PROCEEDINGS OF THE TWENTY-THIRD ANNUAL MEETING

OF THE

BERKELEY LINGUISTICS SOCIETY

February 14-17, 1997

SPECIAL SESSION

ON

SYNTAX AND SEMANTICS IN AFRICA

Berkeley Linguistics Society
Berkeley, California, USA

PROCEEDINGS OF THE TWENTY-THIRD ANNUAL MEETING

OF THE

BERKELEY LINGUISTICS SOCIETY

February 14-17, 1997

SPECIAL SESSION
ON
SYNTAX AND SEMANTICS IN AFRICA

edited by

Ashlee C. Bailey

Kevin E. Moore

Jeri L. Moxley

Berkeley Linguistics Society

Berkeley Linguistics Society, Inc.
1203 Dwinelle Hall
University of California
Berkeley, California 94720-2650 USA

ISSN 0363-2946

LCCN 76-640143

Printed by Braun-Brumfield, Inc.
P.O. Box 1203
Ann Arbor, Michigan 48108 USA

TABLE OF CONTENTS

SPECIAL SESSION

• • •

The following paper was presented at the conference but does not appear in the volume:

CONTENTS OF GENERAL SESSION TO BE FOUND IN BLS 23

GENERAL SESSION

PARASESSION
ON PRAGMATICS AND GRAMMATICAL STRUCTURE

PREFACE

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

We would like to thank Matt Juge and David Peterson for their help in organizing the Special Session. We are grateful also to the participants, volunteers, and attendees. We hope you enjoy the volume.

Ashlee C. Bailey
Kevin E. Moore
Jeri L. Moxley

The unergative-unaccusative distinction and the benefactive applicative in Amharic*

Mengistu Amberber
Australian National University

1. Introduction. In this paper, I discuss the interaction between monadic predicates and the benefactive applicative construction on the basis of data from Amharic (Ethio-Semitic). I will show that the benefactive applicative construction is available for both unergative and unaccusative predicates. However, I will show that important morphosyntactic and semantic asymmetries exist between the two types of monadic predicates. I propose a structural account for the asymmetries by appealing to independently motivated principles of grammar.

The paper is organized as follows. In §2, I present general background facts regarding transitivity alternation in Amharic. In §3, I discuss some of the theoretical issues regarding the applicative construction. In §4, I examine the interaction between monadic predicates and the benefactive applicative construction. In §5, I argue that a preposition incorporation account of applicatives, à la Baker (1988a) is not tenable for the Amharic facts. I show that the conceptual status of the benefactive argument differs according to the predicate: in the case of unergatives, the benefactive is a canonical Path, in the sense of Jackendoff (1990), whereas in the case of unaccusatives the benefactive is an 'extra' affected argument.

2. Split intransitivity in Amharic. There are two types of monadic verbs in Amharic which can be classified along familiar lines as unaccusative and unergative. The unaccusative verbs are further subdivided into two sub-classes mainly on the basis of morphological criteria. I call these two sub-classes Pattern I and Pattern II. Descriptively, the Pattern I unaccusatives can be regarded as 'basic' intransitives, whereas the Pattern II unaccusatives are 'derived' intransitives. Pattern II unaccusatives are ambiguous between an inchoative and a passive reading. Representative examples of monadic verbs are presented below:

(1)	Unaccusative:	Pattern I	mət't'a	'come (intr)'
		Pattern II	tə-səbbərə	'break (intr)' or 'was broken'
(2)	Unergative:	č'əffərə	'dance'	

There are a number of morphosyntactic tests for split intransitivity in Amharic. For the present purposes, I will mention only one important diagnostic test which is associated with the distribution of causative affixes.

Amharic has two productive causative affixes. The first of these is ə-, which I call the *l-syntax* causative (cf. Hale and Keyser 1993), for reasons that do not concern us here (but see Amberber 1996). This affix attaches to Pattern I unaccusatives like √mt' 'come' to derive a causative verb such as ə-mət't'ə 'bring'. It does not attach to Pattern II unaccusatives nor to unergative verbs as the examples in (3) show.

