'John is easy to please'

Besides direct objects, NP locatives with the markers *xu-*, *mu-*, and *ha-* may undergo Tough Movement. Thus, for example, Tough Movement applies to (20) to yield (21):

- (20) oxu -leera eshi -taβo mu -shi -iro ni oxw -aangu
 CM15 -bring CM7 -book LM -CM7 -market is CM15 -easy
 'to bring a book in the market is easy'

²Note that earlier we cited (13) to demonstrate that prepositional phrase locatives cannot control subject agreement in a sentence; (17), then, would simply serve as another such example.

³CM15 is the infinitive marker.

- (21) mu -shi -iro ni omw -aangu mw -oxu -leera -mwo eshi -taβo
 LM -CM7 -market is LM -easy LM -CML5 -bring -LC CM7 -book
 'in the market is easy to bring a book'

It is not possible, however, to Tough Move prepositional phrase locatives, since this process would yield outputs with the prepositional phrase locative in matrix subject position. We have already noted that prepositional phrases cannot serve as subjects of sentences, so that one must exclude prepositional phrase locatives from the domain of the rules of Passivization as well as Tough Movement, possibly in terms of a general constraint on the language, restricting non-term prepositional phrases from assuming subject position. Given such a constraint, a sentence like (22), wherein a prepositional phrase locative has undergone Tough Movement to become subject, is ungrammatical:

- (22) *inyuma ya in -zu ni yi -angu yi -oxu -leera eshi -taβo
 behind of CM9 -house is CM9 -easy CM9 -CML5 -bring-LC CM7 -book
 ('behind the house is easy to bring a book')

The third syntactic process to which NP locatives are accessible while the prepositional phrase type are not is a Pronominalization rule that deletes a NP, copying it in terms of a clitic pronoun attached onto the verb. This type of Clitic Pronominalization applies to direct objects such as eshi-taβo in (14) to produce a sentence like (23):

- (23) jon a -shi -leer -a xulua mary
 John SM -CM7 -bring -T for Mary
 'John brings it for Mary'

The direct object, eshi-taβo, is pronominalized into an object pronoun, -shi-, in class agreement with its antecedent; -shi- then appears as a prefix on the verb. A locative NP such as mu-shi-iro 'in the market' in a sentence like (24) would be pronominalized into an agreeing locative pronoun mwo which like the direct object pronoun is cliticized onto the verb, (but unlike the case of the direct object, is cliticized as a suffix) to yield a sentence like (25):

- (24) jon a -leer -a eshi -taβo mu -shi iro
 John SM -bring -T CM7 -book LM -CM7 market
 'John brings the book in the market'

- (25) jon a -leer -a -mwo eshi -taβo
 John SM -bring -T -LC CM7 -book
 'John brings the book (in) there'

Prepositional phrase locatives could not be pronominalized in the same way as NP locatives without violating the general constraint against preposition stranding. If the object of the preposition alone were pronominalized and cliticized onto the verb, then the preposition would be stranded, resulting in an ungrammatical sentence. The deletion of the preposition, however, does not render the sentence any more grammatical than before. Thus, Pronominalization of the prepositional phrase locative in (2), for example, in the form of a clitic on the verb, both with or without the deletion of the stranded preposition, generates the ill-formed constructions in (26):

- (26) *jon a -tsi -a -yiiyo (inyuma ya)
 John SM -go -T -CM9 LC (behind of)
 ('John goes behind (it)')

The only way to pronominalize the prepositional phrase locative *inyuma ya omu-saala* in (2) is (27), which is not a clitic-type Pronominalization:

- (27) jon a -tsi -a inyuma ya yu -kwo
 John SM -go -T behind of D -C3
 'John goes behind it'

Another syntactic process to which NP locatives, contrary to prepositional phrase locatives, are accessible is Topicalization. Topicalization in OluTsootso applies on a post-verbal NP, moving it to the left of the sentence, inserting after it a demonstrative pronominal element in agreement with it, and attaching as a prefix to the verb an object marker in agreement with the Topicalized NP.⁴ To illustrate, in (28) the direct object *aβa-ana* 'child' and in (29) the NP locative *mu-shi-iro* 'in the market' are Topicalized:

⁴Due to the constraint against preposition stranding only non-prepositional phrase NP's are eligible for Topicalization; thus chômeurs resulting from the equivalent of the English Dative Movement (ex-direct objects) can undergo Topicalization, while those resulting from Passivization (ex-subjects), being preceded by the preposition *neende*, cannot.

- (28) aβa -ana yaaβo em -ba -lol -a
 CM2 -child C2D I -them -see -T
 'those children ... those children, I see them'
- (29) mu -shi -iro yumwo jon a -mu -leer -a -mwo eshi taβo
 LM -CM7 -market LD John SM -L object -bring -T -LC CM7 book
 'in the market, John brings the book (in) there'

Sentence (29) indicates that Topicalization bestows on the topicalized NP some object-like properties--at least insofar as verbal marking is concerned. Such phenomena tempt one to speculate that Topicalization in this language perhaps involves a promotional step, such that the Topicalized NP is promoted to direct object position, at least with respect to coding properties, and possibly with respect to some behavioral properties as well.⁵ This is why in (29) the topicalized NP locative mu-shi-iro 'in the market' triggers both an object marker prefix and a locative clitic as a suffix on the verb.⁶ In any case, a prepositional phrase locative, unlike the NP locative, cannot undergo Topicalization, as (30) indicates:

- (30) *inyuma ya omu -saala yukwo jon a -ku- -leer -a -yukwo eshi -taβo
 behind of CM3 -tree C3D John SM -C3OM -bring -T -C3D CM7 -book
 ('behind the tree, John brings the book there')

3. NP Locatives and Direct Objects

3.1. Shared properties of NP locatives and direct objects. We have thus far shown that in coding as well as behavioral properties NP locatives stand distinctly apart from prepositional phrase locatives. Furthermore, it has become apparent from their behavioral properties in terms of accessibility to certain syntactic rules (whether the rules directly affect grammatical relations or whether they are sensitive to them in some way) that NP locatives share many behavioral characteristics with

⁵The justification for analyzing Topicalization as a promotion rule in OluTsootso, though a very interesting issue, will not be taken up in this paper, due to its rather marginal relevance to the topic as weighed against its complexity.

⁶We will show later on in the paper that the appearance of a locative clitic as a suffix on a verb indicates that the NP locative in the sentence has undergone a change in grammatical relation.

direct objects--properties not shared by other grammatical relations in the language.⁷ It can be shown, for example, that besides the NP locatives in question, only direct objects (both derived and basic) can undergo the relation-changing rules of Passivization and Tough Movement to be promoted to subject position.⁸

It has been shown in examples (15) and (16) that Passivization can apply to direct objects and NP locatives respectively. It must be noted that in the case of NP locatives the verb governing Passivization is a directional intransitive verb, rather than a transitive verb, which usually governs this rule. A number of directional intransitive verbs in OluTsootso govern Passivization, with some exceptions, among them -its 'come', for example, whose passive version sounds "funny" to native speakers.⁹ In order to substantiate the claim that only direct objects and NP locatives undergo Passivization, one has to show that Passivization of the indirect object *mary* in (14), for example, yields the ungrammatical constructions in (31):

- (31) **mary* -a -leer -w -a eshi -taβo (xulua)¹⁰ neende jon
Mary -SM -bring -PM -T CM7 -book (for) by John
('Mary is brought the book by John')

However, if the indirect object *mary* in (14) is promoted to direct object position prior to Passivization, then Passivization yields a grammatical sentence. Though we do not provide irrefutable evidence for the

⁷We are assuming Postal and Perlmutter's [unpublished lectures] classification of grammatical relations to be correct, so that the possible universal grammatical relations are subject, direct object, and indirect object, with all other NP's being non-terms--either oblique ones which were never terms at any stage of the derivation of the sentence, or chômeurs which held a grammatical relation to the verb only at some earlier stage of the derivation of the sentence.

⁸In OluTsootso there are some verbs that take double objects, one or the other of which is generally more accessible to certain processes. The exact basis on which the grammatical relations of the two objects of a given verb are determined is not of any immediate concern in this paper.

⁹See Dalgish [1976] for a more complete list of directional intransitive verbs governing Passivization in this language.

¹⁰Notice that the sentence is still ungrammatical if the preposition *xulua* meaning 'for' is deleted, so that the ungrammaticality of (31) is not due to a violation of the constraint against preposition stranding.

derivation of sentences like (32) from those like (14), there is reason to believe that there exists a productive process in OluTsootso, somewhat equivalent to Dative Movement, whereby an "applied" suffix, (-il/-el) is attached onto the verb making it benefactive, instrumental, or directional, the case marker denoting benefactive, instrumental or directional is deleted, and the NP whose marker has been deleted is moved to direct object position.¹¹ The Dative Movement rule would apply to (14), for example, to produce (32):¹²

- (32) jon a -leer -el -a mary eshi -taβo
 John SM -bring -applied -T Mary CM7 -book
 marker(AM)
 'John brings Mary the book'

Mary in (32), being a direct object, is accessible to Passivization, as shown in (33), while eshi-taβo the ex-direct object which has been demoted to chômeur status is inaccessible to the rule. This turns out to be indeed the case, as indicated by the ungrammatical (34) where eshi-taβo has been Passivized:

- (33) mary a -leer -el -w -a -eshi -taβo neende jon
 Mary SM -bring -AM -PM -T -CM7 -book by John
 'Mary is brought the book by John'

¹¹This type of proposal is made by Kimenyi [1976] and Givon [1976] for the promotion of indirect objects ("datives" in Givon's terminology) and of instrumentals in the Bantu languages of KiNyarwanda and Bemba, which are of course related to OluTsootso.

¹²It is neither within the scope nor within the goals of this research to show whether the structure in (32) is a derived or underlying one--that is to say whether or not there exists a rule of Dative Movement in the language. The arguments for and against either position are not decisive and clear-cut. In any case, it should be made clear that the absence of a rule of Dative Movement does not affect our argument in any crucial way. Without the Dative Movement analysis, (31) would still be ungrammatical because mary an indirect object has been passivized. In (33), on the other hand, mary would be considered as a basic direct object, undergoing Passivization to yield a grammatical output. However, we strongly suspect that the analysis which postulates a rule of Dative Movement is more motivated; so for the purposes of this paper we shall assume such a rule to exist.

- (34) *eshi -taβo shi -leer -el -w -a mary neende jon¹³
 CM7 -book CM7 -bring -AM -PM -T Mary by John
 ('the book is brought Mary by John')

Having established one behavioral characteristic, involving accessibility to Passivization, shared by only the NP locatives and the direct objects in OluTsootso, the next step is to bring out other such properties exclusively shared by the two relations. Tough Movement is another process that demonstrates such behavior. We have already shown in an earlier part of this paper that direct objects and NP locatives can undergo Tough Movement, whereas prepositional phrase locatives cannot; see sentences (19), (21), and (22) respectively. In order to establish that only direct objects and NP locatives can be promoted to become the subject of a higher clause, it is sufficient to show that indirect objects cannot undergo Tough Movement.¹⁴ Consider sentence (35):

- (35) oxu -leera eshi -taβo xulua mary ni oxw -aangu
 CM15 -bring CM7 -book for Mary is CM15 -easy
 'to bring a book for Mary is easy'

The Tough Moved version of (35) is the ungrammatical (36):

- (36) *mary ni omw -aangu w -oxu -leera (xulua)¹⁵ eshi -taβo
 Mary is CM1 -easy CM1 -CM15 -bring (for CM7 -book
 ('Mary is easy to bring a book for')

Example (37), however, a paraphrase of (35), where mary is the direct object of a benefactive "applied" form of the verb leer 'bring', can undergo Tough Movement to produce a grammatical (38):

- (37) oxu -leer -ela mary eshi -taβo ni oxw -aangu
 CM15 -bring -AM Mary CM7 -book is CM15 -easy
 'to bring Mary a book is easy'

¹³Please note that whereas such sentences are ill-formed in OluTsootso and in ChiMwini, they are grammatical in other Bantu languages such as Swahili and KiMeru, for example.

¹⁴Indirect objects are nevertheless objects of prepositions. In order to show that only direct objects can undergo Tough Movement, one should test the objects of double object verbs. The behavior of such double object constructions is complex, and though if properly analyzed, it could shed some light on the issue, such an analysis has not yet been conducted by us.

¹⁵See footnote 10.

- (38) mary ni omw -aangu w -oxu -leer -ela eshi -taβo
 Mary is CMI -easy CMI -CM15 -bring -AM CM7 -book
 'Mary is easy to bring a book to/for'

Besides the relation-changing rules of Passivization and Tough Movement, there are syntactic processes which, though not relation-changing, are sensitive to grammatical relations, affecting in similar ways direct objects and NP locatives only. Two such syntactic processes are Pronominalization and Relativization.

Clitic Pronominalization has been shown to apply to direct objects and NP locatives in sentences (23) and (25) respectively, but not to prepositional phrase locatives, as illustrated by sentence (26). It remains to be shown that Clitic Pronominalization does not apply to indirect objects either, so as to establish that the behavioral properties of NP locatives and direct objects in Clitic Pronominalization are not only similar, but also that they are not shared by other grammatical relations and non-terms in the language.¹⁶ Clitic Pronominalization, if applied to the indirect object *mary* in (14), would yield the ungrammatical (39):

- (39) *jon a -mu -leer -a eshi -taβo (xulua)¹⁷
 John SM -DM -bring -T CM7 -book (for)
 ('John brought for her a book')

We can show that even if the preposition is not stranded (as it is in (39)), the indirect object is still not acceptable to Clitic Pronominalization. By adding the demonstrative *wuuwo* 'that one (person)' after *xulua*, we have prevented the stranding of the preposition; but the sentence, with the indirect object cliticized onto the verb, is still ungrammatical:

- (40) *jon a -mu -leer -a eshi -taβo xulua wuuwo
 John SM -DM -bring -T CM7 -book for Cl-D
 ('John brought her a book for her')

¹⁶It is unnecessary to check Clitic Pronominalization of subjects for its similarities to and differences from Clitic Pronominalization of direct objects and NP locatives, because the appearance of a subject clitic pronoun (an agreement marker) on the verb is obligatory in all Olu-Tsootso non-imperative sentences, so that the Pronominalization of a subject would simply involve its deletion.

¹⁷See footnote 10.

The argument involving Relativization follows along the lines of the Clitic Pronominalization argument. Having shown two different Relativization strategies, one for prepositional phrases and the other for NP's, due to the constraint against preposition stranding, one expects indirect objects to be relativized by the strategy used for prepositional phrases, since the indirect object case markers appear in front of the indirect objects in the form of a separate word.¹⁸ Indeed, unlike the direct objects and the NP locatives, the indirect objects require a pronominal copy of the target NP when relativized, so that the relative clause would be ungrammatical without it:¹⁹

- (41) *aβa -saatsa a -ʔa -a en- -deer- -a eshi -taβo (xulua) ...
 CM2 -man RM -C2M -RM I -bring -T CM7 -book (for)
 ('the men for whom I brought the books ...')

We have seen, therefore, that direct objects and locative NP's use a different Relativizing strategy than that used by indirect objects and other prepositional phrases. In order to completely distinguish direct objects and NP locatives as one set, different from all other grammatical relations with respect to Relativization, the Relativization strategy used for subjects remains to be examined. Relativization of subjects involves a strategy slightly different from that which is used for direct objects and NP locatives, the difference being that in the former, the relative pronoun has only one relative marker affix on it, as example (42) indicates:

¹⁸Charles Kisseberth pointed out to us that datives like xulua NP, being prepositional phrases, are expected to behave like all other prepositional phrases in the language. In order to distinguish direct objects from other objects of the language, one's attention has to be focussed primarily on unmarked NP's. The lack of distinction between marked and unmarked objects, therefore, reflects one of the limitations of Postal and Perlmutter's theory of Relational Grammar.

¹⁹The grammatical version for (41) would be:

- (i) aβa -saatsa a -βa -a en -deer -a eshi -taβo xulua ya -aβo
 CM2 -men RM -CM2 -RM I -bring-T CM7 -book for D -CM2
 'the men for whom I brought the books . . .'

where a pronoun, demonstrative copy yaaβo 'they (class2)' appears after the preposition.

- (42) aβa -saatsa a -βa -xol -a emi -limo ...
 CM2 -man RM -C2SM -do -T CM4 -work
 'the men who worked ...'

Thus locatives and direct objects are set apart from other NP's with respect to the rules of Relativization and Clitic Pronominalization, which are sensitive to grammatical relations, as well as with respect to relation-changing rules such as Passivization and Tough Movement, making imperative an analysis which accounts for the similarities between the two. Any analysis that sets apart NP locatives from direct objects, as different grammatical relations (or categories totally distinct from each other), misses the generalization that the NP locatives undergo the same relation-changing processes (Passivization and Tough Movement) that underlying and derived direct objects do. Such an analysis would also miss the generalization that the two relations/categories employ the same strategies for (each of) the Relativization and Pronominalization processes, as opposed to subjects, on the one hand, and as opposed to other grammatical relations like indirect objects, etc. on the other.

The commonly accepted analysis, e.g. Trithart [1975], that such NP locatives, typically found in Bantu languages, are prepositional phrases not only misses the above generalizations, but also violates the Universal Subjectivization Constraint [Johnson 1974] and questions its universality. The Universal Subjectivization Constraint states:

- (43) If a certain position on the Relational Hierarchy²⁰ subject > direct object > indirect object > non-term undergoes a subjectivization rule, then all non-subject positions above it on the Relational Hierarchy must be able to undergo that rule.

An analysis wherein the NP locatives are considered to be prepositional phrases would necessarily violate (43). As prepositional phrases, such

²⁰The Relational Hierarchy, which roughly correlates with Keenan and Comrie's [1977] NP Accessibility Hierarchy, ranks the relations in such a way (with subjects in the highest position), as to enable the formulation of linguistic universals using the hierarchy, the constraint in (43) constituting such an example.

NP locatives would be included under non-terms, forcing an analysis that allows for some non-terms and direct objects only to subjectivize via Passivization and Tough Movement, thus leaving a gap of non-subjectivizable NP's on the Relational Hierarchy over the range of indirect objects--indirect objects not being directly accessible to these rules. Such an analysis would either have to reject altogether the Universal Subjectivization Constraint as invalid, or resort to some ad hoc revision of it, e.g. Trithart [1975].²¹

3.2. Differences between NP locatives and direct objects. Evidence has been provided in terms of behavior with respect to relation-changing rules, and from rules which are sensitive to grammatical relations, to suggest that NP locatives have the same grammatical relation to the verb as direct objects do. On the other hand, there is also evidence suggesting that NP locatives and direct objects are distinct from each other.

One such piece of evidence is the fact that NP locatives and direct objects trigger different coding processes. First of all, a relative pronoun formed on a direct object agrees with the class of the head NP, whereas a relative pronoun formed on a NP locative agrees with the locative marker of the head NP. Contrast example (5) with (6) and (7). The same discrepancy in agreement behavior can be found in verbal agreement triggered by a direct object and that triggered by a NP locative, once they are subjectivized as in passive sentences. A subjectivized direct object triggers class agreement on the verb whereas a subjectivized NP locative triggers verbal agreement with the locative marker.²² Contrast

²¹Trithart's Revised Subjectivization Constraint is circular:

- (i) If a language can subjectivize an NP low in the Relational Hierarchy, then it can subjectivize NP's in all intermediate positions, where subjectivizable is defined recursively as follows:
- a) X is subjectivizable if L has a rule $X \rightarrow \text{Subject}$
 - b) X is subjectivizable if L has a rule $\rightarrow Y$, where Y is subjectivizable.

²²Note that it is possible to analyze the locative marker as a class prefix, so that the coding processes for NP locatives and direct objects would appear to be similar. However, it can be shown that the locative NP has a class prefix of its own, other than the locative marker, which

example (15) with (16). Another difference between the two is that Clitic Pronominalization involves the appearance of a verbal agreement prefix for direct objects, whereas the verbal agreement for NP locatives is in the form of a suffix, as shown in (23) and (25) respectively.

There exists another major coding difference between direct objects and NP locatives. A NP locative leaves a locative clitic on its verb, if it undergoes a change in grammatical relation, whereas a direct object leaves no clitic on its verb when undergoing a change in grammatical relation. Contrast the subjectivized direct object and NP locative in the passive sentences (15) and (16) respectively. The verb in (15), *shi-leer-w-a*, is made up of the subject marker, followed by the verb stem, the passive suffix, and the tense marker. The verb in (16), *xu-tsii-bw-a-xwo*, has one suffix in addition to the elements found in the verb in (15), namely *-xwo*, a locative clitic indicating the underlying grammatical relation of the subject. The same phenomenon is true of NP locatives when subjectivized via Tough Movement, as sentence (21) shows. The embedded verb in (21) carries a locative clitic *-mwo* indicating the underlying grammatical relation of the matrix subject with the embedded verb.

Besides the above differences in the coding properties they trigger, there is another not less substantial argument for considering direct objects and NP locatives as distinct relations, despite their main similarities. There is sufficient motivation for positing a rule of Locative Marker Deletion in OluTsootso, whereby a NP locative is transformed into a direct object. The immediate output of this rule is not attested in the language, so that such an output serves as an intermediate structure for rules such as Pronominalization, Relativization, Topicalization, Passivization, and Tough Movement, which treat the NP locatives whose markers have been "stripped off" as if they were direct objects.

must under the circumstances, be interpreted as a case marker. There are sentences in which the locative NP can be separated from its locative prefix, and then undergo the rules discussed above. When this happens, all agreement is with the NP, and not with the locative marker. See Dalgish [1976b] for some examples. Also, see the discussion, following in the text.

Clitic Pronominalization of *mu-shi-iro* in (24), results in (25) and a syntactic variant of it (44), where the clitic pronoun is in form of an object pronoun reflecting class agreement with *shi-iro*, thus indicating that *shi-iro* has become the object, probably at a stage in the derivation prior to Clitic Pronominalization. In addition, there is a locative clitic *-mwo* on the verb, indicating that the underlying NP locative has undergone some change in grammatical relation.

- (44) *jon a -shi -leer -a-mwo eβi -taβo*²³
 John SM -C7OM -bring -T -LC CM8 -books
 'John brings the books in it'

Likewise, relativized NP locatives as in (6) and (7) have syntactic variants wherein the locative markers have been deleted from the NP locatives, so that the relative pronoun includes an object marker agreeing in class with the "stripped" NP, and the verb has a locative clitic attached to it, indicating a change in the grammatical relation borne by the target of Relativization. The syntactic variants of (6) and (7) are (45) and (46) respectively:

- (45) *eshi -iro e -shy -a jon a -leer -a -mwo eshi -taβo ...*
 CM7 -market RM -C7M -RM John SM -bring -T -LC CM7 -book
 (object)

'the market which John brings a book in ...'

- (46) *in -zu e -yi -a jon a -tsi -a -xwo ...*
 CM9 -house RM -C9M -RM John SM -go -T -LC
 (object)

'the house which John goes on ...'