(3) The distribution of the causative a-:

Pattern I:	mət't'a	'come (intr)'	(a-mət't'a 'bring')
Pattern II:	tə-səbbərə	'break (intr)'	(*a-tə-səbbərə)
Unergative:	č'əffərə	'dance'	(*a-č'əffərə)
Transitive:	səbbərə	'break (tr)'	(*a-səbbərə)

The second causative affix, which I refer to as the *s-syntax* causative, is as-. It has a wider distribution: with the exception of Pattern II unaccusatives, it occurs with both types of monadic verbs. A felicitous English translation of this causative affix is 'make/cause'.

(4) The distribution of the causative as-:

Pattern I:	mət't'a	'come' (intr)	(as-mət't'a 'cause x come')
Transitive:	səbbərə	'break (tr)'	(as-səbbərə 'cause x break y')
Unergative:	č'əffərə	'dance'	(as-č'əffərə 'cause x dance')
Pattern II:	*as-tə-səbbərə		

Thus, Amharic has a morphologically transparent diagnostic for unaccusativity which can be informally stated as in (5):

(5) *CAUS*-selection:
Intransitive verbs which can be causativized only by the s-syntax *CAUS* affix are unergative.

The unergative/unaccusative distinction exhibits itself in various areas of the grammar. In this paper, I will focus on how split intransitivity interacts with one valency changing derivation, namely the applicative.

3. The applicative. The basic facts about the applicative construction are by now familiar to many. Essentially, in the typical applicative construction an erstwhile oblique argument of a predicate (such as instrumental, beneficiary, locative) becomes the direct object. The verb of the applicative construction is morphologically more complex than its non-applicative variant. Consider, for instance, the examples in (6) from Chichewa, a Bantu language (from Baker 1988a:229):

(6) a. *Mbidzi zi-na-**perek**-a msampha **kwa** nkhandwe*
zebras SP-PAST-**hand**-ASP trap **to** fox
the zebras handed the trap to the fox

b. *Mbidzi zi-na-**perek-er**-a nkhandwe msampha*
zebras SP-PAST-**hand-to**-ASP fox trap
the zebras handed the fox the trap

The goal argument in (6a), the beneficiary, occurs as a direct object in (6b), the applicative construction. The verb is complex, comprising the applicative affix -er.

Likewise, consider the Chamorro (Austronesian) example in (7), (from Baker 1988a:237, original due to Gibson 1980):

(7) a. *Ha puunu' si Miguel i bəbui* ***pəra*** *guahu*
3sS-kill PN Miguel the pig **for** me
Miguel killed the pig for me

b. *Ha* ***punu'-i*** *yu' si Miguel nu i bəbui*
3sS-**kill-for** me PN Miguel OBL the pig
Miguel killed the pig for me

The beneficiary of the action denoted by the verb in (7a) occurs as the object of the preposition 'for', whereas in (7b), it occurs as the direct object of the sentence.

Amharic has a construction which exhibits similar formal properties to the applicative. This can be demonstrated by the instrumental applicative as shown in (8):

(8) a. *əster bə-mət'rəgiyə-wi məskot*
A. with-broom-DEF window
t'ərrəgə-č-(ɨbb-əti)
clean.pf.-3fS-(with-3mO)
Aster cleaned a window with the broom

b. *əster mət'rəgiyə-*(wi-ɨn) məskot*
A. broom-(DEF-ACC) window
t'ərrəgə-č-(ɨbb-əti)*
clean.pf.-3fS-with-3mO
Aster cleaned a window with the broom

In (8a) the instrumental NP 'the broom' occurs with the prepositional prefix bə-. The verb is optionally marked with the element -bb-ət which consists of a form similar to the preposition bə- and a pronominal suffix which agrees with the instrumental NP. Notice that the -bb- form and the agreement affix occur as a unit, i.e., one cannot occur without the other. For ease of exposition, I will call this unit the *B-complex*. Now consider (8b): the instrument is no longer marked by the prepositional element. Rather it occurs with the accusative suffix -n. Notice also that the B-complex which was optional in (8a) is obligatory in (8b).