Topicalized locative NP's as in (29) were hypothesized earlier in the paper to have undergone objectivization, even without being "stripped" of their locative markers, so that they trigger object agreement and leave a locative clitic on the verb, indicating that they have undergone a change in grammatical relation. Sentences like (29) have their syntactic

²³The object, *eβi-taβo* 'books' is substituted for *eshi-taβo* 'book' simply to eliminate the possibility that the C7OM would be "agreeing" with 'book', and not *-shi-iro* (class 7) 'market'. If the OM in (44) were agreeing with 'books', the C8OM *-βi* would appear on the verb, instead of the *-shi* C7OM, which does occur.

variants in sentences such as (47), where the Topicalized NP locative, "stripped off" of its locative marker, is clearly a direct object, triggering object agreement and leaving a locative clitic on the verb indicating that the locative NP has undergone a change in grammatical relation:

- (47) eshi -iro yiisho jon a -shi -leer -a -mwo eþi -taþo
 CM7 -market C7D John SM-C7OM -bring -T -LC CM8 -books
 'that market John brings the books in it'

Passive and Tough Moved sentences with NP locatives as subjects also have syntactic variants in which the subjectivized NP's are "stripped off" of their locative markers, the verbs agreeing with them in class, as with subjectivized direct objects. Thus, for example, the Passive (16) has a variant in (48), while the Tough Moved (21) has its variant in (49):

- (48) omu -saala ku -tsii -bw -a -xwo neende jon
 CM3 -tree C3SM -go -PM -T -LC by John
 'the tree was gone on by John'
- (49) eshi -iro ni eshi -aangu shi -oxu -leer -a -mwo eshi -taþo
 CM7 -market is CM7 -easy CM7 -CM15 -bring -T -LC CM7 -book
 'the market is easy to bring a book in'

It should be clear, then, that NP locatives are different from direct objects, not only because they have different coding properties from those of direct objects, but also because NP locatives can be transformed into direct objects by "stripping off" their locative markers, as the above data has shown.²⁴

3.3. NP locatives and direct objects as subrelations. An analysis which provides the basis for capturing the generalization that direct objects and NP locatives share certain behavioral properties and which at the same time allows for differences in coding properties, can only be one

²⁴Note that prepositional phrase locatives, unlike NP locatives, cannot be "stripped off" of their prepositions due to the general constraint against preposition stranding.

wherein NP locatives and direct objects constitute distinct subrelations within a single but broader grammatical relation which we shall call "supra direct objects" simply for lack of a better term.²⁵

There are other works, e.g. Sheintuch [1976], which show the need for a finer subclassification of grammatical relations for certain languages. Such language-specific sub-grammatical relations are generally semantically specifiable, basic (as opposed to derived) grammatical relations. Further investigation of grammatical relations in various languages should clarify whether such language-specific subrelations are motivated and governed by universal principles, and if so, an attempt should be made towards the discovery of such principles.

²⁵There seems to be one type of counterexample, indicating that NP locatives can also be underlying subjects, as shown by the lack of appearance on the verb of the locative clitic which designates that the NP locative has undergone a change in grammatical relation. See, for example, (i):

- (i) xuu -n -zu xu -bal -a
 LM -CM9-house LSM -be warm -T
 'on the house was warm'

Though we admit that we have not looked into this matter carefully, we offer a suggestion of what may be happening here. Based on the observation that the equivalent of such sentences in English and many other languages are generally either subjectless or have a dummy element, e.g. it in English, for subject, as (ii) demonstrates:

- (ii) It was warm in the house.

we speculate that sentences such as (i) are underlyingly subjectless, and that the NP locatives have been promoted to subject position via a once productive syntactic process, triggering a locative clitic on the verb. Historically, however, this process might have lost its productivity so causing the gradual loss of the locative clitic on the verb. In any case, more research must be devoted to this matter before any conclusions can be drawn.

REFERENCES

- Dalgish, G.M. 1976a. "Passivizing locatives in Olutsootso." In Charles W. Kisseberth, et al. (eds.), *Studies in the Linguistic Sciences*, 6(1):57-67. Urbana, Illinois: University of Illinois Department of Linguistics.
- Dalgish, G.M. 1976b. "Locative NP's, locative suffixes, and grammatical relations." In Henry Thompson, et al. (eds.), *Proceedings of the Second Meeting of the Berkeley Linguistic Society*, pp. 139-48. Berkeley: Berkeley Linguistics Society.
- Givón, T. 1976. "Promotion, accessibility, and case marking: toward understanding grammars." Ms., University of California, Los Angeles.
- Johnson, D. 1974. "Toward a theory of relationally based grammar." Doctoral dissertation, University of Illinois, Urbana, Illinois.
- Keenan, E. 1975. "Towards a universal definition of 'subject'." In Charles N. Li (ed.), *Subject and Topic*, pp. 303-33. New York: Academic Press.
- Keenan, E., and B. Comrie. 1977. "Noun phrase accessibility and universal grammar." *Linguistic Inquiry*, 8:63-99.
- Kimenyi, A. 1976. "A relational grammar of Kinyarwanda." Doctoral dissertation, University of California, Los Angeles.
- Sheintuch, G. 1976. "On the gradation of grammatical relations." In Charles W. Kisseberth, et al. (eds.), *Studies in the Linguistic Sciences*, 6(1):186-94. Urbana, Illinois: University of Illinois Department of Linguistics.
- Trithart, L. 1975. "Relational grammar and Chichewa subjectivization rules." In Robin E. Grossman, et al. (eds.), *Papers from the Eleventh Regional Meeting, Chicago Linguistics Society*, pp. 615-25. Chicago: University of Chicago Department of Linguistics.

SOUTH CENTRAL NIGER-CONGO: A RECLASSIFICATION

Patrick R. Bennett and Jan P. Sterk
University of Wisconsin

A major revision is proposed of the genetic classification of the languages grouped by J. Greenberg under Niger-Kordofanian. Special attention is given to the groups traditionally called Kwa and Benue-Congo, which are now mostly reclassified under the single heading South Central Niger-Congo. Traditional comparative methods and lexicostatistical data based on a computer-aided study were used. Much importance is given to lexical innovations and clusters of innovative isoglosses marking significant subgroupings. The greatest departures from previous classifications lie in the Central Niger group, where Idoma has been included, and among the Bantoid languages, now grouped under the heading Benue-Zambesi, where Guthrian Bantu does not appear to constitute a valid grouping.

1. Content and Methodology

This is a revised version of a paper presented to the 8th Conference on African Linguistics under the title "Benue-Kwa: Internal and External Relations".¹ Its purpose is to present the results of a preliminary re-evaluation of a part of the genetic classification of the languages of Africa proposed by Westermann [1927] and expanded and modified by Greenberg [1966]. We focus on the languages classified by Greenberg as Kwa and Benue-Congo, profiting from ongoing research in the field, including comparative studies of individual subgroups, and from a larger corpus of language data than has been previously available, including unpublished materials. We have been able to reach some specific conclusions as to the position of these languages within the larger framework of Niger-Congo (perhaps Niger-Kordofanian) and as to their interrelationships.

¹We regret the inconvenience caused by the change of title. However, we feel that the term "Benue-Kwa" still incorporates the traditional division between "Kwa" and "Benue-Congo" which we suggest should be abandoned. Hence: South Central Niger-Congo, or SCNC.

It should be pointed out that the broad outlines of preceding classifications are generally confirmed; the areas of controversy examined here are at the level of subgrouping. Any new insights gained in subgrouping are due not so much to improvement in comparative technique as to the more reliable data which permit refinement of cognate recognition. However, the technique of what Greenberg has called "mass comparison", while often adequate for demonstrating relationship, is not best suited for investigation of degrees of relationship and subgrouping.

It is not the intention of the present paper to dwell on the methods used to arrive at our conclusions. A full account of the procedures followed and their theoretical justification is being prepared for publication elsewhere. However, some account of the methodological background is necessary here.

The study of South Central Niger-Congo (SCNC) falls into two main subdivisions, one focussing on external, the other on group-internal relationships. The former rests primarily on a number of lexicostatistical studies, including a short-list general survey of the languages classed by Greenberg as Niger-Kordofanian, a study of body-part vocabulary (some 60 items) in a group of about 45 languages covering Niger-Kordofanian, a large-scale study of lexical relationships within Benue-Congo, and a study covering Niger-Congo generally but focussing on Kwa and Benue-Congo.

The last-mentioned was done specifically for the present investigation, and was a computer-aided weighted-count study of an 87-gloss list involving the fifty languages listed in Table 1 [Tables are in the Appendix at the end of this article]. Multiple entries from individual languages were included, which allowed some compensation for ambiguous or overlapping items; 'arm' and 'hand' were treated as a single gloss even where languages differentiated the two, as were 'bee' and 'honey'. In the latter case, some languages treat 'honey' as primary, with 'bee' equalling, for example, 'fly of honey'. In others, 'bee' is primary, with 'honey' equivalent to 'water (or fat) of bee'. In yet others, the two were distinct. Treatment of 'bee-honey' as a single gloss and admission of multiple entries permitted recognition of cognacy even with semantic skewing, while avoiding double recognition of cognacy as would frequently have occurred had the two been treated as separate glosses. The list of the 87 glosses is given in Table 2.

Cognacy evaluation was on a three-level scale; the first level marked as cognate all items for which a probable common origin could reasonably be assumed. Wherever possible, the necessarily somewhat subjective assessment of similarity was checked against the known or observable patterns of regular correspondence. Thus such recurrent sound shifts² as that of assumed original *t in some environments to r ~ l in Yoruba-Igala, or the shift of assumed original *ku to fu in certain Ijo items, were used to confirm or in some cases deny impressionistic evaluations of cognacy. The second level divided cognate sets into subgroups. Thus in the case of the gloss 'tongue' a reflex of the pattern *me| would be marked at this level as different from items matching a pattern *|em, though the assumed original connection between the two alternants was reflected in their first level evaluation as cognate. The third level was used to distinguish subdivisions of finer detail, as for example, separating nominals with identical stem but non-corresponding class.

The computer program (designed by P. Bennett) analyzes the data in different ways, the two most important being a calculation of the percentage of entries from each language shared with each other (the results given in Table 3) and the determination for each language of the ten languages sharing the most cognates (results given in Table 4).³ Both calculations were done at the first and second level of cognacy. However, for

²It should be remembered that in our present state of knowledge there are only a few groups where it is as yet safe to speak of truly regular correspondences.

³This chart should be read as follows: following the names of the languages are ten columns (numbered from 1 through 10). The figures in the first column indicate the closest languages, those in the second column the second closest languages, etc. Figures are given in pairs separated by a colon: the first figure of the pair indicates the language (numbered from 1 through 50 as in Table 1), the second figure of the pair indicates the number of cognates shared by the two languages. A and B are the two levels of cognacy. For example: the language closest to Grebo (number 1) is number 2 (Newole). The two share 36 cognates at the first level of cognacy (A), and 21 at the second level (B). The language which is second closest to Grebo is Yoruba at the first level of cognacy with 26 items in common, at the second level of cognacy it is Kassene with 7 items shared, etc.

the percentage chart a 18% cut-off level was chosen as a minimum below which any distinctions become meaningless.⁴ Blanks in Table 3 indicate that the degree of cognacy dropped below the cut-off level. Because of this procedure the run-through of this part of the program at the second level of cognacy was only marginally interesting: almost all relations fell below the 18% mark and the print-out was nearly completely blank except within closeknit subgroups. A third level run-through was not therefore attempted.

The percentages in Table 3 are not presented in the usual form of a lower matrix without the diagonal. The reason is that although the number of glosses for each language is the same, the total number of entries may vary since some languages may have more than one entry for one gloss, while other languages may lack entries for particular glosses (see above). Therefore it usually happens that the percentage of items in language A shared with language B may be different from the percentage of items in language B shared with A. The figures in Table 3 should be read horizontally to obtain the figures each language (listed in the lefthand column by its number assigned in Table 1) shares with each other language (listed in the top row, also by number).

For the second calculation, that of the ten closest languages (Table 4) actual numbers of cognates are given, not percentages, and no cut-off point was chosen. This permits a calculation at first (A) and second level (B) of cognacy (third was again omitted). The tabulation of the second level, B, gives evidence for closeknit subgroupings within the clusters indicated at the more generous level A.⁵

⁴The precise figure of 18% was chosen rather arbitrarily. Depending on the languages and data, cut-off at anywhere from 15% to 25% is normally reasonable. The 18%, while a little low, seemed appropriate to the diversity of the languages studied.

⁵For example: both on the percentage chart and even more clearly on the "10 closest" chart (on both levels of cognacy) Idoma patterns consistently as a regular Niger-Kaduna language. The decision to include Idoma as a subgroup with Niger-Kaduna under a common Central Niger node was made primarily on the basis of these data. Compared to Idoma, Yoruba, (looking at the "10 closest" chart) although close to the Central Niger languages at the first or most generous level of

Table 5 shows the geographical location of the fifty languages and the isoglosses for the word for 'head' which is diagnostic for the overall classification.

Lexicostatistics, while useful for preliminary gross subgrouping, is not--if used alone--adequate for indication of fine degrees of relationship. The nature of lexicostatistics is such that geographic and social proximity tends to increase cognacy scores significantly. While the results thus obtained are of course useful as indications of the length of time a particular geographic relationship has existed, or, where present contact does not exist, of past movements, they do not by themselves give adequate evidence on which to base decisions as to genetic relationships. The portion of the study dealing with group-internal relationships, therefore, rests not only on the statistical studies mentioned, but also on investigations aimed at establishing the distribution of shared lexical and phonologic innovations. Where the two types of study disagreed, the innovation-based evidence was usually given priority.

The principal innovation-based study used was also made specifically in connection with this project.⁶ The 145-item word lists were collected for approximately 150 languages. The languages were as far as possible chosen to give an average of four or five members of each of the probable subgroups revealed by the statistical studies. All the languages in this study were chosen from Kwa and Benue-Congo. It was of course not possible to obtain all items for all the languages concerned; items for which there was inadequate representation were not discarded (as had been done in the statistical studies), but also were not given too much weight. For

cognacy, falls far behind at the second level of cognacy. This is a clear indication of the phonological changes and skewings that have distanced Yoruba from the Central Niger group although the two maintain a relatively high percentage of common lexical stock. As an example: the four languages closest to Igbira at the first cognacy level are the other Central Niger languages. Following them is Yoruba (number 16) with a high number of cognates shared (33 in fact). When looking at the second level of cognacy, however, Yoruba does not even appear among the ten closest, evidence that the 33 lexical items that Igbira shares with Yoruba are phonologically quite differentiated.

⁶Other studies were uniformly focussed on individual subgroups, and so are not enumerated here.

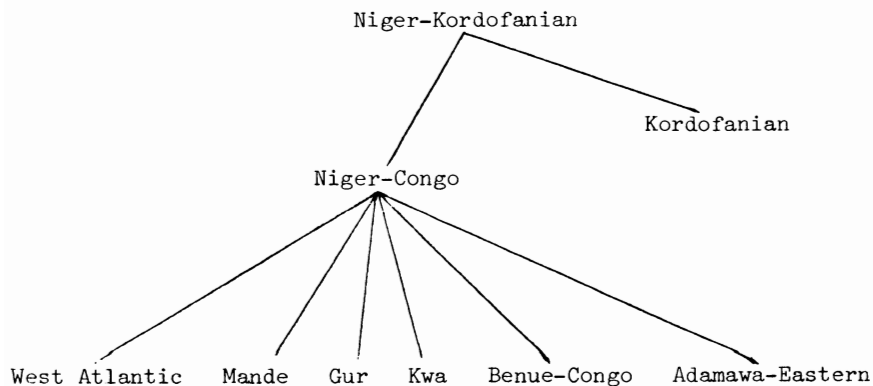
the items whose distribution patterns were least ambiguous and apparently most significant isogloss maps were drawn. In many instances, where it was desired to establish the precise boundary between two forms, Benue-Congo or Kwa languages other than those in the sample were consulted. It was possible to check many items against a 109-word list covering a range of representative Mande, West Atlantic, Gur and Adamawa-Eastern languages. This allowed us to discriminate between inherited Niger-Congo items and local innovations.

For the most part typological criteria were avoided as being often unreliable. In some cases typological boundaries were found to match those based on statistical analysis and innovation distribution; where this was the case the supporting evidence was sometimes noted. However, the majority of typological boundaries are not confirmed by other evidence. It is interesting, however, to note that our data show that the true distribution of features used in previous largely typologically based studies does not show the clear-cut pattern one would expect.

2. SCNC: External Relationships

2.1. Niger-Kordofanian. The majority of the statistical evidence and some of the innovation-based evidence suggests a significant realignment of the Niger-Kordofanian subgroups presently recognized. The Greenbergian groupings are shown in Table 6.

Table 6 - Niger-Kordofanian



Our evidence does not give conclusive support for the relationship of Mande and Kordofanian. Kordofanian shows, besides its typological similarity in overall arrangement (and some details) of concord systems, but few really striking lexical correspondences. Some of those, however, are of a relatively high level of probability, as with the word for 'eye' which may be "reconstructed"⁷ as **-git* in some Kordofanian subgroups and as **-kit* or **-git* in many forms of West Atlantic; or 'tooth', which appears as **-nyi(n)* in much of Kordofanian and most of Niger-Congo. Interestingly, many of the more likely cognates are closest in form to West Atlantic items, though this is the group geographically furthest removed from Kordofanian.

Mande shows less typological similarity. Again, however, there are a few lexical similarities of rather high probability, including the numerals 'two', 'four', 'five', 'neck', 'tooth' and some others. While some items, such as 'goat', can be due to spread of cultural vocabulary across linguistic boundaries, many others are not so liable to borrowing. The Mande group is linguistically much closer knit than Kordofanian, though covering a much wider area.

Mande and Kordofanian show percentages of correspondence with Niger-Congo proper so low that the probability of relationship is very slight. Nonetheless, we have not rejected the hypothesis of a relationship between Niger-Congo, Mande, and Kordofanian, in view of the considerable typological evidence, and the striking similarity in a few basic lexical items. We prefer to reserve further judgment until more detailed examination of a larger body of evidence. In the mean time we suggest setting up Mande, Kordofanian, and Niger-Congo as coordinate branches under the central Niger-Kordofanian node (see Table 9).

⁷Like most of the starred forms in this paper, this is in fact "pseudo-reconstruction", i.e. not a form based on rigorous examination of regular correspondences, but one which is impressionistic and mnemonic. At the level of cognacy involved, and with such limited material, little true reconstruction is now possible.

2.2. Niger-Congo. The evidence for a group containing Greenberg's West Atlantic, Gur, Kwa, Adamawa-Eastern, and Benue-Congo is strong. A large number of lexical and grammatical similarities link these groups, including basic nominal, verbal and numeral vocabulary, precisely corresponding concord systems with cognate class markers, and identical systems of verbal derivation. This group we continue to call Niger-Congo.

Within Niger-Congo proper the evidence of statistics, innovation-sharing, and typology as well indicates two principal divisions. The first corresponds to West Atlantic. This is a very diverse group, containing at least three major subdivisions. It is possible that some language groups traditionally assigned to West Atlantic are in fact coordinate branches of Niger-Congo. There is no apparent common innovation linking West Atlantic, and evaluation of its status as a well-defined subgroup of Niger-Congo must await further detailed investigation. At present, all that can be said is that the lexicostatistical distance between branches of West Atlantic is nearly as great as that between West Atlantic and the remainder of Niger-Congo (see Table 3).

The situation is quite different in the case of the group here called Central Niger-Congo (CNC), which comprises Kwa, Gur, Adamawa-Eastern and Benue-Congo of Greenberg's classification. Here we have a group which can, on the basis of shared innovative lexicon, be considered well-defined.⁸ Statistically as well, the group is comparatively coherent.

Within CNC we recognize principal groups, definable on the basis of statistics and shared innovation. South CNC (SCNC) includes most of Greenberg's Kwa and Benue-Congo, though Kru certainly, and perhaps some other languages must be excluded from this group. SCNC is, as will be shown, a well-defined group; this family and its subgroupings are the major foci of this paper.

North CNC (NCNC) certainly includes the languages classed as Gur and

⁸We consider a grouping of languages "well-defined" if and only if there can be found clear evidence for one or more linguistic innovations, lexical, phonologic, or morphologic, shared by all or most of the languages included in the group but not to be found outside the group.

Adamawa-Eastern, and it is probable that the Kru languages also belong here. The special case of Kru will be discussed in detail below. The well-defined status of NCNC is less certain than that of SCNC. Statistically, there is some evidence for a link between Gur and Adamawa-Eastern. Though the percentages are not large enough to be impressive, the two groups seem to form a continuum. Longuda and Tula, of Adamawa-Eastern, for example, show significantly higher percentages of cognacy with More (Gur) than with some other Adamawa-Eastern languages. The evidence of specific lexical sharings and typology support the link. The word 'two' appears as *sõ in Kru and parts of Adamawa-Eastern. The word for 'head' appears as *du in Kru and Adamawa-Eastern, *yu- in Gur. This is in contrast with the form *to which links SCNC, where the form seems to represent the same item with devoiced initial (see Table 5). In all three groups class markers or their vestiges are consistently suffixed, in contrast to the predominance of prefixing in SCNC, West Atlantic and Kordofanian.

Concerning Kru, two studies [Bennett 1974, Voegler 1974] have pointed out the numerous lexical similarities between Kru and Mande (with which Kru was classified by early researchers) and Gur relative to the comparatively few sharings with SCNC. SCNC sharings are mostly with the contiguous languages of Western SCNC. Mande resemblances, while frequent, are mostly of dubious validity or else in areas of cultural vocabulary. The similarities with certain parts of Gur, especially Lobi, however, are not apparently due to proximity nor confined to cultural items. There is also a certain amount of sharing with Adamawa-Eastern, including such items as *du- 'head', *sõ 'two', and *bidi 'cow' (only in Kru and Longuda); the item *to 'salt' is less significant, as it also appears in Gur and West Atlantic as well as Kru and Adamawa-Eastern.

A more conservative view would be to consider Kru, Gur, and Adamawa-Eastern to be coordinate branches of Central Niger-Congo. Their similarities might well be due to retention of older forms, rather than to any common innovation. It is apparent that Gur and Adamawa-Eastern are very diverse groups; as with West Atlantic, it is possible that some subgroups will prove to be fully coordinate. It is clear that there are striking links between some parts of Adamawa-Eastern and some parts of Gur, between Kru and other parts of Gur, and between Kru and some of Adamawa-Eastern.

It seems that both Gur and Adamawa-Eastern form continua, where extreme languages appear hardly related until geographically and linguistically intermediate groups are examined.

While the data are not conclusive, we hold that Kru is probably one of three branches, with Gur and Adamawa-Eastern, of NCNC. It has influenced and been influenced in turn by its Mande, SCNC, and even West Atlantic neighbors. A future study focussing on NCNC as the present one has focussed on SCNC should clarify its position.

3. SCNC: Internal Relationships

Table 9 contains our proposed classification of South Central Niger-Congo. A representative language of each group is given in most cases, except where, as with Volta-Comoe and Zambesi, the group is large and well known, or where, as in Kru and Imo, the group name is also a language name.

3.1. Main subdivision. The remainder of this study will focus on Greenberg's Kwa and Benue-Congo, and the subgrouping of SCNC. The close relationship of the two Greenbergian groups has long been recognized, and often, as in de Wolf [1971], it has been proposed to set up a higher grouping, "Benue-Kwa", to include the two.⁹ Here we propose a more radical grouping, eliminating the division into "Kwa" and "Benue-Congo", instead recognizing a SCNC none of whose subgroupings can reasonably be interpreted as paralleling the Kwa/Benue-Congo division.

In origin the Kwa/Benue-Congo distinction was typological. Most of the coastal (and hence early-documented) "Kwa" languages have no functioning concord system, or at most have a singular/plural concord distinction; most also have greatly reduced syllable structure with loss of final consonants, and the development of nasalized vowels the norm. Most of the early-documented "Benue-Congo" languages, on the other hand, were known to

⁹The term "Benue-Kwa" has been attributed to Larry Hyman by K. Williamson. De Wolf's argument for the separation between Kwa and Benue-Congo, based on such features as the canonical forms of morpheme structure, is not supported by this study. K. Williamson and B. Elugbe [personal communication] in a paper to appear in the *Greenberg Festschrift* apparently hold a position similar to ours.

have either functioning class systems or well-preserved systems of singular/plural marking. Most also show typically a CVC or even CVCV morpheme structure. If we compare the words for 'arm' in (1), we see good reason for a distinction between 'Kwa' and 'Benue-Congo':

- | | | | | | |
|--------|-----|----------|----|-------------------|-----------|
| (1) a. | úwo | (Ahlõ) | b. | òbóók / ìbóók | (Mbembe) |
| | owó | (Yoruba) | | kùbók / ìbòk | (Afusare) |
| | òvó | (Igbira) | | ukubóko / amabóko | (Rundi) |

The "Kwa" forms in (1a) show a common phonetic reduction when compared with the "Benue-Congo" forms of (1b). If we add, however, the forms of (2), and recognize the presence of numerous "Kwa" languages showing operative class systems, and "Benue-Congo" languages which no longer operate even singular/plural distinctions in nominals, the reason for a Kwa/Benue-Congo distinction becomes less clear.

- | | | | | | |
|--------|--------------|---------|----|-----------|----------|
| (2) a. | gùbwo / ìbwo | (Gade) | b. | avo | (Wukari) |
| | wòbó / àbó | (Yala) | | òḃḃ / àḃḃ | (Ekoi) |
| | óbó / ábó | (Isoko) | | àgwò | (Koto) |

The evidence available today shows no strong typological divisions corresponding to "Kwa" and "Benue-Congo", but rather a continuum of types (to be further discussed under 3.3.1. below). Neither lexicostatistics nor the evidence of shared innovation supports such a distinction. Instead we recognize a three-way division into West SCNC,¹⁰ East SCNC, and Ijo.

The position of Ijo, like that of Kru, is ambiguous. While its affiliations with SCNC seem reasonably certain, obvious cognates are few. Some evidence links it with East SCNC, other data indicate that it should be classed as a coordinate branch, and a position as an independent branch of Central Niger-Congo or even Niger-Congo is not inconceivable. Many of the possible cognates are much changed, thereby distinguishing Ijo sharply from the remainder of SCNC:

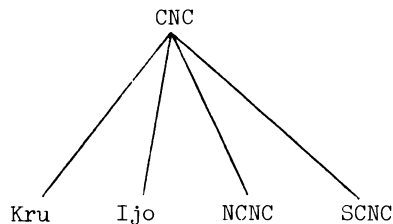
¹⁰Corresponding closely to the languages of Greenberg's Western Kwa, including Volta-Comoe, Gã-Adangme, Ewe-Fon, the Lagoon and Togo-Remnant groups.

- (3) *fɛnɛ (Ijo) / *kónì (SCNC) 'firewood'
 *mɛl (Ijo) / *lɛm (Niger-Congo) 'tongue'
 *tɛbɛ (Ijo) / *toi (SCNC) 'head'

Where cognates are obvious, they are either general Niger-Congo items like *mɛn 'swallow' (*mɛnɛ in Ijo) or probable loans in one or the other direction, as *ovĩ 'cow' in Ijo and parts of Edo and Delta-Cross. As in the case of Kru, there are numerous correspondences with non-SCNC languages. The numeral 'five' is *sorõ in Ijo and much of Mande; 'moon' is *kalo in Ijo and Mande. The metathesis in 'tongue' to *mɛl ~ ɛl in Ijo is common in Kru, Adamawa-Eastern, and parts of West Atlantic. Kru shares with Ijo *bo 'foot' (also possibly found in Adamawa-Eastern), *mu 'go'; 'sleep (verb)' which is *munu in Ijo and *moɔ in Kru,¹¹ and perhaps 'fish' which is *indi in Ijo and *(h)ne in Kru.¹²

We have chosen to treat Ijo as a coordinate branch of SCNC. Statistical evidence and the dearth of clear isoglosses do not let us link it with either the Eastern or the Western subgroup of SCNC. A reasonable alternative (as mentioned earlier) would be to treat it as a branch of CNC coordinate with North and South CNC. If North CNC including Kru should prove not to be a valid grouping as assumed here, it is possible that the best grouping would be as in Table 7:

Table 7 - Central Niger-Congo



¹¹In Niger-Congo generally *ĩa.

¹²For the correspondence compare *mindi (Ijo) / *ni (Kru) 'water'. The connection is remote but not impossible.

3.2. Western SCNC. The distinctive status of the Western SCNC languages is clear. Both the statistical evidence and the evidence of innovation-based isoglosses point to a strong border located between the Western SCNC Ewe-Fon and the Mo languages on the one hand, and the Eastern SCNC Niger-Kaduna and Yoruboid groups on the other. Sample isoglosses are:

- (4) *tã̃ (Western SCNC)¹³ / *tát(o) (Eastern SCNC)¹⁴ 'three'
 (isogloss number 1 (henceforth Is.1) in Table 8)
- *nye ~ nyo (Western SCNC) / *bɛɛĩ (Eastern SCNC) 'breast' (Is.1)
- *je ~ nye (Western SCNC) / *kónì (Eastern SCNC)¹⁵ 'firewood' (Is.1)
- *kɔŋgo (Western SCNC)¹⁶ / *tɔŋ (Eastern SCNC) 'neck' (Is.1a)

Western SCNC includes a number of subdivisions. Up to a point we can follow previous work, as reflected in the Volta-Comoe tree in Trutenau [1976]. We recognize an initial division between the Nyo group and the Togo languages.¹⁷ The division is based on:

- (5) *nyo (Nyo) / *bade (Togo) 'two' (Is.2, Table 8)

The *bade* form is general in SCNC. Other isoglosses of significance are few in our data, owing to internal differentiation. A number of isoglosses, however, separate the two groups without directly supporting common innovation.

Some isoglosses link portions of Nyo with portions of Togo, as

¹³Shared with Kru and some Gur, Adamawa-Eastern languages.

¹⁴Shared with West Atlantic, Kordofanian, Adamawa-Eastern and some Gur. All SCNC forms with final vowel show *tato, while West Atlantic and Adamawa-Eastern forms generally show *tati. If this is a reduplication of *tã̃, we may here have several independent but convergent innovations.

¹⁵But note Adjukru líkù 'tree', a probable cognate.

¹⁶Also in Ijo, Mande, and (as kíngò) in Zambesi and Equatorial. Here the Eastern item must be considered an innovation after some differentiation of Eastern subgroups had occurred.

¹⁷This group should not be confused with the "Togo-Remnant" languages which are not a single group but are distributed between the two subdivisions.

- (6) *nu (Guang branch of Volta-Comoe, / *te ~ se (Volta-Comoe, Ewe-Fon) 'hear'
 Tu, Gã-Adangme, Mo; also in
 Lower Niger, Kru, Gur)
- *nyo (Volta-Comoe, Tu) / *mo-bõ~mi (Togo, Gã-Adangme) 'oil'
 *tõ (Togo, Adele) / *nu (Nyo except Adele) 'five'

Some such items may reflect the close cultural links in this area, in effect borrowing. However, it is conceivable that further data will force reevaluation of relationships at this level. An interpretation viewing such isoglosses as the item 'two' as innovation originating in Volta-Comoe and spreading across linguistic boundaries would not be unreasonable. Supporting this is the fact that in many cases the form predominant in the Togo branch is shared with the remainder of SCNC.

The status of the Yi languages as a valid subgroup is less dubious. This group includes the Volta-Comoe group and the Tu languages, the latter being a subset of the Togo-Remnant languages. There are important innovations in both this and the Gã-Adangme branch of the Nyo group. Some typical isoglosses include:

- (7) *yi (Yi) / *ti (Gã-Adangme) 'tree' (Is.3. This isogloss is diagnostic for Yi branch)
 *tu (Yi) / *ni (Gã-Adangme) 'water' (Is.3)
- (8) *na (Yi) / *jwe (Gã-Adangme) 'four' (Is.3a)
 *bĩ~mi (Yi) / *fĩ (Gã-Adangme) 'excrement' (Is.3a)

In (7) the Yi form is innovative, that in Gã-Adangme shared with the rest of SCNC; in (8) the Gã-Adangme form is innovative, the Yi form inherited. Enough such cases exist to convince us of the viability of these subgroups.

The border between Volta-Comoe and Tu is equally clear. Volta-Comoe shows a number of phonetic shifts, as of *t to *s in such items as:

- (9) *sã (Volta-Comoe) / *tã (Tu) 'three' (Is.4)
 *sú (Volta-Comoe) / *tú (Tu, origin of the name of the 'water' (Is.4)
 group)

Related is the isogloss:

- (10) *wú (Volta-Comoe) / *kú(i) (Tu) 'die' (Is.4)

It is apparent that Volta-Comoe is the innovator in most instances. The Tu languages, therefore, may not constitute a coherent group, but rather a cluster of coordinate conservative groups. Within Volta-Comoe we may recognize a distinction between Guang and Akan based on numerous lexical, statistical, and typological arguments. For subdivisions of Akan, see Trutenaу [1976].

In the Togo languages it is easier to differentiate the Ewe-Fon and Mo subgroups than it is to justify the valid subgroup status of the latter or their relationship. The isogloss:

(11) *mi (Ewe-Fon) / *mo ~ bõ (Mo, origin of the name of the group) 'oil'

as indicated in the discussion of the isoglosses in (6) above, does not clearly indicate common innovation in either branch; such innovations as *dũ 'tooth' linking Ewe-Fon and Mo are not general in the latter. Statistically there is some basis for the grouping, but this may simply be due to contact. Further investigation is needed to settle this point.

The position of the so-called Lagoon languages is uncertain. Data are still few, in most cases unreliable, and the languages are highly diversified. Avikam is clearly West SCNC, probably closely affiliated with, if not part of, Volta-Comoe. Adjukru, though it has been classed as West Atlantic by some, is clearly a part of the Nyo group, but probably deserves the status of a coordinate branch. We consider this a prime area for future investigation.

The subclassification of Western SCNC here discussed is presented in Table 9.

3.3. Eastern SCNC. This is a larger, more complex group than Western SCNC. The principal isoglosses separating Eastern SCNC from Western SCNC have already been discussed under the latter group. It is significant that no certain innovation common to the group has been found. This raises the possibility that some or all of the subgroups included are in fact branches of SCNC coordinate with West SCNC and Ijo. Further data may, however, reveal such innovations. It is noteworthy, however, that Western SCNC (in whole or in part) often is seen to share with NCNC where Eastern SCNC does not. This may support the status of Eastern SCNC as a subgroup.

The nine divisions of Eastern SCNC recognized here (see Table 9) form a

lexicostatistical and typological continuum. Each language group shares lexical and structural characteristics with its geographical neighbors. Investigation of shared innovations reveals a singular lack of pattern in the distribution of isoglosses, in contrast to the clustering tendency of major Western SCNC isoglosses. We therefore have treated the nine as coordinate branches. It is possible that further data may allow a grouping of Efikoid with Delta-Cross or Eastern Cross. These groups, included in Greenberg's Cross River groups, are but poorly represented in the data.

3.3.1. Typology: three branches or nine? The typological evidence at first glance supports grouping the nine divisions of Eastern SCNC into larger units. Such groupings, however, are not supported by detailed investigation. Most of Central Niger, Yoruboid, and Edo form a cluster of "Kwa-type" languages, characterized by the absence--or great reduction--of class systems, typologically similar developments of the verbal system emphasizing serialization and tense-aspect-mood auxiliaries, and a fairly high proportion of shared vocabulary, whose significance is lessened by the fact that much of it is cultural.

Benue-Zambesi and Eastern Cross (and perhaps parts at least of Efikoid) form on the other hand a group of "Benue-Congo-type" languages, for the most part with well preserved class systems, a tendency to more inflection in verbal systems, and again a fairly high degree of shared vocabulary.

Between, however, are the transitional groups Lower Niger, Jukunoid, Delta-Cross, and Efikoid, whose characteristics are intermediate between the two extremes. Study of the statistical data indicates that the majority of the sharings apparently supporting the Kwa-like and Benue-Congo-like groupings are due to geographical proximity rather than common innovation.

A closer look at the principal typological feature; the degree of retention of the class system, likewise shows the influence of geography. Operative SCNC class systems are preserved in most of Benue-Zambesi, some parts of Western SCNC including the Mo and Tu groups and some of the Lagoon languages, Gade and Yala of Central Niger, and Eastern Cross at least. Indications of recently operative class systems in the form of singular/plural alternations appear in Western SCNC in most of Volta-Comoe, in Central Niger in Gwari, in much of Edo, some parts of Lower Niger and Jukunoid, and

portions of Delta-Cross, Eastern Cross, and Benue-Zambesi. Languages with little or no active trace of class systems occur in all groups. A glance at a map reveals a fairly cohesive belt of "classless" languages running from Cã-Adangme through to Efikoid. Those subgroups which show both class languages and non-class languages normally show a distribution such that the classless languages are those in closest proximity to the uniformly classless subgroups. Of the reduced or classless systems, the eastern-most groups generally show the most numerous signs of recently operative class systems. We have here a clear case for a spread of class reduction probably from a Western SCNC center such as Ewe-Fon, following the coast and to a large extent concentrating on the languages of the city-states.

Many of the Eastern SCNC subgroups show significant internal subgrouping, especially Benue-Zambesi. Of these, the internal relationships of Yoruboid, Edo, Lower Niger, and Jukunoid are adequately treated elsewhere and will not be considered here in detail.

3.3.2. Central Niger. This group is linked by a considerable proportion of common vocabulary. Some specific items include:

- | | | | |
|-----------|----------|---|-----------------|
| (12) *ɣwó | 'ten' | (possible cognates in Yoruboid, Edo) | (Is.5, Table 8) |
| *yé ~ yá | 'year' | (possible cognates in Plateau) | (Is.5) |
| *me | 'hunger' | | (Is.5) |
| *ma | 'bear' | (possible cognates in Jukunoid, Delta-Cross, Lower Niger, Efikoid and Cara) | (Is.6) |

A problem is the relative lack of clear common innovation, coupled with a high degree of statistical cognacy with other groups and significant internal deviation. We have here a loosely knit cluster of well differentiated conservative languages. Most have lost class systems, but Gade and Yala still show them, and other languages show signs of recently operative systems. The grouping of Niger-Kaduna with Idomoid is tentative, but is felt to be justified.¹⁸

¹⁸Note that the classification of Idoma with what is now called Niger-Kaduna has been suggested before by R.G. Latham [1862] in a classification which also anticipated Greenberg's assignment of Fula to West Atlantic.

Idomoid is characterized by frequent retention of Proto-Central Niger *ɣ as g ~ j in such items as

- (13) *gwó (Idomoid) / *wó (Niger-Kaduna) 'ten'
 *ēdyé (Yatye), *yí (Idoma, Igede) / *yé (Niger-Kaduna) 'eye'
 *gù (Yatye), *ɣmó (Idoma, Igede) / *wu ~ wū (Niger-Kaduna) 'kill'

and by the shift of *b to *wy before front vowels in

- (14) *wyi (Idomoid) / *bí (Niger-Kaduna) 'child'
 àwyī (Yatye), *mi (Idoma, Igede) / *mi ~ bi (Niger-Kaduna) 'faeces'

Compare the examples in (15) with Proto-Central Niger *-pé 'moon'; we see the same shift with the voiceless labial:

- (15) ðtywā (Yatye), *wyá (Idoma, Igede) / *pwyá (Gade, Nupe), *fé (Igbiroid)

Lexical isoglosses separating Idomoid from Niger-Kaduna include:

- (16) *ku (Idomoid) / *cúkũ (Niger-Kaduna) 'bone'
 *í (Idomoid) / *bí (Niger-Kaduna) 'grass'
 *pà (Idomoid) / *bà (Niger-Kaduna) 'two'
 *hwó (Idomoid) / *kõ (Niger-Kaduna) 'war'

Niger-Kaduna may be divided into Gade, Nupe-Gwari, and Igbiroid. Though the three groups are easily distinguishable on the basis of lexical and phonologic isoglosses, their unity is attested by such isoglosses as that for 'bone' above, and:

- (17) *hĩ ~ hú (Niger-Kaduna) / *nu (CNC) 'drink'
 *tókpa (Niger-Kaduna, also in Mo) / *toŋ (CNC) 'ear'
 *wó (Niger-Kaduna) / *pó (Idomoid) 'hear'

The group as whole is phonologically rather conservative, but drops *ɣ as noted above, and tends to reflect original stops as fricatives; *ti, for example, often appears as Niger-Kaduna *ci (for example in 'tree'); *ku as *cu in 'die'.

Gade is characterized by loss of nasal vowels, much ideosyncratic vocabulary, and an operative class system. No other member of this subgroup is known.

Nupe-Gwari is set off by a number of shared innovations, such as

- (18) *nyikã (Nupe-Gwari) / *nyi (CNC) 'tooth'
 *gya (Nupe-Gwari) / *b̄ayi (Gade), *nyá (Igbiroid) 'blood'
 *yá (Nupe-Gwari) / *kpà (Gade, Igbiroid) 'boat'

covering cultural and non-cultural vocabulary.

Igbiroid shows innovative lexicon such as *mù 'tail', *h̄é 'nose', *nyá 'blood', and a tendency to introduce a *re- prefix, especially in body parts:

- (19) *r̄èsó (Igbiroid) / *to (SCNC) 'head'
 *ìrénú (Igbiroid) / *nu (CNC) 'mouth'

Igbiroid typologically is intermediate between the rest of Niger-Kaduna and Idomoid, but on the basis of shared innovation is clearly Niger-Kaduna.

See Table 9 for the subclassification of Central Niger.

3.3.3. Benue-Zambesi. Benue-Zambesi is a large and important group of languages. Lexicostatistically it is well defined. On the basis of typology and shared innovation it is not. Eastern Cross languages show nearly identical class systems, and we have found no significant isogloss matching the statistically determined border. Further, as the border languages in Eastern Cross and Cara are poorly represented in the statistical study, it may well be that even the statistical evidence is illusory.

The true status of Cara is unclear. The group corresponds to parts of Greenberg's Plateau; the name derives from the isogloss

- (20) *cara (Cara) / *bók (Niger-Congo) 'hand' (Is.7, Table 8)

Typologically, some parts of Cara, especially Western Cara (including Kambari, Dakarkari and Reshe), show similarities to NCNC. It is quite possible that Cara should be classed as least as a tenth coordinate branch of SCNC. However, it shares with the Nyama languages the retention of *s in 'four' and 'breast'¹⁹ (Is.8, Table 8), and considerable vocabulary, so the relationship is not rejected. Further evidence is needed. Evidence available seems to support a division into Eastern and Western Cara but the data are inadequate for a firm classification.

¹⁹*s in 'four' and perhaps 'breast' is found in Gur. Many parts of Nyama have lost the *s in these items, but traces remain.

Much more important as a subdivision of Benue-Zambesi is the Nyama group. This is based on the isogloss

(21) *nyàmà (Nyama) / *nàmà (Niger-Congo) 'animal, meat' (Is.9)

The isogloss marking palatal-initial consonant in this item is apparently exactly paralleled by the isogloss indicating the presence of tonally contrasting concord markers in the class system. The group shows considerable lexical and typological cohesiveness. It includes portions of Greenberg's Plateau group (our Plateau), Cross River (Bendi), and all of Bantoid. There is considerable internal differentiation.

3.3.4. Plateau. The most important boundary in Nyama seems to be that dividing Plateau from the Wɛl languages. This is based on the isoglosses:

(22) *pɔk (Plateau)²⁰ / *wɔk (Wɛl) 'hear' (Is.10)
 *pɛĩ (Plateau and CNC) / *wɛl (Wɛl) 'moon' (hence the name of (Is.10)
 the group)
 *kɔm (CNC) / *jála (Wɛl) 'hunger' (Is.10)

These (and others) for the most part indicate shared innovation in Wɛl. Because of this and the still poor documentation of Plateau, it is again possible that Plateau is not a single valid subgroup but a cluster of coordinate groupings. The subgroups of Plateau are shown in Table 9.

3.3.5. Wɛl. Within the Wɛl group the primary cut is hard to find. Perhaps most important is the question of the position of Bendi, comprising the Boki and Alege dialect clusters formerly assigned to Cross River. Their status is unclear. They share a large body of vocabulary with Wɛl, including *jwó 'hear' reflecting Proto-Wɛl *wók²¹, and perhaps *jála 'hunger'. But

²⁰Also in Efikoid, Idomoid, and Jukunoid. Notice the parallel between the Central Niger isogloss separating Niger-Kaduna *wó from Idomoid *pó and this. The Proto-CNC form for 'hear' seems to have been *nu (innovative in CNC). This is retained in West SCNC, Lower Niger, and much of NCNC. The stem *pók seems to have been an innovation in SCNC. If the Yoruboid *gbó is related we would then have an interesting situation with voiceless initial over a uniform central area, and peripheral cognates showing a voiced consonant. The parallel shift of *pɛĩ > *wɛl, however, is limited to this group.

²¹The j presumably reflects an earlier form such as Proto-Ekoid *ɣwók.

the other item in our data attesting the Wɛl *p ~ *w shift, 'moon', appears in Alege (the only Bendi language available with a reflex of this item) as kofi²² whose voiceless consonant would seem to indicate non-Wɛl affiliation. The group also shares *mo 'water' with Efikoid and Delta-Cross. Pending further investigation we class Bendi as a subgroup of Wɛl.

The remainder of Wɛl constitutes a fairly clearly defined grouping essentially identical to Greenberg's Bantoid, which name we will accordingly retain. Lexicostatistically and typologically this may be divided into Mambiloid (comprising Mambila and Bute) and the Bin languages, named by the isogloss of *bín 'to dance' which follows this line,²³ as apparently does *kádì 'woman' (Is.11, Table 8).

The Bin languages are divided into two groups based on a number of isoglosses which include the items:

- (23) *wɔk (Jarawan, Ekoid, Mbam-Nkam, / *ungwa (Tiv and the re- 'hear' (Is.12)
Bantu zones A, B, C and remainder of Bantu)²⁵
part of D)²⁴

(this isogloss provides the names of the groups)

*-ɔŋ (Wɔk)	/	*nyúéié (SCNC)	'hair'
*-baŋ (Wɔk)	/	(no agreement elsewhere)	'red'

The picture is complicated by areally spreading items covering parts of the northwestern Bin languages, such as the innovative *díbà 'water' (elsewhere 'lake'), *-songa 'tooth', and *páde 'two'. It is further complicated by the great number of languages involved. We will not attempt to discuss the subgrouping of either Bin or Ungwa languages here. Our data on this area have not yet been fully analyzed, and it is not certain that

²²This form, however, is from Koelle whose transcriptions are not always reliable in the area of labial continuants.