The applicative is quite productive in Amharic. The preposition bə- and the associated B-complex do not occur only with instrumentals. The prepositional prefix bə- has a range of spatial and temporal meanings which include 'on, at, by, with, in', among others. It paradigmatically contrasts with another prepositional element, lə-. The classic minimal pair that shows the distribution of these two forms, bə- and lə-, is presented in (9) and (10) respectively:

(9) a. *dəññəw bə-əster fərrədə-(bb-ət)*
judge-DEF on-A. judge.pf.3mS-(**on**-3fO)
the judge judged against Aster (=he sentenced her)

b. *dañnaw aster-ɨn farrada-**bb**-at*
judge-DEF A.-ACC. judge.pf.3mS-**on**-3fO
the judge judged against Aster (=he sentenced her)

(10) a. *dañna-w la-aster farrada-(ll-at)*
judge-DEF for-A judge.pf.3mS-(**for**-3fO)
the judge judged in favor of Aster (=he acquitted her)

b. *dañna-w aster-ɨn farrada-ll-at*
judge-DEF A.-ACC judge.pf.3mS-**for-**3fO
the judge judged in favor of Aster (=he acquitted her)

In general, when a verb is marked by -bb-, as in (9), the construction has a malefactive interpretation. (9a) is the non-applicative version, whereas (9b) is the applicative counterpart. On the other hand, when a verb is marked by -ll-, as in (10), the construction has a benefactive meaning. Again the (b) example is the applicative version. For the sake of brevity, I will use the term *Benefactive* as a superordinate term to refer to both the malefactive and the benefactive constructions.

Let us now turn to the central issue: how the applicative interacts with split intransitivity.

4. The applicative and split intransitivity. It has been noted in the literature that in a number of languages the Benefactive applicative of intransitive predicates is ungrammatical. Thus, compare (11) and (12) from Bahasa Indonesian (cf. Baker 1988a:252, original due to Chung 1976):

(11) a. *Mereka mem-bawa daging itu kepada dia*
they TRANS-bring meat the to him
they brought the meat to him

b. *Mereka mem-bawa-kan dia daging itu*
they TRANS-bring-to him meat the
they brought him the meat

(12) a. *Ajah saja menj-umbang kepada rumah sakit*
father my TRANS-donate to house sick
my father donated to the hospital

b. **Ajah saja menj-umbang-kan rumah sakit*
father my TRANS-donate-to house sick
my father donated to the hospital

The verbs meaning 'bring' and 'donate' differ in transitivity: the latter cannot take a direct object. Thus, the NP 'hospital' in (12) cannot occur as a direct object of the complex verb, as shown in (12b).

Similar facts pertain in Chichewa as well, as presented in Baker (1988a:255). Consider (13) and (14):

(13) a. *Mlenje a-na-gon-a*
hunter SP-PAST-sleep-ASP
the hunter slept

b. *Mlenje a-na-gon-er-a kalulu
hunter SP-PAST-sleep-for-ASP hare
the hunter slept for the hare

(14) a. Mkango u-ku-yenda-a
lion SP-PRES-walk-ASP
the lion walked

b. *Mkango u-ku-yenda-er-a anayani
lion SP-PRES-walk-for-ASP baboons
the lion walked for the baboons

Again, the verbs meaning 'sleep' and 'walk' which are typical intransitive predicates do not permit the applicative as the ungrammatical sentences in (13b) and (14b) show.[1]

Interestingly, certain predicates in Chichewa can permit the applicative despite their intransitive valency. A case in point is the verb meaning 'dance' in (15), (cf. Baker 1988a: 258).