²³Note that possible reflexes of *bín have been found in Edo. The validity of this specific isogloss is uncertain, though the existence of a boundary here is clear.

²⁴A retention also found in Mambiloid and Bendi.

²⁵Possible cognate with Kambari uwwa ; in parts of Bantu it appears as *-ígu- .

they will prove adequate for subgrouping at this level of relationship.

4. Conclusion

The use of statistical and comparative data has allowed us to make what seems a significant step toward the subgrouping of SCNC. It has yielded some unexpected revisions of the traditional classification, while confirming previous conclusions in a number of areas. Some of the more significant departures from previous work are in fact not unexpected. The non-SCNC status of Kru, the absence of a true division between "Kwa" and "Benue-Congo", and even the presence of a distinction between the Northwest and the remainder of Bantu has been the subject of speculation before this.²⁶ The hypothesized affiliations of Kru with Gur and Adamawa-Eastern, and of Zambesi and Equatorial Bantu with Tiv and Jarawan respectively rather than with each other are rather more controversial. The break-up of "Cross-River", the "Togo-Remnant" group and Plateau, and the joining of Niger-Kaduna and Idomoid are less startling but still relatively innovative.

Many questions remain. This is a very preliminary study, still based in many cases on old and inadequate word lists. Even where full lists are available, some areas are not well enough documented to allow safe subgrouping. Even where good coverage is possible, rigorous comparison is impossible. For some items we can recognize cognacy over wide areas and construct "pseudo reconstructions" with some hope of accuracy in statements of distribution and probable original form. We can recognize--and use in cognacy evaluations--recurrent sound shifts in many cases, some of which have been explicitly used above. But in our short lists over this wide area there are not enough cognates to testify enough recurrent shifts to permit certain postulation of regularity and actual reconstruction. It is doubtful if even the sound system of Guthrian Bantu could be reconstructed from these data.

²⁶See, for example, Henrici [1973]. Note that the high statistical percentages and typological similarities within "Bantu" are explained by the proximity of the two branches and their relative isolation from innovative centers in Nigeria.

Within SCNC it remains necessary to perform follow-up investigations of the validity and exact limits of these groupings, based on close study of larger bodies of data. Outside SCNC it is necessary to perform this type of preliminary study of NCNC and West Atlantic, and after that to consider the problem of the relationship of Mande and Kordofanian.

REFERENCES

The following bibliography is not intended to be exhaustive. It would be impossible to list all the language sources consulted. Only previous comparative works relevant to South Central Niger-Congo have been included, though some of these are also major sources of the vocabularies forming the bases of this study.

In addition to the published works cited below, this study rests on a number of previously unpublished comparative studies by the authors and others at the University of Wisconsin, Madison. These included a general survey of Niger-Kordofanian interrelationships by P.R. Bennett based on 31-item vocabularies from most documented Niger-Kordofanian languages; a study of Niger-Kordofanian based on body-part vocabulary by Joko Sengova; a comparative Benue-Congo list (based on the glosses to Guthrie's P.B. "starred forms") by P.R. Bennett; a comparative study of Plateau languages by L. Townsend, M. Higbie and P.R. Bennett; a comparative study of "Kwa" languages, and a study of three Edo dialects by J.P. Sterk.

We are grateful to H. Stahlke for sending us several long word lists of Eastern SCNC languages, especially Idomoid. Other sources of lexical materials include a number of unpublished word lists collected by the authors and others, including I. Dihoff, M. Higbie and J. Ellington. All those who consciously or unconsciously have contributed materials and constructive criticism are herewith gratefully acknowledged.

- Armstrong, R.C. 1963. "The Kwa working group at Dakar." *Actes du Second Colloque International Linguistique Négro-Africaine*. Dakar: Université de Dakar.
- Armstrong, R.C. 1964. *The Study of West African Languages*. Ibadan: Ibadan University Press.
- Bennett, P. R. 1973. "Identification, classification, and Bantu linguistics." *African Language Studies* 16:19-25.
- Bennett, P.R. 1974. "A reevaluation of the linguistic classification of the Kru-speaking peoples." Paper presented to the 6th Annual Liberian Studies Conference.
- Coupez, A., E. Eurard, and J. Vansina. 1975. "Classification d'un échantillon de langues bantoues d'après la lexicostatistique." *Africana Linguistica* 6:131-58.

- Crabb, D.W. 1965. *Ekoid Bantu Languages of Ogoja: Part I*. West African Language Monographs, 4. Cambridge: University Press.
- de Wolf, P.P. 1971. *The Noun Class System of Proto-Benue-Congo*. The Hague: Mouton.
- Fresco, E.M. 1970. *Topics in Yoruba Dialect Phonology*. *Studies in African Linguistics*, Supplement 1.
- Gerhardt, L. 1967-70. "Analytische und vergleichende Untersuchungen zu einigen zentralnigerianischen Klassensprachen." *Afrika und Übersee* 51:161-98; 52:23-57, 125-43, 207-42; 53:44-65.
- Gerhardt, L. 1972-73. "Das Nominalsystem der Plateau-4-Sprachen: Versuch einer Rekonstruktion." *Afrika und Übersee* 56:72-89.
- Greenberg, J.H. 1955. *Studies in African Linguistic Classification*. Branford: Compass Publishing Company.
- Greenberg, J.H. 1963. "History and present state of the Kwa problem." *Actes du Second Colloque International Linguistique Nègro-Africaine*. Dakar: Université de Dakar.
- Greenberg, J.H. 1966. *The Languages of Africa*. Bloomington: Indiana University.
- Greenberg, J.H. 1972. "Linguistic evidence regarding Bantu origins." *Journal of African History* 13:189-216.
- Greenberg, J.H. 1974. "Bantu and its closest relatives." In W.R. Leben (ed.), *Papers from the Fifth Conference on African Linguistics*, pp. 115-19. *Studies in African Linguistics*, Supplement 5.
- Guthrie, M. 1967-70. *Comparative Bantu*, 4 Vols. Farnborough: Gregg Press.
- Heine, B. 1968. *Die Verbreitung und Gliederung der Togorestsprachen*. Berlin: Dietrich Reimer Verlag.
- Heine, B. 1974. "Historical linguistics and lexicostatistics in Africa." *Journal of African Languages* 11(3):7-20.
- Henrici, A. 1973. "Numerical classification of Bantu languages." *African Language Studies* 16:82-104.
- Hoffmann, C. n.d. "The languages of Nigeria by language families." Ms., Department of Linguistics and Nigerian Languages, Ibadan.
- Ibadan, University. n.d. "Language families of Africa." Ms., Department of Linguistics and Nigerian Languages, Ibadan.
- Koelle, S.W. 1854. *Polyglotta Africana*. Graz: Akademische Druck: U. Verlagsanstalt. [Reprinted 1963.]
- Latham, R.G. 1862. *Elements of Comparative Philology*. London.
- Maddieson, I., and K. Williamson. 1975. "Jarawan Bantu." *African Languages/Langues africaines* 1:124-63.
- Shimizu, K. 1969. "Preliminary Report on the Comparative Study of Jukun, Tigong and Mbembe." Paper presented to the Eighth West African Congress, Abidjan.

- Sterk, J.P. 1977. "Elements of Gade Grammar." Doctoral dissertation, University of Wisconsin.
- Stewart, J.M. 1976. *Towards Volta-Congo Reconstructions*. Leiden: Universitaire Press.
- Trutenau, H.M.J. 1976. "Introduction to languages of the Akan area." *Papers in Western Kwa Linguistics and on the Linguistic Geography of the Area of Ancient Begho*. Transactions of the Linguistic Circle of Accra, Vol. VIII. Basel: Basler Afrika Bibliographien.
- Voorhoeve, J., ed. 1971. *The Mbam-Nkam Languages*. *Journal of African Languages* 10(2).
- Voorhoeve, J. 1976. "Bantu/Bane." Paper presented to the 12th West African Languages Congress, Ile-Ife.
- Voorhoeve, J., and P.P. de Wolf, eds. 1969. *Benue-Congo Noun Class Systems*. Leiden: West African Linguistic Society.
- Westermann, D. 1927. *Die Westlichen Sudansprachen und ihre Beziehungen zum Bantu*. Berlin: Walter de Gruyter.
- Westermann, D. 1952. "African linguistic classification." *Africa* 22:250-56.
- Williamson, K., ed. 1968, 1973. *Benue-Congo Comparative Wordlist*, 2 vols. Ibadan: West African Linguistic Society.
- Williamson, K. 1971. "The Benue-Congo Languages and Ijo." In T.A. Seboeok (ed.), *Linguistics in Sub-Saharan Africa*, 245-306. *Current Trends in Linguistics*, Vol. 7. The Hague: Mouton.
- Williamson, K. 1973a. "The Sound System of Proto-Lower Niger." Paper presented to the 4th African Linguistics Conference, New York.
- Williamson, K. 1973b. "The Lower Niger Languages." *Òdúmá Magazine* 1:1.
- Williamson, K. 1976. "Noun Class Prefixes in Proto-Lower Niger." Paper presented to the 12th West African Languages Congress, Ile-Ife.

APPENDIX

Table 1: <u>The fifty languages</u>		Table 2: <u>List of the glosses</u>			
1. Grebo	26. Kaje	animal	egg	moon	tongue
2. Newole	27. Chori	arrow	eye	mother	tooth
3. Asante	28. Afusare	back	father	mountain	tree
4. Larteh	29. Aten	belly	fire	mouth	two
5. Lelemi	30. Kambari	bird	fish	name	urine
6. Ga	31. Jukun	black	five	neck	war
7. Ewe	32. Kpan	blood	four	night	water
8. Gwari	33. Tiv	body	go	nose	white
9. Gade	34. Mambila	bone	goat/he-goat	oil/fat	wife
10. Nupe	35. Eloyi	bow	guest	one	woman
11. Igbira	36. Tunen	buy	hair	person	yam
12. Idoma	37. Jarawa	chicken/cock	hand/arm	rain	year
13. Igbo	38. Nyanja	child	head	red	
14. Igala	39. Bobangi	chin/jaw	hear	river	
15. Ife (Togo)	40. Kikuyu	come	heart/liver	saliva	
16. Yoruba	41. Kwanyama	crocodile	honey/bee	see	
17. Ora	42. Fula	dance	house	shoulder	
18. Bini	43. Dyola	day	hunger	six	
19. Urhobo	44. Temne	die	husband	skin	
20. Isoko	45. Mossi	dog	iron	slave	
21. Degema	46. Kassene	dream	kill	snake	
22. Ijo	47. Mamprusi	drink	king/chief	stone	
23. Abua	48. Tula	ear	leaf	sun	
24. Efik	49. Gbaya	earth	leg/foot	ten	
25. Ogoni	50. Ndogo	eat	man	three	

Table 4: "10 closest", i.e. for each language, the ten languages sharing the most cognates (See Fn. 3 for instructions on how to read chart)

		1	2	3	4	5	6	7	8	9	10
1. Grebo	A.	2:36	16:25	14:23	15:22	11:22	36:21	19:20	33:20	12:20	10:19
	B.	2:21	46:7	16:6	33:6	34:6	15:6	50:6	38:5	14:5	24:5
2. Newole	A.	1:36	36:24	11:21	7:21	40:20	21:19	24:19	12:19	18:19	3:18
	B.	1:21	7:7	12:7	15:6	16:6	21:6	40:6	18:5	14:5	23:5
3. Asante	A.	4:39	5:31	7:27	16:27	45:26	9:25	12:25	6:23	36:23	14:23
	B.	4:32	5:14	7:12	6:11	24:9	37:9	23:8	45:8	19:7	36:7
4. Larteh	A.	3:39	6:25	7:25	5:24	36:24	45:21	12:20	40:20	30:20	9:19
	B.	3:32	6:14	7:13	5:12	36:9	12:8	45:8	21:7	24:7	46:7
5. Lelemi	A.	3:31	7:29	12:28	16:27	37:25	13:24	15:24	10:24	4:24	14:23
	B.	3:14	7:13	4:12	12:12	23:10	16:8	28:8	35:8	47:8	33:7
6. Ga	A.	4:25	3:23	12:22	5:19	11:17	7:17	30:17	45:17	40:16	10:16
	B.	4:14	3:11	7:10	11:8	5:6	9:5	8:5	30:5	33:5	45:5
7. Ewe	A.	5:29	3:27	12:27	20:27	16:26	14:26	21:26	19:25	8:25	4:25
	B.	4:13	5:13	12:13	8:12	3:12	14:12	6:10	9:10	13:9	16:9
8. Gwari	A.	10:49	9:42	12:38	11:34	16:33	14:32	15:28	7:25	18:25	19:25
	B.	10:38	9:25	12:14	11:13	7:12	14:10	32:10	16:8	28:7	15:7
9. Gade	A.	8:42	12:40	11:39	10:37	16:37	14:34	15:33	36:30	35:27	31:26
	B.	8:25	11:22	12:21	10:20	35:13	45:11	7:10	18:10	28:9	29:9
10. Nupe	A.	8:49	11:42	12:38	9:37	16:35	15:31	14:29	36:25	5:24	20:24
	B.	8:38	9:20	11:17	12:13	7:7	15:7	16:7	32:7	35:7	25:6
11. Igbirra	A.	10:42	9:39	12:38	8:34	16:33	15:32	14:28	36:27	13:26	21:25
	B.	9:22	10:17	12:15	8:13	18:12	17:11	13:11	15:8	7:8	6:8
12. Idoma	A.	9:40	16:40	10:38	11:38	8:38	14:35	35:35	15:34	32:31	13:31
	B.	9:21	35:20	11:15	8:14	7:13	10:13	14:12	5:12	17:11	13:10
13. Igbo	A.	16:34	15:33	12:31	14:31	11:26	23:26	30:26	33:26	36:26	40:26
	B.	11:11	14:11	12:10	7:9	9:8	16:8	23:8	35:8	21:7	18:7
14. Igala	A.	16:81	15:73	12:35	9:34	33:34	8:32	21:31	13:31	37:31	36:30
	B.	16:69	15:61	7:12	12:12	13:11	8:10	18:9	17:8	9:7	19:7
15. Ife Togo	A.	16:94	14:73	12:34	9:33	13:33	11:32	18:32	19:32	20:32	10:31
	B.	16:89	14:61	17:11	18:11	33:9	11:8	7:8	8:7	19:7	10:7
16. Yoruba	A.	15:94	14:81	12:40	18:39	9:37	10:35	19:35	20:35	33:35	21:34
	B.	15:89	14:69	18:16	17:13	12:10	33:10	19:9	7:9	8:8	5:8
17. Ora	A.	18:68	20:50	19:48	21:47	16:33	15:29	12:29	14:27	9:25	8:24
	B.	18:62	19:35	20:31	21:27	16:13	15:11	11:11	12:11	7:9	9:9
18. Bini	A.	17:68	19:52	20:50	21:46	16:39	15:32	14:30	12:29	31:26	9:25
	B.	17:62	19:40	20:34	21:26	16:16	11:12	15:11	9:10	7:9	12:9
19. Úrhobo	A.	20:78	18:52	21:49	17:48	16:35	15:32	12:28	14:28	7:25	8:25
	B.	20:67	18:40	17:35	21:30	12:9	16:9	7:8	11:8	13:7	14:7
20. Isoko	A.	19:78	18:50	17:50	21:49	16:35	15:32	12:29	14:28	7:27	11:25
	B.	19:67	18:34	17:31	21:30	12:9	16:8	29:8	13:6	15:6	11:6
21. Degema	A.	19:49	20:49	17:47	18:46	16:34	15:31	14:31	12:27	40:27	24:26
	B.	19:30	20:30	17:27	18:26	23:10	25:10	24:8	22:8	30:8	7:7
22. Ijo	A.	12:17	14:17	16:17	17:17	23:17	18:16	21:16	15:16	32:16	7:15
	B.	21:8	23:7	32:7	16:6	19:5	20:5	9:5	15:5	8:5	24:4
23. Abua	A.	25:34	24:27	13:26	12:24	21:24	36:23	40:23	39:22	29:21	5:21
	B.	25:16	21:10	5:10	24:9	48:9	12:8	13:8	36:8	3:8	28:7
24. Efik	A.	36:32	12:29	25:28	28:27	33:27	33:27	40:27	21:26	16:25	26:25
	B.	25:13	23:9	3:9	21:8	30:8	4:7	40:7	48:7	36:6	27:6
25. Ogoni	A.	23:34	24:28	33:25	36:24	16:23	29:23	21:23	14:23	39:23	40:23
	B.	23:16	24:13	21:10	9:9	27:8	32:8	19:7	18:7	16:6	3:6

26. Kaĵe	A.	28:61	29:40	27:37	35:32	16:31	14:29	36:29	39:29	15:28	40:28
	B.	28:53	29:24	27:19	37:13	38:12	40:10	39:9	33:9	7:8	36:8
27. Chori	A.	26:37	28:37	29:29	36:29	40:29	39:28	16:27	33:26	15:26	35:25
	B.	26:19	28:18	30:10	41:10	32:9	29:9	33:8	39:8	25:8	38:7
28. Afusare	A.	26:61	29:40	27:37	35:32	12:30	33:30	16:30	39:30	30:28	14:28
	B.	26:53	29:27	27:18	37:11	30:10	33:10	12:10	38:10	40:10	9:9
29. Aten	A.	26:40	28:40	27:29	36:28	35:25	37:25	30:24	24:24	33:23	12:23
	B.	28:27	26:24	36:12	33:10	40:10	27:9	9:9	37:9	30:9	38:8
30. Kambari	A.	36:33	40:33	41:29	28:28	13:26	26:26	29:24	27:23	24:23	33:22
	B.	40:12	28:10	27:10	29:9	41:9	26:8	21:8	24:8	23:7	37:7
31. Jukun	A.	32:51	16:29	14:29	12:28	33:28	28:26	9:26	18:26	36:26	26:24
	B.	32:29	17:8	18:8	12:8	35:8	27:7	39:6	41:6	16:5	33:5
32. Kpan	A.	31:51	16:31	12:31	14:29	33:28	26:27	28:27	15:26	8:25	9:25
	B.	31:29	12:10	8:10	28:9	27:9	7:8	16:8	25:8	10:7	22:7
33. Tiv	A.	37:41	36:40	40:38	38:36	41:36	16:35	14:34	39:32	15:31	34:31
	B.	34:20	38:17	41:17	37:15	40:14	36:14	16:10	39:10	28:10	29:10
34. Mambila	A.	33:31	37:30	36:29	39:26	40:26	14:24	16:23	15:22	38:22	41:22
	B.	33:20	36:10	40:9	29:7	14:7	15:7	37:7	38:7	39:7	16:7
35. Eloyi	A.	12:35	26:32	28:32	9:27	36:27	39:27	16:26	29:25	27:25	14:25
	B.	12:20	9:13	5:8	13:8	18:8	28:8	31:8	10:7	11:7	17:7
36. Tunen	A.	40:57	41:48	33:40	39:40	37:37	38:34	30:33	24:32	9:30	14:30
	B.	40:29	41:23	37:20	39:18	38:17	33:14	29:12	34:10	4:9	28:8
37. Jarawa	A.	33:41	36:37	38:35	41:33	39:32	40:32	16:32	14:31	34:30	12:29
	B.	36:20	38:18	41:18	39:16	40:15	33:15	26:13	28:11	3:9	7:9
38. Nyanja	A.	40:41	41:40	39:39	33:36	37:35	36:34	14:27	16:27	9:26	26:24
	B.	40:29	41:24	29:22	37:18	33:17	36:17	26:12	28:10	29:8	7:7
39. Bobangi	A.	40:45	41:45	36:40	38:39	37:32	33:32	28:30	26:29	27:28	35:27
	B.	41:30	40:23	38:22	36:18	37:16	33:10	26:9	28:9	27:8	29:8
40. Kikuyu	A.	36:57	41:46	39:45	38:41	33:38	30:33	37:32	27:29	16:28	26:28
	B.	36:29	38:29	39:23	41:23	37:15	33:14	30:12	28:10	29:10	26:10
41. Kwanyama	A.	36:48	40:46	39:45	38:40	33:36	37:33	30:29	16:26	14:25	27:24
	B.	39:30	38:24	36:23	40:23	37:18	33:17	27:10	30:9	7:8	28:8
42. Fula	A.	36:13	50:11	39:10	40:10	45:10	49:10	2:10	11:9	43:9	29:9
	B.	26:4	30:4	49:4	2:3	27:3	7:3	36:3	38:3	39:3	40:3
43. Dyola	A.	40:9	42:9	33:8	45:8	44:7	1:7	19:6	20:6	21:6	31:6
	B.	44:4	45:4	40:3	42:3	23:2	37:2	5:2	47:2	48:2	16:1
44. Temne	A.	33:14	14:13	3:13	18:12	16:12	40:12	28:11	21:11	36:11	38:11
	B.	3:5	28:5	33:5	38:5	23:4	26:4	27:4	43:4	24:3	25:3
45. Mossi	A.	47:59	46:27	3:26	5:23	16:23	14:22	9:22	12:21	18:21	21:21
	B.	47:51	46:12	9:11	12:8	4:8	3:8	5:7	7:7	28:6	36:6
46. Kassene	A.	45:27	47:25	9:22	12:21	3:20	33:20	36:19	38:19	41:19	21:19
	B.	45:12	47:10	9:9	48:9	12:7	34:7	3:7	4:7	1:7	11:6
47. Mamprusi	A.	45:59	46:25	12:21	37:21	5:20	39:20	3:20	36:20	14:19	16:19
	B.	45:51	46:10	5:8	9:8	12:7	7:6	49:6	4:5	8:5	36:5
48. Tula	A.	45:17	46:17	23:15	36:15	50:15	7:14	49:14	5:14	3:13	47:13
	B.	23:9	46:9	24:7	49:7	36:6	45:6	3:6	4:6	25:5	12:5
49. Gbaya	A.	36:21	11:19	15:19	16:19	9:19	12:18	19:17	20:17	7:17	41:17
	B.	30:7	36:7	37:7	48:7	7:6	23:6	41:6	45:6	46:6	47:6
50. Ndogo	A.	14:20	15:20	36:20	39:20	40:20	19:19	21:19	25:19	33:19	13:19
	B.	9:7	11:7	8:6	1:6	7:6	12:6	15:6	19:6	25:6	17:5

Table 5:

GEOGRAPHICAL LOCATION OF THE FIFTY
LANGUAGES

with isoglosses for
'HEAD'

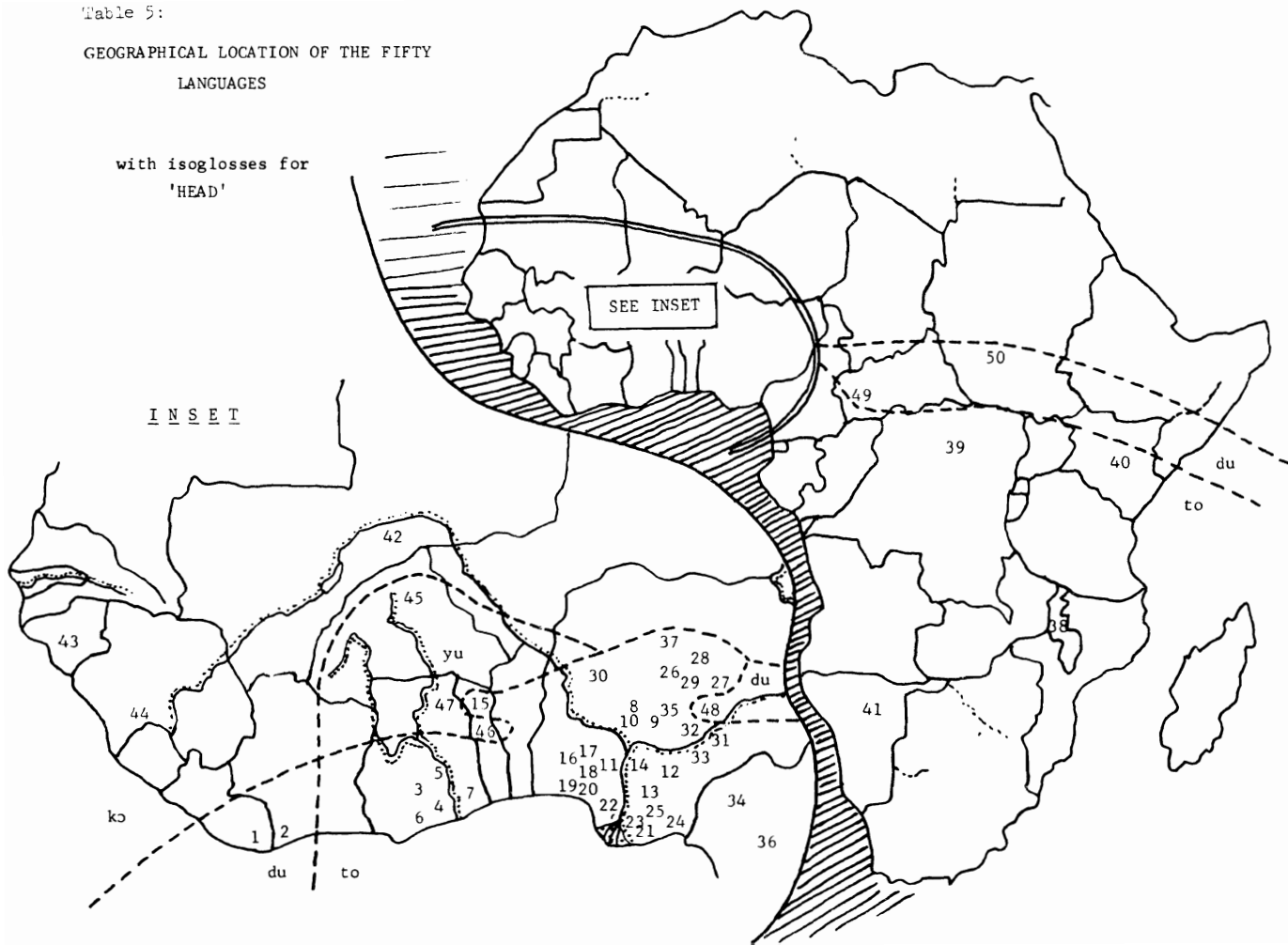


Figure 1. Language families

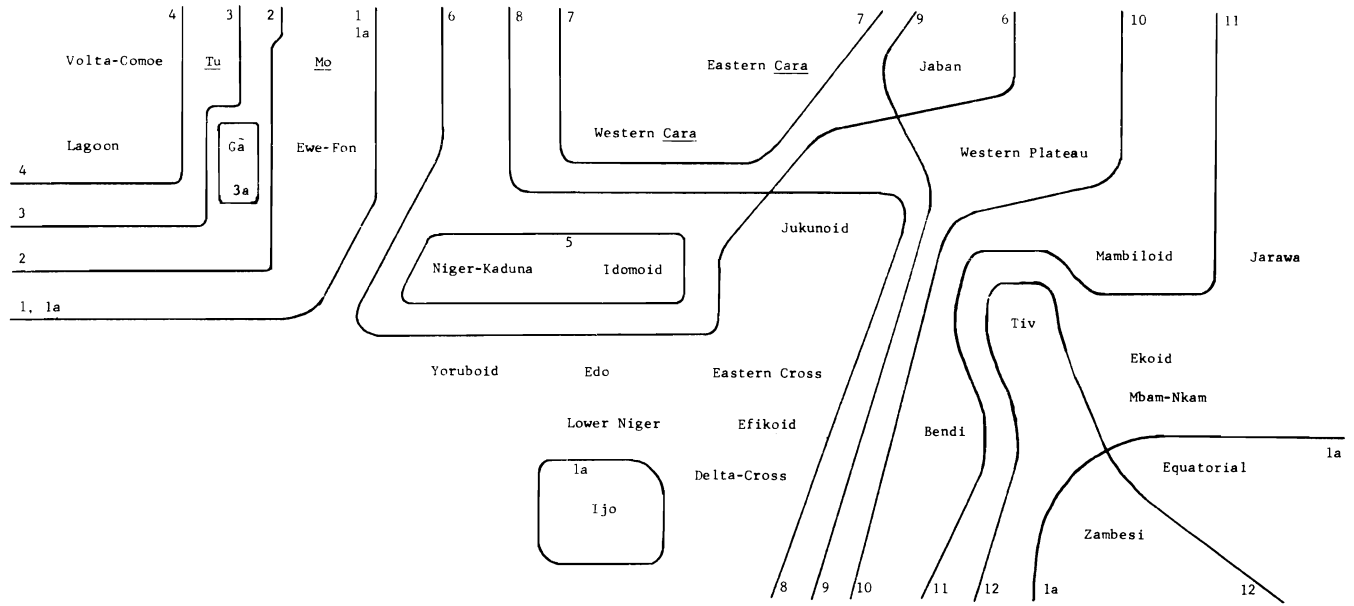


Table 9: Bennett-Sterk proposed classification

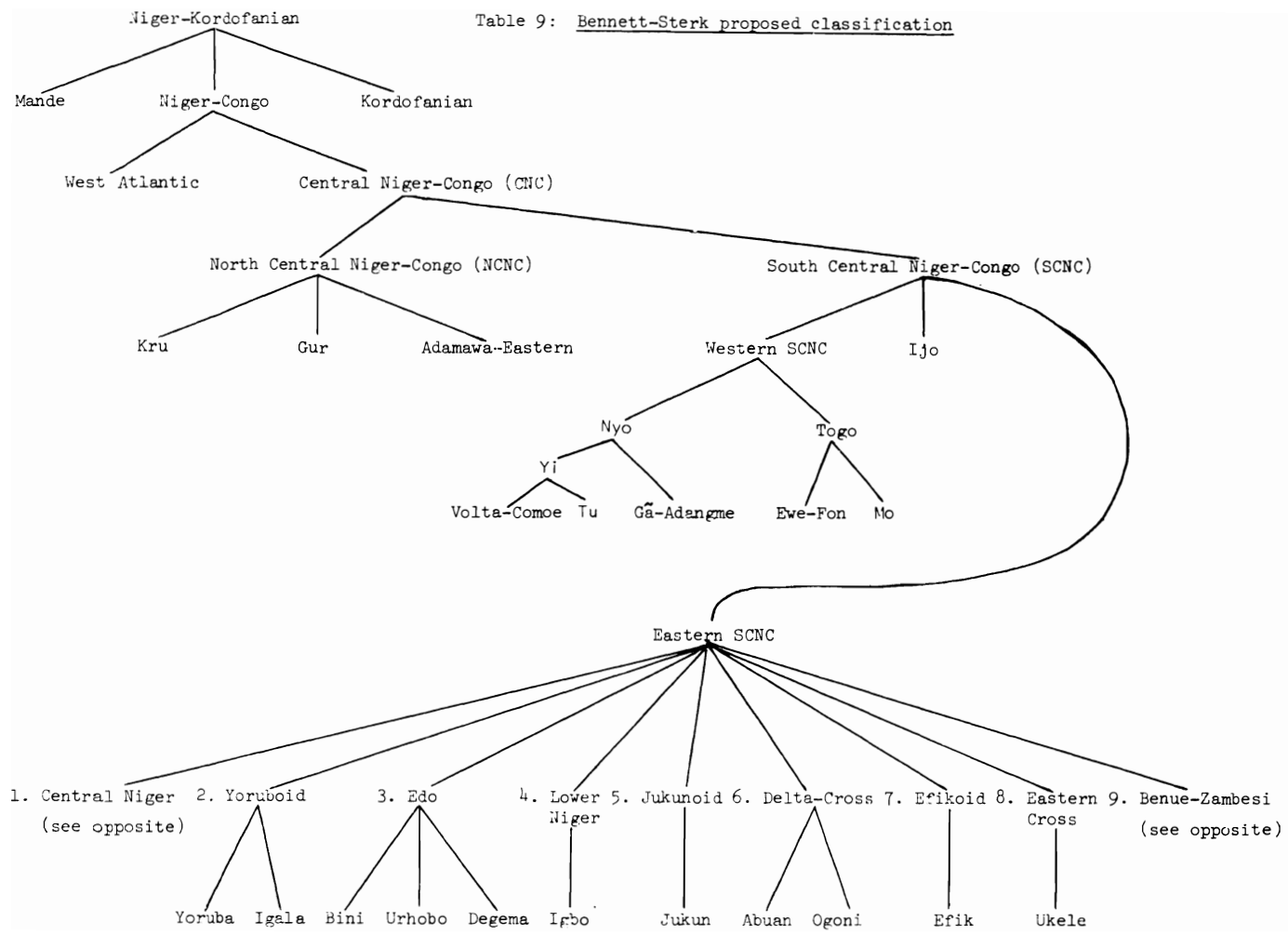
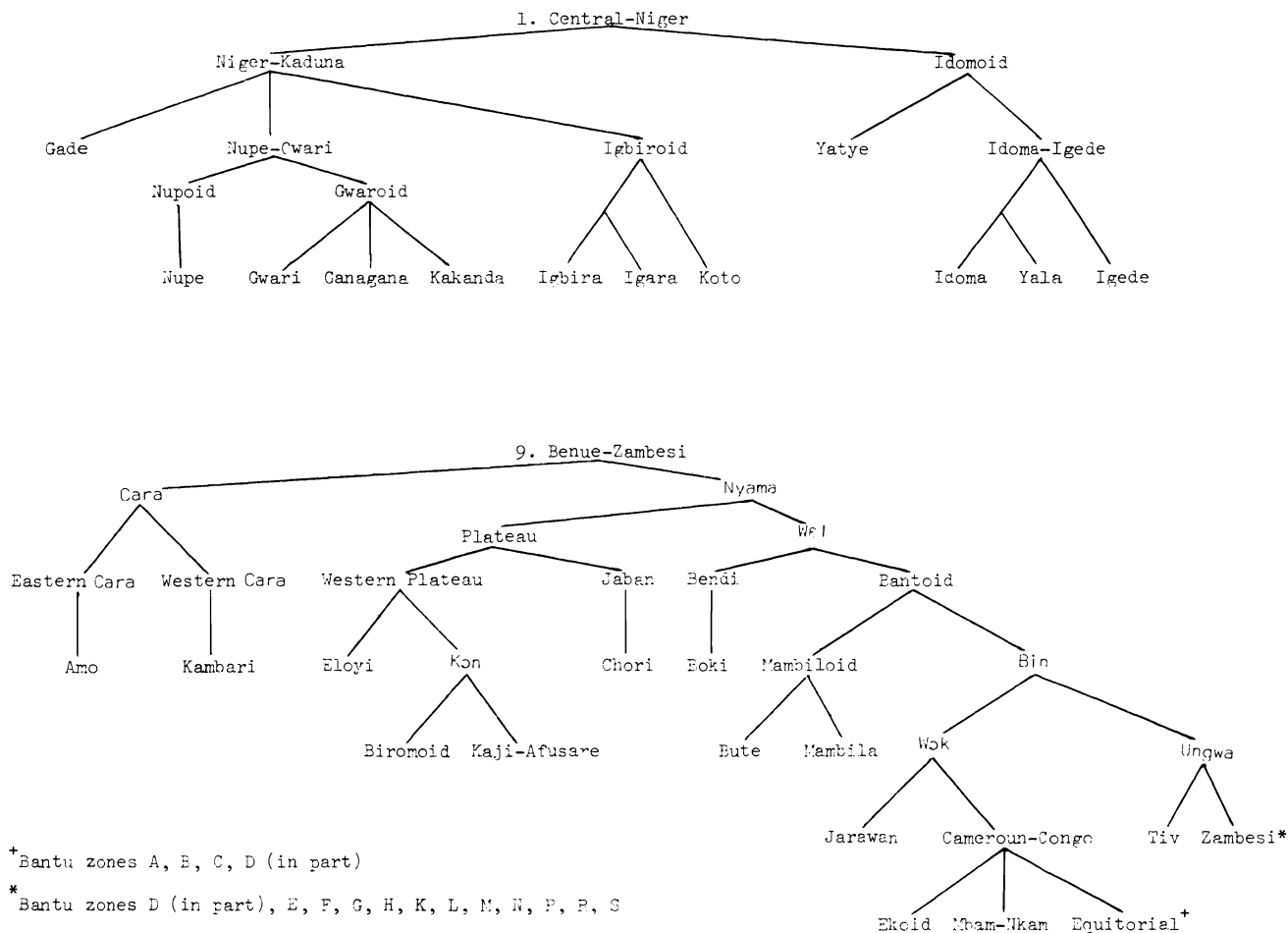


Table 9 (cont.)



CHADIC EXTENSIONS AND PRE-DATIVE VERB FORMS IN HAUSA*

Paul Newman
Rijksuniversiteit te Leiden

Two derivational extensions are reconstructed for Proto-Chadic: a Distant extension *(a)wa, which places the action of a verb at some distance from or in the direction of the speaker, and a Destinative extension *in, which relates the action to a person as its destination, beneficiary, or otherwise affected party. These extensions are illustrated by descriptions of the form and function of their reflexes in present-day Chadic languages belonging to two of the major branches of the family (West and Biu-Mandara). Hausa, which retains the Distant extension in the grade 6 form of the verb, appears to have lost the Destinative. It is argued, however, that a reflex of the Destinative can be identified in Hausa, although it no longer functions as a derivational extension. This is the unusual pre-dative form with final -ř/-m of grade 2, 3, and 7 verbs. The interpretation of this pre-dative form as a Destinative is offered as a counterproposal to Parsons' analysis of this form as a "borrowed" grade 5.

1. Introduction

In most Chadic languages the meaning of a verb can be developed or modified by the addition of derivational extensions, e.g. Hausa fít-óo 'come out' < fítá 'go out'; Bole yàw-tú 'take down' < yàwwú 'get down'; Margi sà-nyà 'drink all' < sà 'drink'. The extensions are usually indicated by bound suffixes, but separable particles, prefixes, vocalic replacements, and tonal changes are also used. In some languages the extensions are highly productive derivational affixes, while in others, they have become lexically restricted or have moved in the direction of

*Field research incorporated in this study was supported by an NSF grant no. GS-2279. Presentation of the paper at the 8th African Linguistics Conference (UCLA) was made possible by a travel grant from the Faculty of Letters, Leiden University. I am grateful to Roxana Ma Newman for valuable comments on an earlier draft of this paper.

becoming inflectional/grammatical markers. Viewed comparatively, one finds that some of the extensions found in individual languages are unique and show no resemblance to extensional elements in other languages, while other extensions (such as the "transitizer/causative" *da*, the "totality" *anya*, and the "completive" *kwa*) have a wide distribution throughout the family.¹ It can be assumed that Proto-Chadic made use of verbal extensions, although exactly how many and which ones is not known.

In this paper, I shall focus on two particular extensions that I believe can be reconstructed for Proto-Chadic with a high degree of certainty. These are the DISTANT extension **(a)wa* and the DESTINATIVE extension **in*. First I shall describe the presumed nature of these extensions in Proto-Chadic and then present evidence from present-day languages to support the reconstructions. In so doing, I shall also be illustrating the various ways in which these extensions have altered over time. Finally, I shall turn to a peculiar problem in Hausa morphology and suggest a solution based on the identification of a previously unsuspected reflex of the Proto-Chadic "Destinative".

2. DISTANT **(a)wa*

2.1. Function. The original meaning of the Distant extension would have been to indicate that the action of the verb was done "there" at some distance from the speaker. With motion verbs, it may also have indicated movement in the direction of the speaker. In many present-day languages, this "Ventive" sense of "do towards someone" or "do and come to someone" has been generalized so that it applies to all verbs. Moreover, where this directional "Ventive" sense is now primary, as opposed to the original spatial "Distant" sense, the extension sometimes also has accompanying benefactive/dative connotations, i.e. indicating action in the direction of or for the benefit of someone. It is unlikely that these benefactive/dative meanings pertained to the proto Distant extension, which was probably entirely locative in nature.

¹The exact phonological shape of these three extensions in Proto-Chadic is yet to be reconstructed. The forms given here are simply approximations noted for referential purposes.

2.2. Position. In Proto-Chadic, the Distant was probably the extensional suffix most closely bound to the verb stem, i.e. occurring immediately after the verb (which might be inflected for aspect or number) and before pronominal suffixes or other extensions. Note the following example from Gisiga, where the presumably original position of the Distant marker has been preserved: a m-o taŋ de /he/return-Dst/them/Transitizer/ 'he returned them back here'.

2.3. Reconstruction. Reconstruction of the Distant morpheme *(a)wa² is supported by evidence from the following languages, representing two of the three major branches of Chadic.³

(1)

West Branch		Biu-Mandara Branch	
Sha	-ó	Gisiga	-awa/-o(o)
Hausa	-óo/-wóo	Bachama	-á(a)
Ngizim	-àyi/-ài/-ée	Tera	á

2.3.1. Sha. In Sha [Jungraithmayr 1970], one of the languages in the Ron group, a verb stem indicating action at a distance from or in the direction of the speaker (= "*Distanzstamm*") is formed by adding a suffix -ó to the verbal base, whether simple (the "*Grundaspektstamm*") or reduplicated (the "*Habitativstamm*"), e.g. (mid tone unmarked):

(2)	mbùt	'seek'	mbùt-ó	'seek (for me)'
	bên	'jump'	bên-ó	'jump this way'
	dèŋ	'pull'	dèŋ-ó	'pull towards here'
	dyàŋàŋ	'pull (hab)'	dyàŋàŋ-ó	'pull (hab) here'
	lwágág	'go up (hab)'	lwágág-ó	'come up (hab)'

²The proto-form cannot be reconstructed with a final -o, in spite of its repeated occurrence in a number of present-day languages, since this vowel almost certainly did not exist in the Proto-Chadic phonemic inventory.

³In Newman [1977], the Chadic family is subdivided into three major branches (West, Biu-Mandara, and East) plus an isolated group (Masa) that is treated as an independent branch. Unfortunately, I was not able to find information on East or Masa languages that had a bearing on the two extensions in question.

In forming verb stems used in the perfect tense, the extensional suffix *-ó* is added before the perfect marker *-(h)i*.⁴

- | | | | | |
|-----|---------------|----------------|------------------|-----------------------------|
| (3) | <i>mbùt-í</i> | 'seek (perf.)' | <i>mbùt-ó-hí</i> | 'seek (for me) (perf.)' |
| | <i>bén-i</i> | 'jump (perf.)' | <i>bèn-ó-hí</i> | 'jump this way (perf.)' |
| | <i>dèṅ-í</i> | 'pull (perf.)' | <i>dèṅ-ó-hí</i> | 'pull towards here (perf.)' |

2.3.2. Hausa. In Hausa, the proto Distant extension is to be found incorporated in the "grade 6" Ventive verb stems [Parsons 1960/61], i.e. those with (Hi) ... Hi tone ending in *-óo/-wóo*,⁵ e.g. *fít-óo* 'come out' < *fítá* 'go out'; *fáad-óo* 'fall this way' < *fáadí* 'fall'; *sáy-óo* 'buy and bring' < *sàyí* 'buy'; *háng-óo* 'espy at a distance' < *hàngí* 'espy'; *w-óo* (= *γ-óo* = *γí-wóo*) 'do and come' or 'do for me' < *γí* 'do'; *jáa-wóo* 'pull here' < *jáa* 'pull'; *kíráa-wóo* 'call here' < *kíráa* 'call'. As just illustrated, the full suffix *-wóo* is only used with *Cí*, *Cáa*, and *CVCáa* verbs. All others simply replace the final vowel of the verb by *-óo*. Contrary to the presumed proto-pattern, the Ventive suffix in Hausa is not necessarily the first extension to be attached to the verb when more than one is used, e.g.

- | | | |
|-----|-------------------|--|
| (4) | <i>já-ny-óo</i> | 'pull all this way' (vb-Totality-Vnt) |
| | <i>káry-ák-óo</i> | 'break up and bring' (vb-Intensive-Vnt) ⁶ |
| | <i>fíd-d-óo</i> | 'bring out' (vb-Causative-Vnt) |

⁴Jungraithmayr treats *-i* as the perfect suffix and the /h/ as an epenthetic consonant (and similarly for the perfect marker (h)e in the closely related Kulere language). Historically, *hi* is the original form (cognate with Pero *ko*, Bole *Wo*, Ngizim *-w*, etc.), of which the isolated vowel *-i* represents a phonological reduction.

⁵The final vowel of grade 6 verbs is long in all positions. It does not shorten before noun objects, as originally stated but later corrected by Parsons, nor can it be said to be short in pre-pausal position in spite of the fact that it is characterized there by a non-distinctive glottal closure usually associated with final short vowels.

⁶The function of this extension *-ak-* (or *-ikk-*) is unclear. It is not found in standard Hausa and has received only sketchy description in the dialects in which it occurs. According to Taylor [1959:102], "it intensifies the meaning of the simple verb." Historically, it may be a reflex of the Proto-Chadic Completive extension, preserved in many West Chadic languages as a perfect marker (see footnote 4).

The third example in (4), with the dialectal "decausative" [Gouffé 1962], can be contrasted with the equivalent standard Hausa form *fít-óo dà* 'bring out' with the order of extensions reversed.⁷

2.3.3. Ngizim. In Ngizim [Schuh 1972], as in closely related Bade, the extension termed "Ventive" has two suppletive allomorphs, a form -EN used in the imperative and "2nd subjunctive", and a form -AY used in all other tenses. This latter allomorph itself has grammatically conditioned subvariants: -ée in the perfective, -àì in the subjunctive, and -àyí in the imperfective (verbal noun), e.g. (taken from Schuh 1972:26-27):

(5)	<i>tàf-é-w</i> [təfú]	'he went in'	<i>tàf-ée-w</i>	'he came in' ⁸
	<i>tàf-é-dù</i>	'he took in'	<i>tàf-ée-dù</i>	'he brought in' (vb-Vnt-Transitizer)
	<i>jà màsé márdù</i>	'we bought millet'	<i>jà màs-ée márdù</i>	'we bought (and brought) millet'
	<i>dà máshí</i>	'that he buy'	<i>dá màs-àì</i>	'that he buy and bring'
	<i>zèdàpù</i>	'alighting'	<i>zèdàp-àyí</i>	'alighting here'

This -AY allomorph is almost certainly a relex of Proto-Chadic *(a)wa . The -EN allomorph, to be described in section 3.3.1, can be related to the Proto-Chadic Destinitive. Thus what synchronically are suppletive allomorphs of a single extension are the historical result of a merger in Ngizim/Bade of two originally distinct extensions.

2.3.4. Gisiga. In Gisiga [Lukas 1970], the Distant extension (= "Entfernungsmorphem") is indicated by a suffix -awa (tone not indicated). In non-final position, *awa* → *o(o)* , e.g.

⁷The distinction between ventive-causatives and simple ventives followed by the sociative preposition *dà* 'with' is generally neutralized in surface structure. It does, however, show up in the continuous tenses, e.g. *yànaa fítóo dà káayáa* 'he is bringing out the loads' (causative) vs. *yànaa fítódwáa dà káayáa* 'he is coming out with the loads' (sociative). Arguments for the interpretation of the *dà* one finds in Hausa causatives as a genuine transitizer/causative marker, historically distinct from the preposition *dà* , are presented in Newman [1971].

⁸The suffix -w is a perfective marker used in certain positions. For details of tense/aspect formation in Ngizim, see Schuh [1971].

- (6) m-awa 'return here' < me 'return'
 a m-oo le 'he has returned here'
 ngəl-awa 'pick at a distance' < ngal 'pick'⁹
 a zəl-oo gawla pal 'he called a servant here' < zal 'call'

The Distant marker occurs before any of the other extensions but after the morpheme *ak* that indicates plurality of the subject, e.g.

- (7) b-o-de 'take it here' (vb-Dst-Causative)
 a b-ak-awa 'they came out' (subject vb-pl-Dst)
 a s-ak-o-de 'they brought it' (subj. vb-pl-Dst-Caus)

2.3.5. Bachama. In Bachama [Carnochan 1970], the Distant extension (= "adessive") is formed by a suffix *-á(a)* plus some associated internal vowel changes not accounted for in Carnochan's description. The suffix is equally used with simple stems and with plural stems (inflected forms generally agreeing in number with the object of transitive verbs or the subject of intransitive verbs), e.g.

- | | | | | |
|-----|------------------|------------------------|-------------------|-------------------------------|
| (8) | nda dùmè | 'he went out' | nda dùmá | 'he came out' |
| | lèrì | 'hurry and go' | làrá | 'hurry and come' |
| | nda ngəl sàlàkey | 'he pulled the rope' | nda ngàlá sàlàkey | 'he pulled the rope this way' |
| | ngga 'úsè | 'she cooked' | ngga 'úsá | 'she cooked and came' |
| | taa 'ósè | 'they cooked' | taa 'wásá | 'they cooked and came' |
| | nda pyér vonye | 'he thatched the huts' | nda pyáará vonye | 'he thatched the huts there' |

When co-occurring with other extensions, such as the Causative *dè* or the Deprivative *ge*, the Distant is always the first added to the verb, e.g.

- (9) ngàl-áa-dè 'cause to pull this way'
 dàw-á-ge 'go there and cut down without person's knowledge'

⁹The change in the verb stem vowel is due to a more general phonological alternation in Gisiga between /a/ and /ə/, and is not a specific property of the Distant extension.

2.3.6. Tera. In Tera [P. Newman 1970], the Distant extension is marked not by a suffix, as is usually the case, but by a preverbal marker á,¹⁰ which phonologically becomes attached to preceding and/or following pronoun/tense markers, e.g.

- | | | | | |
|------|-----------------------------|---------------------------|-------------------|---------------------------------|
| (10) | késu | 'catch it' | á késu | 'catch it there' |
| | tá-ŋ vi né dà̃m | 'and I went out' | tá-ŋ-á vi né dà̃m | 'and I came out' |
| | /sequential-I/leave/to/out/ | | | |
| | nùn masá nda bá | 'you (pl.) didn't buy it' | nùn-á masá nda bá | 'you (pl.) didn't buy it there' |
| | wà kaḅ gàwà | 'he planted corn' | á-à kaḅ gàwà | 'he planted corn there' |

In Ga'anda [R. Newman 1971], a language belonging to the same group as Tera, the difference between Distant and neutral verbs has been reduced to a tonal distinction, Hi vs. non-Hi on the first syllable of the verb, e.g.

- | | | | | |
|------|----------|-----------------------|----------|------------------------------|
| (11) | yimu sé | 'go into' | yímu sé | 'come into' |
| | tèrù xár | 'take some and leave' | térù xár | 'fetch some there and bring' |

Since Ga'anda and Tera are quite closely related, we can assume that the Hitone and the á are cognate forms of the same Distant morpheme. Whether they are in fact reflexes of Proto-Chadic *(a)wa is a more difficult assertion to prove, although given the match with the Bachama Hi tone marker -á(a), the historical connection is not unlikely in spite of the change in position from suffix to prefix.

3. DESTINATIVE *in

3.1. Function. The primary meaning of this extension was to indicate that the action of a verb was destined for, done for the benefit of, or otherwise affected or pertained to someone. It was probably most commonly used in sentences containing an indirect object. Originally, the notion "destination of action", as expressed by this extension, would have been "destination"

¹⁰Tera also has a Distant marker dá. This, however, is probably made up of the verb dē 'go' + á, i.e. .té dā gwà 'and he found it there' < *té dē á gwà.

in a purposive rather than a locative sense. The Destinitive would thus not have been used with motion verbs to indicate simple action in the direction of the speaker, for which one would have used the Distant extension. Later, particularly in the West Chadic branch, the Destinitive did pick up "Ventive" meanings and functions, with the result that in many languages the originally clearcut distinction between the Destinitive and the Distant was lost.

3.2. Position. The Destinitive was presumably a bound suffix, although less closely bound to the verb than the Distant marker. It was probably more bound and positionally closer to the verb than extensions such as the Completive and the Totality (and the Transitizer/Causative (?)), but this is still to be established.

3.3. Reconstruction. Reconstruction of the Destinitive *in is supported by evidence from the following languages, belonging to the West and Biu-Mandara branches of the family.

(12)

West Branch	Biu-Mandara Branch
Ngizim -én/-ìiná	Ga'anda ín
Bole -N	Daba -eŋ
Pero -ínà	Margi -əri

3.3.1. Ngizim. In Ngizim, the extension labelled "Ventive" by Schuh [1972] indicates both action in the direction of and for the benefit of a person. The extension has two grammatically conditioned suppletive allomorphs: -AY, the presumed reflex of the Proto-Chadic DISTANT extension (see section 2.3.3.), and -EN, the presumed reflex of the Proto-Chadic Destinitive *in. Although originally derived from two different extensions, the two allomorphs now equally indicate the full range of meanings of the combined extension. The -EN allomorph itself has two grammatically conditioned forms: -én in the singular imperative, and -ìiná in the plural imperative and the 2nd subjunctive, e.g.

(13) á ràurí 'call'	á ràur-én 'call here'
à wání 'send'	á wàn-én 'send here'
à wána 'send (pl)'	á wàn-ìiná 'send (pl) here'
wà màsá 'let's buy it'	wà màs-ìiná 'let's buy (and bring) it'

When co-occurring with the Transitizer extension, the Ventive (whether marked by -EN or -AY) is attached to the verb first, e.g.

- (14) dà kàlàkt-ìiná-dù 'they should return it here' < kàlakte 'return'
á vèr-én-dù 'bring it out' < vèrə 'go out'
cf. vèr-ée-dù 'he brought it out'

3.3.2. Bole. Bole [Lukas 1971, P. Newman field notes¹¹] has one extension, which I shall call the "Ventive", that covers the range of distant/ventive/destinative/benefactive meanings. This extension has three suppletive allomorphs: kó used in the continuous, tú used in the subjunctive and imperative, and -N used in the perfective, e.g.

- (15) 'ísìη màa-kó 'he will be coming back'
dóolè kà mée-tú (= mét-tú) 'you must come back'
'ísì mé-η gò¹² 'he came back'
cf. 'ísì máa wò 'he went back'

The -N is realized as a homorganic nasal and, with a-final verbs, is accompanied by the change of preceding a's to e. If the verb has a pronoun object, the suffix -N appears twice, once immediately attached to the verb and once after the pronoun, e.g.

- (16) 'ísì bèsé-m Béliò 'he shot Bello (there)'
'ísì bèsé-n-nì-η gó 'he shot him (there)'
cf. 'ísì básaá nì wó 'he shot him'

¹¹Examples are taken from my notes on the Gombe dialect. They conform very closely to Lukas's description (of the Fika dialect) except for minor details. Interestingly, Lukas uses exactly the same term ("*Entfernungsmorphem*") for this Bole extension, derived from the proto Destinative, as for the -awa extension in Gisiga (see section 2.3.4), which is derived from the proto Distant extension.

¹²The perfective/completive marker has the form wo after a vowel, go after a nasal consonant, and ko when abutting with k. I am not clear about the rules governing the tone and vowel length.

(16) cont.

'ísì wùndí-η gòò 'í ¹³	'he called (to come here)'
cf. 'ísì wùndú wòò 'í	'he called'
'ísì wùndí-η-ká-η gò	'he called you (to come here)'
cf. 'ísì wùndí káa wó	'he called you'

In forming verb stems with more than one extension, Bole adds the Transitizer first, then the Ventive, and then the Totality (which behaves like a pronoun as far as the repetition of the -N marker is concerned), e.g.

(17) 'ísì yàw-tí-η gòò 'í	'he brought it down this way' < yàwwú 'get down' (vb-Trans-Vnt)
'ísì ngòrí-n-tù-m bàayà	'he tied up all the slaves there' < ngóru 'tie' (vb-Vnt-Tot-Vnt)
'ísì bòi-tí-n-tù-n dóbó	'he broke all the pots there' < bólu 'be broken' (vb-Trans-Vnt-Tot-Vnt)

The historical sources of the kó and tú allomorphs of this extension have not been identified. The -N allomorph is presumably a reflex of the Destinative *in. The destinative/benefactive character of this extension in Bole shows up in the fact that in the perfective (where -N occurs), the extension is obligatory when indirect object pronouns are used, e.g.

(18) 'ísì bèsé-n-nì-η kòsúm	'he shot a rat for him'
'ísì sùrrí-n-tá-n lò	'he fried meat for her'

In sentences such as the above with an indirect object, it is not possible to make a distinction between a neutral verb and one with a spatial/directional meaning since the use of the -N is obligatory.

3.3.3. Pero. In Pero,¹⁴ a language belonging to the same larger group as

¹³In my notes, but not in Lukas's description, final u changes to i: before -N, e.g. my wùndíη = L's 'ùndúη. This is probably due to a more general process of final vowel weakening in Bole and is not a specific property of the -N morpheme.

¹⁴My information on Pero is drawn almost entirely from unpublished field materials of Zygmunt Frajzyngier, to whom I am most grateful. In the examples, I have taken the liberty to modify somewhat the transcription system used in Frajzyngier's notes.

Bole, but to a different subgroup, the Destinative has also been preserved as one of the suppletive allomorphs of the "Ventive" extension, a form -ínà.¹⁵ As with Bole -N, this -ínà form is limited to verbs in the perfective, the other allomorph -tù occurring in all other tenses. In Pero, however, the close association of the -ínà allomorph and the perfective tense has been carried further, to the extent that -ínà now functions as a combined "ventive-perfective" marker in contrast to the "neutral-perfective" marker kò, with which it may not co-occur. This development from derivational extension to tense/aspect marker, still in an incipient stage in Pero, is interesting since it illustrates a process that has been repeated independently a number of times in Chadic, especially with the Completive and Totality extensions. For example,

- | | | | | |
|------|------------|----------------------|------------|--------------------------------|
| (19) | léé kò | 'he cut it' | léé-nà | 'he cut it (there)' |
| | lá | 'cut' | lí-tù | 'cut it (there)' |
| | pétò kò | 'he went out' | pét-ínà | 'he came out' |
| | nì ádò kò | cùrà 'I ate peanuts' | nì ád-ínà | cùrà 'I ate peanuts (there)' |
| | nì pécò kò | tùrùmì | nì péc-ínà | tùrùmì |
| | | 'I shot a lion' | | 'I shot a lion (there)' |
| | cf. ndà | pícu-tù | tùrùmì | 'I shall shoot a lion (there)' |

3.3.4. Ga'anda. In Ga'anda [R. Newman 1971], the presumed reflex of the Destinative extension is a particle ín that indicates action directed to or for a person. While ín sometimes indicates action "towards the speaker", it is not essentially a locative/spatial extension and thus it can be functionally and semantically distinguished from the Hi tone Distant marker (section 2.3.6) with which it may co-occur. The marker ín is a free adverbial-like particle that occurs late in the sentence,

¹⁵Pero also has a suffix -n that in some of its uses could be characterized as benefactive, thereby suggesting a comparison with the Destinative. While this comparison seems reasonable at first sight, I would suggest that this -n is historically unrelated to the Destinative, rather being derived from an old transitizer extension or, in other cases, from a frozen third person masculine singular pleonastic indirect object pronoun.

separated from the verb by the subject, indirect object pronoun, direct object, etc., e.g.

- (20) è hlán-úcé-ndá wànbəbá ín 'they sent medicine to you'
 /past/send/you/they/medicine/ín/
 è yàrk-àn-ndá òířshà ín ì kútìrá 'they stole and brought the horse
 to the chief'
 /past/steal/him/they/horse/ín/to/chief/

(Note: -an is a pleonastic i.o. pronoun co-occurring with the noun i.o., a grammatical phenomenon extremely common in Chadic.)

kè sàr-án hlìwdí ín 'you should fry this meat for him'
 (i.e. and give to him)

cf. kè sàr-án hlìwdí sé 'you should fry this meat for him'
 (i.e. on his behalf)

In Ga'anda, indirect object is a functional category. The semantic reading (or underlying case relationship) of the indirect object in relation to the verb depends to a great extent on the choice of the extensional particle, whether ín 'to', sé 'for', fá 'at', or xár 'from'.

From phonological and semantic points of view, the identification of Ga'anda ín with the Proto-Chadic Destinitive *in could hardly be better. The only doubt concerns the equation of a free particle with what must originally have been a bound suffix. However, given the adjustments that Ga'anda presumably underwent in its shift from an SVO to a VSO language, the forced move of the extension out of the verb stem and its change in status to a free particle is not so difficult to imagine.

3.3.5. Daba. In Daba, the extension -eŋ, called "destinatif" by Mouchet [1966]--from whom I have taken the term--has preserved essentially the form and function of the original extension. According to Mouchet, the Destinitive in Daba is no longer productive--this being a good indication of its archaic nature--having been found with less than ten verbs, e.g. (tone not indicated):

- (21) mitis a hem-eŋ ta 'hunger troubled them' < ham 'to itch'
 ə njə ke ced-eŋ ta ma 'he spoke to them'
 tæv ku wud-eŋ ni 'I shall pay him'
 kat hi ami vəl-eŋ ni 'I gave it to him'

3.3.6. Margi. Margi has a productive extensional suffix $-əri$ (with copy tone) indicating "that action is done for somebody or to somebody ... [and] also to indicate direction of the action 'to', 'towards' ..." [Hoffmann 1963:143]. (In phonologically conditioned environments, the initial and/or final vowel(s) of the suffix are dropped.) Interestingly, when the extension is used, a suffixed object pronoun (which is formally unmarked for case) is generally interpreted as an indirect object; otherwise it is understood to be a direct object, e.g.

(22)	ɓàl	'break'	ɓàl-əri	'break for person'
	nə	'say'	nə-rì	'tell a person'
	tìdè	'creep'	tìdè-rì	'creep up to a person'
	kəhìlè-dá	'chase me'	kəhìlè-r-dá	'chase to me'
	ɲà-mèr	'call us'	ɲà-r-mèr	'call for us'

Since Margi /r/ is a regular reflex of Chadic $*n$ [Newman and Ma 1966:227], and since the Margi suffix so closely matches the presumed original meaning and function of the Proto-Chadic Destivative, the interpretation of $-əri$ as being derived from $*in$ seems justified in spite of Hoffmann's earlier suggestion [1963:143] that Margi $-əri$ be compared with a somewhat similar Bura suffix $-tà$.

4. The Destivative in Hausa

Taken as a whole, the evidence presented above points unmistakably to the existence in Proto-Chadic of two semantically related but clearly distinct verbal extensions: a Distant $*(a)wa$ and a Destivative $*in$. One of these, the Distant, has a readily identifiable reflex in Hausa (see section 2.3.2). By contrast, the uncovering of the Destivative in Hausa requires delving into a puzzling problem of Hausa grammar, to which we now turn.

For most "grades" of Hausa verbs (see [Parsons 1960/61]), the pre-dative form (henceforth the D-form), i.e. the form of the verb occurring before the markers $wà$ (+ noun) and $má$ (+ pronoun),¹⁶ is the same as

¹⁶In normal speech, the $má$ usually assimilates to the vowel of the following pronoun, e.g. $mákà$ 'to you (masc.)', $míkì$ 'to you (fem.)', $múkù$ 'to you (pl.)'. In northern and western dialects of Hausa, the dative marker before nouns is $mà$ rather than $wà$.

the citation form,¹⁷ e.g.

(23)	Gr. 1	yáa káamàa 'he caught it'	yáa káamàa wà sárkíi 'he caught it for the chief'
	Gr. 4	yáa rufèe 'he closed it'	yáa rufèe mínì 'he closed it for me'
	Gr. 5	yáa sáyáĩ 'he sold it'	yáa sáyáĩ másà 'he sold it to him'
	Gr. 6	yáa káawóo 'he brought it'	yáa káawóo mànà 'he brought it to us'

With grades 2, 3, and 7, however (plus some "irregular" verbs), the D-forms are invariably different from the underlying forms, e.g.¹⁸

(24)	Gr. 2	nèemí 'seek'	yáa néemàa másà or yáa néemám másà 'he sought it for him'
	Gr. 2	zàabí 'choose'	náa zàabàa wà sárkíi or náa zàabáĩ wà sárkíi 'I chose it for the chief'
	Gr. 3	dìrá 'swoop down'	táa díram másà 'she pounced on it'
	Gr. 3	túubá 'repent'	yáa túubám mínì 'he apologized to me'
	Gr. 7	'àukú 'happen'	yáa 'áukám mátà 'it befell her'

In most Hausa manuals, the facts regarding the unusual D-forms of these verbs are simply described as such without any attempt at explanation, e.g. "They [Gr. 2 verbs, etc.] may either become high-high(-high) and suffix -r (which assimilates to -m) ... or they may become high-low (-high) ..." [Kraft and Kirk-Greene 1973:78]. The only in-depth grammatical study of Hausa in which a serious attempt is made to account for these D-forms is by Parsons [1971/72]. In Parsons' analysis, the D-forms used by Gr. 2/3/7 verbs are forms that have been "borrowed" from other grades,

¹⁷The differences in final vowel length that exist (depending on what one takes to be the underlying/citation form) can be ignored for purposes of this discussion.

¹⁸The presentation of the examples here, specifically the adoption of the citation form with final -i for Gr. 2 and the labelling of Hi-Hi intransitive verbs as Gr. 3, is in accordance with the analysis in Newman [1973].

i.e., *néemàa* (D-form of *nèemí*) would have been borrowed from Gr. 1 (cf. *káamàa*) while *záábáǎ/záábám* (D-form of *záábí*) would have been borrowed from Gr. 5, the "causative grade" (cf. *sáyáǎ*). Thus for Parsons, it is not the forms themselves of the pre-dative verbs of the *néemàa* or *záábáǎ/záábám* patterns that are irregular, but only their usage. In Newman [1973], I adopted Parsons' analysis, as far as the borrowed Gr. 1's were concerned, but suggested that it was better to describe these D-forms in terms of switching *grades* rather than of borrowing *grade forms*. In other words, the D-form *néemàa* is not merely a Gr. 1 form, it *is* Gr. 1, and the change from *nèemí* to *néemàa* is as much a change in grade as, for example, the change to the Gr. 6 *néemóo*. The explanation I offered for the shift to Gr. 1 was that these Gr. 1's were not really semantically neutral basic verbs, but rather were derived stems containing an applicative extension, whose phonological identity with basic Gr. 1's was due to accidental homonymy.¹⁹ In any case, whether described in Parsons' original terms or according to my approach, the interpretation of the D-forms of the *néemàa* pattern as Gr. 1's seems correct and uncontroversial.²⁰

4.1. Are ǎ/m D-forms "borrowed" grade 5's? The problem concerns the analysis of the Hi-Hi consonant-final D-forms (*záábáǎ*, *néemám*, *'áukám*, etc.), that I shall henceforth refer to as the ǎ/m D-forms.²¹

¹⁹The existence of this applicative extension hidden within Gr. 1 verbs was first proposed (with some doubt and hesitation) in Newman [1973]. The question only a few years later is not whether some Gr. 1's contain an applicative extension--which now seems certain--but whether there are any Gr. 1 verbs that etymologically are *not* derived. For a description of the historical/comparative background against which Hausa phenomena must be viewed, see the excellent study by Schuh [1977].

²⁰The analysis that I have suggested of switching to Gr. 1 was taken for granted by Abraham [1959:29] and, following him, by Pilszczikowa [1969:20].

²¹The ǎ with the tilde represents a roll or tap, that in Hausa contrasts with a flap *r*, written without a diacritic. The distinction between the two, commonly (but mistakenly) ignored by Hausaists as being synchronically insignificant, is of great importance historically, as, for example, in treating the problem at hand.

Consistent with his desire to treat essentially everything in the Hausa verbal system within the confines of his seven grades, Parsons proposed that these D-forms be analyzed as forms borrowed from Gr. 5, the causative grade.²² In equating these ř/m D-forms (his "borrowed 5D forms") with the true Gr. 5's, Parsons was nevertheless aware that the forms were not entirely identical. First, as Parsons himself noted [1971/72:74n., 80n.], the supposedly underlying final -ř of the putative borrowed 5D forms almost always assimilates to the *má-* dative marker, whereas the assimilation is much less general in the case of the real Gr. 5's, i.e. one normally finds *záábám másà* (borrowed 5D) but *kóoyář másà* 'teach him' (true 5D), both in everyday pronunciation and in writing.²³ Secondly, while some speakers of some dialects still have -s as the final consonant of the Gr. 5 causative, i.e. *kóoyář* = *kóoyás* (s being the historically older consonant from which the more usual ř is derived), this -s is never found in the supposedly borrowed 5D forms. Parsons [1971/72:203] argues that the reason for this is that the dialects that use -s in the causative are the same dialects whose dative markers are *má-/mà* (instead of *má-/wà*), and thus the final -s that one might expect to find on borrowed 5D forms can never show up since the sequence s-m is ruled out on phonological grounds. However, when one looks into earlier works on Hausa, one does find examples of -s followed by the dative markers *má-/mà* but only in the case of true causatives, e.g. (tone not indicated):

- (25) *ina bayas ma talaka kurdi* (< Mischlich [1906:50])
 'I am giving money to the poor'
sayas mini da doki (< Mischlich [1906:44])
 'sell me a horse'

²²The suggestion had actually been made earlier by Pilszczikowa [1969:20]: "It seems ... that the -R form [of grade 2] is just a variant of the causative form (Grade 5) before Indirect Object."

²³The fact that the distinction between *neemam masa* and *kooyar masa* is consistently made in written works by Hausa people themselves (e.g. Abubakar Imam's *Magana Jari Ce*, Parsons' "Bible" on Hausa linguistic matters) indicates that as far as the linguistic consciousness of Hausa people is concerned, the true causatives and the "borrowed 5D forms" cannot be equated.

(25) cont.

- ka gayaz mini da gidanka (< Taylor [1923:77])²⁴
 'greet your family for me'
 aka seyas musu da ita (< Taylor [1923:77])²⁴
 'one sold it to them'

Finally, there doesn't seem to be any explanation as to why a Gr. 2 or Gr. 3 verb should "borrow" a highly marked causative form to use in pre-dative position, or why this borrowed 5D form should thereby lose all of its original semantic attributes.²⁵ Consider, for example, the following pairs of sentences, the first from Parsons [1971/72:189], the second, drawn from two different dialects, from Gouffé [1962:196].

- (26) sái kà káwám mínì kàdǎn < Gr. 3 kàwá = káu 'to move'
 'just move away from me a bit'
 sái kà káwáǎr mínì dà shíi kàdǎn < Gr. 5 káwáǎr 'move something'
 'just move it away for/from me a bit'
 náa sáyám māsà dóokìi (Agades) < Gr. 2 sáyí 'buy'
 'I bought him a horse'
 náa sáyáǎr māsà dà dóokìi (Kano) < Gr. 5 sáyáǎr 'sell'
 'I sold him a horse'

Having expressed his own doubts at various points about the proposed interpretation of the ř/m D-forms as Gr. 5 verbs, Parsons [1971/72:205] concluded on a strong note: "I see no reason to make any systematic separation (on a synchronic analysis) among the level-toned forms of the verb with a final consonant, -s/-r/-d/-m, diverse as their function, and somewhat erratic as their contextual and dialectal distribution is ... I call them all grade 5 forms ..." It is clear to me, however, that the mass of the evidence points in the opposite direction, and that there is in fact little reason to relate the ř/m D-forms to the causatives. I would argue, rather, that the theory of "borrowed grade 5's" should be

²⁴In a later edition [Taylor 1959], these forms were "corrected" to gayam mini and sayar musu.

²⁵The switch to Gr. 1, on the other hand, is reasonable in this respect since the formal change is accompanied by the addition of the applicative sense proper to that grade (see Newman [1973]).

rejected and that an alternative explanation be sought for these special pre-dative forms.

4.2. Evidence for the Destinative origin of the ř/m D-form. In treating the alternation found in *néemář wà* (+ noun) vs. *néemám má* (+ pronoun), all previous analyses of Hausa have assumed that the ř was basic and that the m was due to a low-level assimilation rule (e.g. Abraham [1959:28]; Kraft and Kirk-Greene [1973:78]). If, however, one broke with tradition and took -n as the basic form of the suffix (with automatic assimilation to m before ma) and ř as the secondary variant (resulting from recent historical developments, partly phonological, partly analogical), then the identification of the suffix as a reflex of the widespread Destinative *in would become an immediate possibility. From the point of view of function and meaning, the analysis of these grammatical pre-dative forms as being derived verbs containing an extension fits nicely with what one finds elsewhere in Chadic, cf. Bole, for example, where the use of the Destinative/Ventive extension -N is obligatory before indirect object pronouns (see section 3.3.2).

Given this analysis, the semantic contrast sometimes found between borrowed Gr. 1 forms and borrowed Gr. 5 forms--inexplicable in Parsons' analysis--could then be understood in terms of the difference between two derivational extensions, the Applicative and the Destinative,²⁶ e.g. (examples from Parsons 1971/72:81):

- (27) *yá sóokàa mínì wúkàa* 'he stuck a knife into me' (Applicative)
yá sóokám mínì ràakúmíi 'he stabbed my camel' (Destinative)
yá góogàa másà mâi 'he rubbed polish on it' (Applicative)
káa góogám mínì káfàa (táa) 'you brushed against my leg' (Destinative)²⁷

²⁶Compare Ga'anda (section 3.3.4), where the meaning of an indirect object depends very clearly on the choice of one or another extension. If one assumes for Hausa that, in appropriate contexts, non-extended, non-derived verbs could have been used in pre-dative position, then forms such as **sòokí(i) másà* should have existed alongside *sóokàa másà* and *sóokám másà* --as they still do! (See the extremely important but neglected description of pre-dative forms of this type in Pilszczikowa [1969:20-22]).

²⁷With some verbs, the Applicative and Destinative have merged, e.g. *náa záabàa/záabám másà dóokìi* 'I chose him a horse'; *táa háifaa/háifám másà 'yaa'yáa 'úkù* 'she bore him three children'. This does not, however, lessen the importance of distinguishing between the two extensions to account for the cases where they do contrast.

The above discussion only proves that the interpretation of these D-forms as containing a derivational extension is plausible, both internally and comparatively. What is of course still needed before one could really be justified in relating the Hausa forms to the Destinitive *in is some evidence that the final consonant of these ř/m D-forms was in fact originally -n . Such evidence does exist.

The first fact, noted earlier, is that the final consonant of these D-forms is almost always pronounced m before the dative marker má , i.e. néemám másà , not néemář másà . While there is nothing unusual in Hausa about the assimilation of ř to m in the environment of m , it is far from an automatic rule, as evidenced not only by pre-dative causatives such as sáyář másà , but also by non-assimilated sequences within a word, e.g. gāřmáa 'large hoe', fāřmákíi 'sudden attack'. The difference between néemám másà (D-form of Gr. 2) and sáyář másà (D-form of Gr. 5) argues for the essential distinctiveness of the two final consonants, -n in the first place, and -ř (< *s) in the second.

Second, Hausa has a small number of fixed (and presumably archaic) verbal expressions made up of a verb plus dative marker, e.g. táasám-mà 'approach, attack', cím-mà 'overtake', 'ím-mà 'control', where the form of the dative marker used before i.o. nouns is mà in all dialects, including those where wà is the usual pre-noun marker, e.g. yáa néemář wà màatářsà 'he sought for his wife', but yáa táasám-mà màatářsà 'he attacked his wife' (< táashì 'get up'). If one assumes that -ř is the original form of the final consonant and that wà rather than mà is the older form of the dative marker (as is generally done),²⁸ there is no way to account for the transition from *táasář wà to táasám-mà . If, however, the final consonant of the verb stem were originally a nasal, then the transition from *táasán wà to táasám-mà²⁹ would be explicable

²⁸Contrary to the usual view that wà is original and that mà is due to analogic levelling under pressure from the pre-pronoun forms másà , mátà , etc., Eulenberg [1972:33-36] has suggested that mà is original and that the change from mà to wà was a result of a general process of lenition affecting the standard Hausa dialect. My guess is that the traditional view is correct, although Eulenberg's proposal cannot be rejected out of hand.

²⁹Assimilation in Hausa is usually regressive; but for the change postulated here, note gánwóo (W. Hausa) > gámmóo (Standard Hausa) 'head pad'.

in terms of a simple assimilation of wà to mà accompanying the fusion of the verb root and the dative marker into what could be considered a new lexeme.

Finally, direct confirmation of the claim that the final consonant was originally -n is provided by examples, previously overlooked in the literature, of ř/m D-forms where the final consonant is -n even though the adjacent dative marker is not ma (which is generally taken as as the necessary conditioning for the appearance of the nasal) but wà ! Note the following remarkable examples cited in Taylor [1959:101], who explicitly describes these verbs as taking an "n form" (tone not indicated):

- (28) gudun wa 'run away from'
 zaaburan wa 'spring upon'
 sanan wa 'inform' (cf. the causative sánád dà)
 taasan wa 'approach' (cf. the more usual táasám-mà)

Taking all the evidence into account, I would conclude that the original final consonant of the Hi-Hi, ř/m pre-dative forms was almost certainly -n . While the circumstances under which -ř developed remain unclear, there is little doubt but that the direction of change was from n to ř and not ř to n.³⁰ Having reconstructed the phonological shape of these pre-dative forms with a final -n , I would suggest that there is no reason not to interpret this -n etymologically as an extensional suffix deriving from the Proto-Chadic Destivative *in . Contrary to the Distant/Ventive extension, which still retains its extensional qualities in Hausa in the form of grade 6, the Destivative has been grammaticalized to the point where its original character as an optional derivational suffix with its own semantic attributes has essentially been lost. Nevertheless, the recognition of the ř/m D-forms as Destinatives not only solves this particularly perplexing

³⁰In any case ř as such could not be basic since it, as opposed to the flap r , is a historically secondary consonant of fairly recent vintage, having developed in syllable final position out of other alveolar consonants.

puzzle in Hausa grammar, but it is certain to prove important in our understanding of other synchronic and diachronic problems in Hausa as well.

5. Conclusion

Among the verbal extensions that existed in Proto-Chadic, two are reconstructed in this paper: a Distant extension *(a)wa and a Destinative extension *in . The former served to place the action of a verb in space at some distance from the speaker or in the direction of the speaker; the latter served to relate a person to the action as the destination, beneficiary, or otherwise affected party. In the course of Chadic linguistic history, these two extensions have encroached on each other's territory, in some cases the result being the demise of one or the other extension, in other cases the result being the merger of the two extensions and the reassignment of the originally distinct suffixes as allomorphs of a single morpheme. In Hausa, the reflex of the Distant extension is still visible as the -oo/-woo Ventive ending on Grade 6 verbs. The Destinative, which at first sight would seem to have been lost in Hausa, was shown to have in fact been preserved, not, however, still functioning as an extension, but buried deep in the grammatical system as an inflectional ending on verbs in pre-dative position.

REFERENCES

- Abraham, R.C. 1959. *The Language of the Hausa People*. London: University of London Press.
- Carnochan, Jack. 1970. "Categories of the verbal piece in Bachama." *African Language Studies* 11:81-112.
- Eulenberg, John B. 1972. "Complementation phenomena in Hausa." Doctoral dissertation, University of California at San Diego.
- Gouffé, Claude. 1962. "Observations sur le degré causatif dans un parler haoussa du Niger." *Journal of African Languages* 1:182-200.
- Hoffmann, Carl. 1963. *A Grammar of the Margi Language*. London: Oxford University Press.

- Jungraithmayr, Herrmann. 1970. *Die Ron-Sprachen: Tschadohamitische Studien in Nordnigerien*. Afrikanistische Forschungen 3. Glückstadt: J.J. Augustin.
- Kraft, C.H., and A.H.M. Kirk-Greene. 1973. *Hausa*. Teach Yourself Books. London: The English Universities Press.
- Lukas, Johannes. 1970. *Studien zur Sprache der Gisiga (Nordkamerun)*. Afrikanistische Forschungen 4. Glückstadt: J.J. Augustin.
- Lukas, Johannes. 1971. "Über das erweiterte Verb im Bolanci (Nordnigerien)." In Paul Newman (ed.), *Special Chadic Issue*, pp. 1-14. *Journal of African Languages* 10(1).
- Mischlich, A. 1906. *Wörterbuch der Hausasprache*. Berlin: Reimer.
- Mouchet, J. 1966. *Le parler daba*. Recherches et Études Camerounaises 10. Yaoundé.
- Newman, Paul. 1970. *A Grammar of Tera: Transformational Syntax and Texts*. University of California Publications in Linguistics 57. Berkeley and Los Angeles: University of California Press.
- Newman, Paul. 1971. "Transitive and intransitive in Chadic languages." In Veronika Six et al. (eds.), *Afrikanische Sprachen und Kulturen-- Ein Querschnitt* (Festschrift J. Lukas), pp. 188-200. Hamburger Beiträge zur Afrika-Kunde 14. Hamburg.
- Newman, Paul. 1973. "Grades, vowel-tone classes and extensions in the Hausa verbal system." *Studies in African Linguistics* 4:297-346.
- Newman, Paul. 1977. "Chadic classification and reconstructions." *Afro-asiatic Linguistics* 5(1):1-42.
- Newman, Paul, and Roxana Ma. 1966. "Comparative Chadic: phonology and lexicon." *Journal of African Languages* 5:218-51.
- Newman, Roxana Ma. 1971. "A case grammar of Ga'anda." Doctoral dissertation, University of California at Los Angeles.
- Parsons, F.W. 1960/61. "The verbal system in Hausa." *Afrika und Übersee* 44:1-36.
- Parsons, F.W. 1971/72. "Suppletion and neutralization in the verbal system of Hausa." *Afrika und Übersee* 55:49-97, 188-208.
- Pilszczikowa, Nina. 1969. *The Changing Form (Grade 2) of the Verb in Hausa*. Warsaw: Państwowe Wydawnictwo Naukowe.
- Schuh, Russell G. 1971. "Verb forms and verb aspects in Ngizim." In Paul Newman (ed.), *Special Chadic Issue*, pp. 47-60. *Journal of African Languages* 10(1).
- Schuh, Russell G. 1972. "Aspects of Ngizim syntax." Doctoral dissertation, University of California at Los Angeles.
- Schuh, Russell G. 1977. "West Chadic verb classes." In Paul Newman and Roxana Ma Newman (eds.), *Papers in Chadic Linguistics* pp. 143-66. Leiden: Afrika-Studiecentrum.

Taylor, F.W. 1923. *A Practical Hausa Grammar*. London: Oxford University Press.

Taylor, F.W. 1959. *A Practical Hausa Grammar*. 2nd edition. London: Oxford University Press.

THE TREATMENT OF GEMINATES:
EVIDENCE FROM BERBER*

Jilali Saib
Department of Linguistics, UCLA

This paper proposes a solution to the indeterminacy problem which has often characterized the treatment of geminate consonants. The indeterminacy consists in treating them--owing to their dual function--as underlying sequences of two-like segments with respect to some rules and as unit segments with respect to others. Harm's [1968:36] solution to the problem is found inefficient since it does not constrain the representation of geminates. Overwhelming evidence is brought forward for a sequential analysis of Berber geminates. A phonetically motivated and potentially universal convention is here presented, whereby the two elements of a geminate cluster fortify each other and as such are redundantly specified [+tense]. This explains why geminates in Berber and other languages (1) are exempt from spirantization viz. tt → θθ, (2) are not broken up by an epenthetic rule, and (3) undergo strengthening processes of their own.

0. Introduction

A longstanding problem concerning geminates is that, within a single language, they can function sometimes as sequences of two like-segments, or as unit segments. Phoneticians and phonologists alike have engaged in lengthy debates over how these segments should be represented. The

*This paper is a slightly revised version of a paper read at the Winter Meeting of the Linguistics Society of America, Dec. 1973. In writing it, I benefited from the comments and suggestions of the following scholars: Victoria Fromkin, Larry Hyman, Thomas Penchoen, Russell Schuh, and Michael Brame. Their help is gratefully acknowledged. Any errors or omissions are my sole responsibility, however. As this paper goes to press, I have received a paper on geminates by Mohammed Guerssel, entitled "Constraints on phonological rules", to appear in *Linguistic Analysis*. In his paper, Guerssel also argues for a sequential analysis of Berber geminates, follows the same line of argumentation, and utilizes the same arguments as those advanced here.

debate among phoneticians centers around the question of whether the production of geminates involves two phases, i.e. rearticulation of the segment, or one phase.¹

Among phonologists, the debate revolves around the issue of whether geminates, when contrasted with single consonants, are to be analyzed as a sequence of two like segments, i.e. CC versus C, or as one segment with a feature specification [+long] or [+tense].² Since there is much confusion surrounding this issue, the discovery of a language where compelling arguments can be made for defending one of these approaches is of major consequence.

One such language, Berber (the data are mainly drawn from the Ayt Ndhir dialect of Tamazight) is treated at length here, where it is argued that, in order to adequately explain a number of phonological facts, geminates must be treated as sequences of two identical consonants. At the same time, however, in order to explain other facts of the language, a general convention for interpreting each of the two identical consonants in the sequences as redundantly [+tense] is introduced and defended.

In what follows, a typology of geminates in Berber is given (Section 1), followed by an outline of general approaches to the analysis of these segments (Section 2). In Section 3, evidence is presented for analyzing Berber geminates as sequences of two like consonants. Some arguments for analyzing these geminates as single [+long] or [+tense] consonants, are given in Section 4. In Section 5 the aforementioned convention is offered, and its implications discussed (Section 6).

1. Geminates in Berber

The approach to be argued in this paper applies equally well to lexical (or underlying) geminates, as seen in (1):

¹A summary of the debate and references for the two opposing viewpoints can be found in Lehiste [1973:131-32].

²Cf. Swadesh [1937], Martinet [1959:223-27], and Mitchell [1957] as well as other papers in the volume where Mitchell's paper appears.

- (1) /ass/ [ass] 'to tie' (cf. /as/ [as] 'to fit')
 /aff/ [aff] 'to be full' (cf. /af/ [af] 'to find')

as well as to geminates derived by rule, as in (2):³

- (2) a. sg. /afus/ [afus] pl. [ɪfassen] 'hand'
 sg. /afud/ [afud] pl. [ɪfadden] 'knee'
- b. /ad + d # ddu + x/ [addeddɪx] 'that I will come over here'
 FUT DIR come I
- c. /ad # t + ddu/ [atteddu] 'she will go'
 FUT she go
- (FUT stands for future, and DIR for directional particle)

In (2a) the geminate consonants in the plural forms arise through a rule of derivational morphology. In (2b), the geminate consonants are the result of two identical consonants coming together across a grammatical boundary. Finally, in (2c), the geminate consonants result from a complete assimilation of one consonant to another. The fact that certain phonetic geminates derive from a sequence of segments raises the question of whether all phonetic geminates should be similarly derived.

2. Approaches to Geminates

Two possible analyses suggest themselves. Under one analysis, which I will call the *feature analysis*, geminates would be assigned a feature [±long] or [±tense].⁴ Under the other analysis, which I will call the *sequential analysis*, the geminates are analyzed as a sequence of two identical consonants.

³The changes in the quality of the vowels are ignored here since they are not crucial to the discussion. Since the present discussion applies to Berber in general, single consonants are represented as stops, not as spirants (cf. Saib [1974]). Throughout this paper schwa is represented as e; **ɗ**, **ɓ**, **ɠ**, correspond to **ð**, **β**, **ɣ**, respectively. A dot underneath a C indicates Pharyngealization.

⁴The feature [±tense] is explicitly proposed for geminate consonants in Berber by Mitchell [1957], Galand [1975], Harries [1966], Abdel-Massih [1968], and Penchoen [1973]. Of these authors, Mitchell bases his use of the feature on some phonetic experiments, mainly with palatograms and kymographs.

Given these two possible approaches, three logical claims can be made about the nature of the underlying representation of geminates. They are given in (3) - (5):

- (3) *Strong Claim*: only one of these analyses is correct and all geminates must be treated as such in *all* languages, i.e. one of these analyses is universally disallowed.
- (4) *Modest Claim*: both analyses are available, but only one can be used in a given language.
- (5) *Weak Claim*: both analyses are available to languages and can co-occur in a single language. (This is in accordance with a convention originally proposed by Harms [1968:36].)

At the outset it is clear from the linguistic literature that the strong claim is too strong. Therefore, I will first address myself to the modest claim. Given this constraint, which of the two approaches (the feature analysis or the sequential analysis) would be correct for Berber? The claim inherent in the feature analysis is that geminate consonants function like single consonants in Berber; the claim inherent in the sequential analysis is that geminate consonants function like sequences of two consonants in Berber. I shall now present evidence that the feature analysis cannot account in a non-ad hoc way for certain aspects of Berber phonology.

3. Evidence for the Sequential Analysis

The first piece of evidence is afforded by the feminine/diminutive formation of nouns. Examples are provided in (6):

	<u>Masculine</u>		<u>Feminine/diminutive</u>			
a.	/axam/	[axam]	'tent'	/t-axam-t/	[taxamt]	'small tent'
b.	/aḥṛrat/	[aḥerræt] ⁵	'plower'	/t-aḥṛrat-t/	[taḥerræt̪t̪]	'plowing'

⁵Masculine singular nouns with t as the final radical are very rare in present day Berber dialects/languages. A t reappears in the plural of a number of nouns (e.g. arba 'boy' arbaten 'boys'), which suggests that 'boy' was probably */arbat/ to which the regular plural suffix /-n/ is added. Nonetheless, a sequence of t(#)t → [tt], e.g. /wwt#t/ 'hit him' [ewwet̪t̪].

(6) cont.

- c. /afud/ [afud] 'knee' /t-afud-t/ [tafutt] 'small knee'
d. /abuɖ/ [abuɖ] 'bottom' /t-abuɖ-t/ [tabuɖtt] 'small bottom'

In (6a) we see that feminine/diminutive formation of nouns is achieved by prefixation and suffixation of a *t*. When two *t*'s come together, as in (6b), the result is a geminate. Note that there is a phonetic contrast between *t* in the masculine and *tt* (or *t:*) in the feminine. Similarly, when non-identical dental stops come together, as in (6c), regressive assimilation takes place, and the end result is a geminate. It is clear from these examples that at least some phonetic geminates must be derived from a sequence. Otherwise the masculine and feminine form would have to be listed separately in the lexicon, which would obscure the regular relationship between them.

A second piece of evidence is provided by the data in (7), taken from the Ntifa and Zayan dialects of Tamazight (cf. Laoust [1918:129, 130]; Loubignac [1924:206]), which illustrate a process of degemination by vowel insertion:

(7) <u>Zero Form</u>	<u>Intensive Form</u>
/fzɣ/ [fezɣ] 'to chew'	/tt-fzɣ/ [ettefzɣ]
/gɣ/ [geɣ] 'to gnaw'	/tt-gɣ/ [ettegɣ]
/bdd/ [bedd] 'to stand up'	/tt-bdd/ [ettebdad]

The insertion of *a*, along with the prefixation of *tt*, is in line with the intensive form derivation of the so-called vowelless stems with a lexically geminate second radical, e.g. /C₁C₂C₂C₃/: /bddl/ 'to change' [beddeɭ]. Put differently, the verbs in (7) pattern along the lines of verbs such as /bddl/, which form their intensive form with prefixation of *tt* and insertion of *a* before the last radical consonant: /tt-bddl/ → /ettebdadɭ]. Since in the case of verbs such as those in (7) the second root consonant is a geminate, it is broken up for the purpose of *a* insertion required by the intensive form derivation.

While this type of degemination rule is only a minor rule in Berber phonology, and while it is motivated by the intensive form derivation, another motivation for it could have been, historically, the simplification of consonant clusters. In a sense, the *a* insertion might also

double up as an epenthesis rule in the intensive form of these stems. There exists in Berber a constraint against three consonants in a row (cf. Saib [1976: Chap. 3]). (For the insertion of a schwa in [ettbedda] 'to change' before the geminate tt instead of between the two t's, see the explanation given below in footnotes 7, 8 and 9).

Returning to the main discussion, we find that, not only is the [+tense] analysis unable to satisfactorily account for the insertion of /a/ between the geminate consonants in (7), but that it also cannot explain the motivation of this process; since, according to the feature analysis, a sequence of a consonant followed by a geminate would be viewed as only two consonants in a row. One could, of course, consider a sequence of a non-tense C and a tense C to function as a three consonant sequence, but this is just what a "geminate" solution is stating.

A third piece of evidence pointing to the necessity of the sequential analysis is drawn from a productive process of schwa insertion,⁶ illustrated in (8):

- (8) /bdu/ → [ebdu] 'to start'
 /gnu/ → [egnu] 'to sew'
 /rzu/ → [erzu] 'to look for'

The rule in (9) states that a schwa is inserted before two consonants followed by a vowel:

⁶For the sake of clarity, the rules of schwa epenthesis proposed here are stated so as to apply mainly to the verb stems in their simplest forms, i.e. Zero Form and Intensive Form without the affixes. Certain clusterings involving geminates and a single consonant do arise during the conjugation. However, they do not pose any problems as stated in footnote 8 below. Certain Berberists (Laoust [1918]) have claimed that in the particular dialect they were investigating, there is a clustering different from that occurring in my dialect, viz [gers] 'to slaughter' instead of [egres]. It should be pointed out that this state of affairs is understandable given that, with very few exceptions, Berber verbs have at least one sonorant consonant. It is a well known fact across languages that more clustering possibilities arise when sequences of consonants contain sonorants. For a complete investigation of the problem of schwa in Berber, see Saib [1974, 1976].

(9) $\emptyset \rightarrow e / \text{ ____ } CCV$

As seen in the examples in (10), this same rule applies when a geminate is followed by a vowel:

(10) /kku/ → [ekku] 'to mow'
/ddu/ → [eddu] 'to go'
/rru/ → [erru] 'to add'

Thus, in order for rule (9) to apply to the forms in (10), it is necessary to interpret these latter forms as beginning with *two* consonants--which in this case happen to be identical. If these geminates were interpreted as single consonants, then there would be no schwa insertion, as indicated in (11):

(11) a. [fa] 'to yawn' (no schwa insertion)
[su] 'to drink' " " "
b. *[kku] 'to mow'
*[ddu] 'to go'

Hence, if we were to view geminates as single consonants, we would expect the schwa-less forms in (11b). Since these do not occur, there is strong evidence for analyzing geminates as double consonants. One could also state the rule as (9a):

(9a) $\emptyset \rightarrow e / \text{ ____ } \left\{ \begin{array}{l} CC \\ C: \end{array} \right\} V$

But this again makes C: equivalent to CC and makes a simple rule complex by stating two environments instead of one.

There is a second rule of schwa insertion operating in Berber, which is stated in (12):

(12) $\emptyset \rightarrow e / CC \text{ ____ } C\#$ ⁷

⁷It might be pointed out that, as stated, rule (12) would not yield the correct output for the citation form of the verbs in (7) above. Indeed, surface forms such as *[efzez] 'to chew' would result, if the final radical consonant is analyzed as a geminate cluster, as is argued here. This suggests that the rule must be reformulated so as to handle these verbs. While the machinery, for example, angle brackets, is available to us to do

In verb stems consisting of three underlying consonants (and no phonological vowel), a schwa is inserted to separate the last consonant from the two preceding consonants. As seen in the derivation in (13), this rule feeds into rule (9):

- (13) Underlying form: /x_dm/
 By rule (12): xdem
 By rule (9): exdem
 Phonetic form: [exdem] 'to work'
 cf. [egmes] 'to cover'
 [eɾʒem] 'to open'

What is important is that words which consist of a geminate consonant followed by another consonant follow exactly the same two rules, as illustrated in (14):

- (14) Underlying form: /kks/
 By rule (12): kkes
 By rule (9): ekkes
 Phonetic form: [ekkes] 'to take off'
 [ellef] 'to repudiate'
 [eqqen] 'to close'

As seen in the forms in (15), when a word consists of two underlying consonants, a schwa is inserted either before or between the two consonants.

- (15) /ng/ → [eng] or [neg] 'to kill'
 /ns/ → [ens] or [nes] 'to spend the night'

Thus, if the underlying geminates in (16a) were analyzed as single consonants, we would expect to find the phonetic forms in (16b):

- | | |
|-----------------------------|-----------------------|
| (16) a. /kks/ 'to take off' | b. *[ekks] or *[kkes] |
| /llf/ 'to repudiate' | *[ellef] or *[lllef] |
| /qqn/ 'to close' | *[eqqn] or *[qqnen] |

so, we do not see any gain in unduly complicating the statement of the rule. Such forms are easily accounted for by the general constraint stated in footnote 8, below. Hence, all that would be needed is to add a condition on the schwa rule (12) saying that it cannot break up geminate consonants.

Since none of the forms in (16b) are found, (at least as citation forms, cf. footnotes 6 and 8), it is obvious that these geminates must be analyzed as sequences of two identical consonants. The one exception is when the schwa placement rule would insert a schwa between the two geminate consonants (e.g. during the conjugation of the verbs in those persons indicated by suffixes, such as -x 'I', etc.). This would, for instance, create such incorrect forms as *[keksex] (< /kks-x/ 'I took off') instead of the correct [ekkses]. A general constraint on the schwa rule will have to indicate that if a schwa would occur between the two elements of a geminate cluster, it automatically moves one place to the left, occurring before them. This is, however, *not* an argument for saying that geminate consonants should be treated as one segment. If this were to be seriously argued, it would mean that geminates would be counted as one segment just in case the schwa rule would not give the desired output, and would therefore be completely circular. Since I have given examples of where geminates must be counted as two segments for the same schwa insertion rule, this would amount to saying that we have to look ahead to the phonetic output and then decide whether we want to call a geminate consonant one or two segments. It is much sounder to start with one representation of geminates, viz. as a consonant sequence and have a condition on the schwa insertion rule which forbids the schwa from occurring between the like consonants. This condition is in itself well-motivated by the mutual effect that the two consonants have on one another, i.e. mutual strengthening.⁸

In summary, it should be clear from the above arguments that treating Berber geminates under the sequential analysis leads to a more general and explanatory account of Berber phonology.

⁸In fact, what the Berber data suggest is the following general constraint: an epenthetic rule would not normally break up a geminate cluster. The generality of this constraint is supported by the fact that it is found to be operative in a number of Arabic dialects. In Moroccan Arabic (Saib [1975], Dahbi [1975]), as is the case in Berber, the schwa insertion rules break up clusters of unlike consonants, e.g. /ktb/ → [ekteb] 'write', but not geminate clusters, e.g. /fkk/ → [fekkk] 'untie' not *[fkek]. See also Abdo's discussion of this phenomenon in the Mukabir dialect of Arabic (as reported in Brame [1971:569-70]), where an epenthetic ð can break up all clusters except geminates, e.g. y++ktb+u → [yik+tbu] 'they write' but radd → *[rad+ð] 'he returned'.

4. Evidence for the Feature Analysis

Despite the evidence given above, the feature analysis does appear to be better suited to handle other aspects of Berber phonology.

One such aspect involves the morphologically conditioned rule by which the intensive form of verbs is derived from the zero form. The examples in (17) illustrate that one way this is achieved is by geminating one consonant of the stem, in this case the second consonant:

(17)	<u>Zero Form</u>	<u>Intensive Form</u>	
	/x <u>dm</u> /	/x <u>ddm</u> /	→ [xɛddɛm] 'to work'
	/g <u>ms</u> /	/g <u>mms</u> /	→ [ɛgmɛs] 'to cover'
	/r <u>zm</u> /	/r <u>zzm</u> /	→ [rɛzzɛm] 'to open'

What is of interest here is that for those verbs which have underlying geminates in their zero forms, a prefix /tt/ is used to derive the intensive form, as seen in (18):

(18)	<u>Zero Form</u>	<u>Intensive Form</u>	
	/kks/	[ɛttɛkɛs] ⁹	'to take off'
	/llf/	[ɛttɛllɛf]	'to repudiate'
	/qqn/	[ɛttɛqqɛn]	'to close'

In order to determine how the facts in (17) and (18) can best be explained, let us examine roughly how the intensive rule is written under the two

⁹The phonetic forms in (18) are derived by rules (12) and (9). The schwa before the prefix tt is obtained by a reapplication of rule (9), since its structural description is met. In cases where the first stem consonant is non-geminate and is followed by a vowel, the intensive prefix tt reduces to t̄ in fast speech, and a schwa is then inserted before it, e.g. /tt-fafa/ 'to wake up' → [ɛtfafa]. It should be pointed out, however, that the existence of the careful pronunciation [ɛttfafa] does not constitute a problem for the analysis of schwa epenthesis proposed here. It is accounted for by the principled constraint proposed in footnote 8. Previous writers do not insert a schwa in this position, thereby implying that there exists none (Laoust [1918], Galand [1975], Abdul-Massih [1968], Penchoen [1973]). In fact Galand [1975], went as far as calling them "voyelles centrales imaginaires". However, an experiment which Jean-Marie Hombert and I ran at the UCLA Phonetics Laboratory clearly showed the existence of schwas word internally and reduced schwas word initially. I would like to thank Dr. Hombert for his help.

analyses. Under the feature analysis the data in (17) are accounted for by the informal rule in (19):

(19) . . . [-tense] . . . \rightarrow [+tense] / morphological information
C

Notice that the rule in (19) correctly predicts that in forms with underlying geminates, which are represented as single tense consonants, we will not obtain forms such as those in (20):

(20) *kekkes (from /kks/ 'to take off')
*lellef (from /llf/ 'to repudiate')
*qeqqen (from /qqn/ 'to close')

Since (19) converts a non-tense consonant to a tense one, it applies vacuously to such inputs as in the zero forms of (18).

On the other hand, under the sequential analysis the data of (17) are accounted for by the informal rule in (21):

(21) . . . C_x . . . \rightarrow C_x C_x / morphological information

In this conceptualization, the intensive form is seen to be derived from the zero form by means of a reduplication of a stem consonant of the verb. However, if (21) represents the correct approach to this derivation, then its prediction is that one of the geminate consonants should be able to undergo a second gemination as in (22):

(22) . . . C_x C_x . . . \rightarrow C_x C_x C_x / morphological information

However, as already seen in (20), this prediction is not borne out by the data.

Thus, one argument for the feature analysis is that it correctly predicts the inability of underlying geminates to further geminate in the derivation of the intensive forms of verbs. The only way a sequential analysis can be made to provide the same prediction is by adding a constraint in rule (21), reproduced now as (21'):

(21') . . . C_x . . . \rightarrow C_x C_x / Y (& morphological information)
where Y \neq C_x

The rule now states that this derivational process can reduplicate a consonant from the zero form of the verb, unless that consonant is

preceded or followed by an identical consonant (using Bach's neighborhood convention).

Although this condition accounts for the data, it should be noted that under the feature analysis no such condition is needed, since it is already built into the rule of (19) by using the feature specification [-tense] in the structural description.

A similar argument can be made concerning the phonological process of spirantization in Berber, which will now be considered from a historical perspective (cf. the discussion of spirantization in Biblical Hebrew reported by Sampson [1973]).¹⁰ In many Berber dialects, including my own, all non-geminate stops have been converted to spirants, as seen in the data in (23a):

(23a)	<u>Non-Spirantizing Dialects</u>	<u>Spirantizing Dialects</u>
a.	[tarik] 'saddle'	[θarišθ]
	[ender] 'to roar'	[ender]
	[tatbirt] 'pigeon (f.)'	[θaθbirθ]
b.	[afetta] 'couscous'	[afetta]
	[asebbab] 'trader'	[asebbab]
	[taddart] 'house'	[θaddarθ]

However, the geminate stops in (23b) are not affected by this rule in the spirantizing dialects.

Under the feature analysis we can assume the historical change in (24):

(24) [-tense] → [+cont]
 C

¹⁰The spirantization rule in Biblical Hebrew is similar to that operating in Berber except that in Biblical Hebrew a storable phonological environment for it exists. According to Sampson, it applies to single consonants post-vocalically, but not to geminates. In his squib, he argues for a feature analysis on the basis of the difficulty one would have in formulating the aforementioned spirantization rule in Biblical Hebrew. However, in a reply to Sampson, Barkai [1974:456-59], refutes Sampson's contention and convincingly argues for a sequential analysis. For a detailed study on spirantization in Berber, cf. Saib [1974, 1976].

Under the sequential analysis the historical change is stated as in (25):

(25) $C_X \rightarrow [+cont] / Y$, where $Y \neq C_X$

The condition in (25) is needed to ensure that a geminate consonant, which is here analyzed as a sequence of identical consonants, will not undergo spirantization. This would yield unacceptable forms such as those in (26):

(26) *[afeθθa] 'couscous'
 *[asebbab] 'trader'
 *[θaddarθ] 'house'

Thus, just as in the formation of the intensive form of verbs, it is the feature analysis which makes the correct prediction.

5. Conclusion

In conclusion, we can say that both analyses provide important insights into Berber phonology. Each analysis handles some aspects of phonological or morphological rules dealing with geminates, but each analysis seems inadequate with respect to other aspects. It would thus appear necessary to abandon even the modest claim made above in (4). In order to state the phonological properties of Berber, we find it necessary to refer to geminates both as sequences of identical consonants and also as [+tense] consonants. In this case, we may find it necessary to use Harms' convention which would permit us to refer to the same phonetic entity either as a single consonant (for some rules) or as a sequence of two identical consonants (for other rules) (cf. also the treatment of vowel length phenomena by Kenstowicz [1970] and Pyle [1970]). This position, represented as the weak claim in (5), would unfortunately not place any constraints on the phonological representation of geminates in languages.

There is, however, a way out which should be considered. The importance of the arguments presented for the feature analysis in the preceding discussion is that we need to refer to geminates sometimes as [+tense]. No arguments were found, however, that these geminates had to be treated as single consonants. Thus, I propose the following convention in (27):

(27) C_X C_X
 ↓ ↓
 [+tense] [+tense]

In Berber, all geminates will be analyzed as sequences of two identical consonants. However, they will be redundantly specified as [+tense]. In fact, (27) will reapply any time geminates are created. Thus, a single /t/, taken here as representing any single consonant, is [-tense], but when one /t/ abuts with another /t/, the two fortify one another and automatically become [+tense]. Since underlying /tt/ is now viewed as two consonants, both of which are specified [+tense], we can not only account for all of the data presented in favor of the sequential analysis, but also for all of the data handled by the feature analysis. Thus, the reason why geminates do not become intensified in the intensive form of verbs is that each of the geminate consonants is already redundantly specified [+tense]. Similarly, since each /t/ of /tt/ is specified [+tense], neither spirantizes to [θ].

6. Some Implications of Convention (27)

In Section 5, convention (27) is proposed to further motivate the sequential analysis argued for in the main body of the paper for Berber geminates. It is also presented to solve the almost endemic problem of indeterminacy concerning the representation of geminates faced by analysts. The question before us now is whether (27) is a language specific device, devised for Berber, or whether it has a universal application?

In the foregoing sections, we have seen, first, that geminates must be treated as consonant sequences in Berber; but, second, that they have characteristics of their own. In particular, having a sequence of two like-consonants is likely to produce a phonetically motivated strengthening of the articulation, so that /tt/ may differ from /t/ not only in duration but also in fortition.¹¹ Convention (27) is but a reflection of this phonetic fact.

As we have seen, one result can be that geminates may be exempt from weakening processes (e.g. spirantization) which other consonant sequences

¹¹This fact is documented in phonetic studies on geminates in languages belonging to different language families, cf. Delattre [1971], Lehiste [1973], and the references cited there.

undergo. This is explained by convention (27). In fact, with its demonstrably solid phonetic basis, (27) may prove to be a linguistic universal, (especially with respect to spirantization). Indeed, data from languages other than Berber, viz. Biblical Hebrew [Sampson 1973], Old Romance languages [Pope 1934], provide support for attributing universal status to (27). In both languages, as is the case for Berber, spirantization affects single consonants but not geminates. The difference between Hebrew and Berber is that storable environments can be found for the weakening rule. An illustration is given in (28) and (29):

(28) Biblical Hebrew [Sampson 1973:101]¹²

/gaadal/	[gaadaal]	'he became great'	but:
/giddeel/	[giddeel]	'he became great'	

(29) Romance (Old French) [Pope 1937:262]¹³

<u>Late Latin</u>	<u>Old French</u>	
gratu	gre [gre]	'liking'
*totta	tote > tute	tut 'all (fem.)'

It appears, thus, that there is indeed something universal about geminate consonants not weakening into a sequence of identical spirants, viz. tt ≠ θθ. This can be explained by the mutual strengthening effect of the two elements of the geminates captured by convention (27). If this claim withstands the test in other languages, it would represent an important constraint on geminates.

¹²So as to be consistent with the way geminates are represented throughout this study, Sampson's representation of these segments is not followed here. Moreover, VV is substituted for Sampson's V: for long vowels.

¹³Pope's other examples seem to suggest that the † in gratu, first weakens to d, then to ɖ in Gallo-Roman. The loss of the ending leaves ɖ in final position, hence the devoicing into θ, before complete disappearance in Middle and Modern French, viz. 'gre' [gRe]. Notice that the geminates in *totta, while not undergoing spirantization, can undergo reduction, a fact also accounted for by (27) (see below). The asterisk before *totta indicates that it is a reconstructed form, while the small zero under the e indicates that e is reduced to a schwa.

A second claim, which in effect is a corollary of the first, is also implicit in convention (27): geminates will reduce (degeminate) first, before undergoing spirantization (or other weakening processes). Thus, there is always going to be a time lag between reduction of geminates and spirantization. This predictive power of (27) explains certain reportedly "odd" alternations between geminates and their corresponding spirants in Berber, as exemplified in (30).¹⁴

(30) a.	<u>Singular</u>	<u>Plural</u>	
	ašəḡḡur	išəḡran	'flap of a cape'
	asekkur	iseḡwran	'partridge'
b.	<u>Zero form</u>	<u>Intensive form</u>	
	qqim	ttgima	'to sit'
	ggal	ttjalla	'to swear'

This second claim implicit in convention (27) is further supported by changes which took place during the development of Modern French from Latin [Pope 1934], as illustrated in (31):

(31)	LL	OFII	Mid F	Mod F		
	toḡḡu	tot tut	tu(t)	tu	'tout'	'all'
	bassu	bas	ba(s, z)	ba	'bas'	'low'

In (31) it can be seen that Late Latin (LL) geminates tt and ss reduce in Old French (OFII), before undergoing first optional, and then obligatory deletion in Middle French (Mid F) and Modern French (Mod F), respectively.

A final piece of corroborating evidence for (27) is the existence of strengthening processes which only geminates undergo. Thus in a number of Berber dialects, among which is Taqbayliyt (=Kabyle), spoken in Algeria, ss > ts (symbolized as [cc]), but s > s. This is illustrated in (32):

¹⁴According to Loubignac [1924:39], some subdialects of Zayyan, where spirantization is not as pronounced as in Tamazight (Ayt Ndir) represented in (30), exhibit reduction but not yet spirantization.

(32) Taqbayliyt (At-Mangellat; Data from De Vincennes and Dallet [1960])¹⁵

<u>Zero form</u>	<u>Intensive form</u>	
f <u>sey</u>	fe <u>cci</u>	'to melt'
x <u>sey</u>	xe <u>cci</u>	'to be extinguished'
r <u>su</u>	re <u>ccu</u>	'to drive in (a stick)'

Supportive data illustrating the particular strengthening of the geminates can be drawn from languages other than Berber. In Spanish for example, Latin /ll/'s palatalize into [ʎ], while Latin /l/'s do not, e.g. caballo [kabalʎo] 'horse', malo [malo] 'bad'.

Since we have argued that the two elements of a geminate cluster are redundantly [+tense], while other consonants are [-tense], the processes illustrated by the data presented in this section are best explained by a theory incorporating convention (27).

REFERENCES

- Abdel-Massih, Ernest. 1968. *Tamazight Verb Structure*. African Series, Volume 2. Bloomington: Indiana University Publications.
- Abdo, Daud A. 1969. *On Stress and Arabic Phonology: A Generative Approach*. Beirut: Khayyats.
- Barkai, Malachi. 1974. "On duration and spirantization in Biblical Hebrew." *Linguistic Inquiry* 5:456-59.
- Brame, Michael. 1971. "Stress in Arabic and generative phonology." *Foundations of Language* 7:556-91.
- Dahbi, Mohamed. 1975. "On epenthesis in Moroccan Arabic." Ms., University of Southern California.

¹⁵In fact, in this dialect of Taqbayliyt (Kabyle), the intensive form prefix, tt, cf. (7) above, is an affricate transcribed here as cc (=tts). So is the lexical geminate tt in the following cognate forms:

<u>Tamazight</u>	<u>At-Mangellat</u>	
netta	necca	'he/him'
ttu-x	ccu-g	'I forgot'

Furthermore, in Tirifiyt (Northern Morocco), geminate /ll/ are rendered as [ʃʃ], e.g. Tamazight [ellef] 'to repudiate' = Tirifiyt: [eʃʃef].

- Delattre, Pierre. 1971. "Consonant gemination in four languages: an acoustic, perceptual and radiographic study." *International Review of Applied Linguistics* 9:31-52, 97-113.
- De Vincennes, Sr. L., and J.M. Dallet. 1960. *Initiation à la langue berbère (Kabyle)*. Fort National: Fichier de documentation berbère.
- Galand, Lionel. 1975. Review of Penchoen—*Tamazight of the Ayt Ndhir*. *Bulletin de la Société de linguistique* 70:345-48.
- Harms, Robert. 1968. *Introduction to Phonological Theory*. Englewood Cliffs, N.J.: Prentice Hall.
- Harries, M. Jeannette. 1966. "Syntactic structure of Tamazight." Doctoral dissertation, University of California, Los Angeles.
- Kenstowicz, Michael. 1970. "On the notation of vowel length in Lithuanian." *Papers in Linguistics* 3(1):73-113.
- Laoust, Emile. 1918. *Etude sur le dialecte des Ntifa*. Paris:Leroux.
- Lehiste, Ilse. 1973. "An instrumental study of consonant gemination." *Journal of Phonetics* 1:131-48.
- Loubignac, Victorien. 1924. *Etude sur le dialecte berbère des Zaian et Ait Sgougou*. Paris: Leroux.
- Martinet, Andre. 1959. "Du role de la gémination dans l'évolution phonologique." *Zeitschrift fur Phonetik* 12:223-27.
- Mitchell, Thomas. 1957. "Long consonants in phonology and phonetics." In *Studies in Linguistic analysis* 182-205. Oxford: Basil Blackwell.
- Penchoen, Thomas. 1973. *Tamazight of the Ayt Ndhir. Afroasiatic Dialects* Volume 1. Los Angeles: Undena Publications.
- Pope, M.K. 1934. *From Latin to Modern French*. Manchester: Manchester University Press.
- Pyle, Charles. 1970. "West Greenlandic Eskimo and the representation of vowel length." *Papers in Linguistics* 3:115-46.
- Saib, Jilali. 1974. "Gemination and the spirantization in Berber: diachrony and synchrony." *Studies in African Linguistics* 5:1-26.
- Saib, Jilali. 1975. "Epenthesis in Moroccan Arabic." Ms, University of California, Los Angeles.
- Saib, Jilali. 1976. "Schwa insertion in Berber: un problème de choix." *Afroasiatic Linguistics* 3(4):1-13.
- Saib, Jilali. 1976. "A phonological study of Tamazight Berber: dialect of the Ayt Ndhir." Doctoral dissertation, University of California, Los Angeles.
- Sampson, Geoffrey. 1973. "Duration in Hebrew consonants." *Linguistic Inquiry* 4:101-104.
- Swadesh, Morris. 1937. "The phonemic interpretation of long consonants." *Language* 13:1-10.