(15) a. Atsikana a-na-vin-a
girls SP-PAST-dance-ASP
the girls danced

b. Atsikana a-na-vin-ir-a mfumu
girls SP-PAST-dance-for-ASP chief
the girls danced for the chief

Baker (1988a) offered a Case theoretic account for the interaction of intransitivity with the Benefactive applicative. The basic idea is as follows. First it is assumed that the applied object needs structural Case from the verb. If the simple verb does not have structural Case to assign, either because it is lexically a non-Case assigner, as in a basic intransitive verb or is derivationally deprived of its Case assigning properties, as in the passive and antipassive, the complex verb cannot assign structural Case. In order to accommodate data such as (15), Baker suggests that unergative verbs such as 'dance' take cognate objects and thus are different from other intransitive verbs. Such unergative verbs behave as transitive predicates and are capable of Case assignment. Thus, for Baker (1988a) the reason why transitive verbs and unergative verbs which take cognate objects allow the Benefactive applicative follows from Case theory.

Now, the Amharic examples above show that the Benefactive applicative is possible with both types of intransitive predicates. Consider for instance the unergative verb 'laugh' in (16a). This verb is unergative as evidenced by the unaccusativity diagnostic of *CAUS*-selection: it cannot take the affix a- as shown in (16b). However, notice that this verb can appear in the Benefactive applicative as in (16c). An otherwise oblique argument occurs as a direct object marked by the accusative Case.

(16) a. aster sak'a-č
A. laugh.pf.-3mS
Aster laughed

b. * *ləmmə aster-ɨn ə- sək'-ət*
L. A.-ACC CAUS-laugh.pf.3mS-3fO

c. *aster ləmmə-n sək'ə-č-*(ɨbb-ət)*
A. L.-ACC laugh.pf.-3fS-(on-3mO)
Aster laughed at Lemma

The Benefactive applicative of unergative verbs is fairly productive. The only obvious non-lexical exception involves definiteness: an indefinite argument cannot occur as the applied object, nor can it co-occur with the B-complex, as shown in (17a) and (17b). The only acceptable construction involves the absence of the B-complex as shown in (17c).

(17) a. **aster bə-səw sək'ə-č-ɨbb-ət*
A. at-someone laugh.pf.-3fS-on-3mO

b. **aster səwɨ-n sək'ə-č-ɨbb-ət*
A. someone-ACC laugh.pf.-3fS-on-3mO

c. *aster bə-səw sək'ə-č*
A. at-someone laugh.pf.-3fS
Aster laughed at someone

What about the Benefactive applicative of unaccusatives? Consider the Pattern I unaccusative verb 'come' in (18).

(18) a. *ɨngɨdə mət't'ə*
guest come.pf.3mS
a guest came (arrived)

b. **ɨngɨdə bə-aster mət't'ə-(bb-ət)*
guest on-A. come.pf.3mS-(on-3fO)

c. *aster-(ɨn) ɨngɨdə mət't'ə-*(bb-ət)*
A.-(ACC) guest come.pf.3mS-(on-3fO)
lit. Aster a guest arrived on her

Notice that an important asymmetry emerges between the Benefactive applicative of unergatives and that of unaccusatives. Unlike the Benefactive of unergative verbs, a Benefactive argument of unaccusatives cannot occur with the positional element bə-, as shown in (18b). However, the applicative construction is possible as shown in (18c). The same is true for Pattern II unaccusatives as in (19).

(19) a. *t'ərmus-u tə-səbbərə*
bottle-DEF INCH-break.pf.3mS
the glass broke

b. **t'ərmus-u bə-aster tə-səbbərə-(bb-ət)*
bottle-DEF by-A. INCH-break.pf.3mS-(on-3fO)

c. *aster-(ɨn) t'ərmus-u tə-səbbərə-*(bb-ət)*
A.-(ACC) bottle-DEF INCH-break.pf.3mS-(on-3fO)
lit. Aster the glass broke on her (she is adversely affected)

There is another non-trivial difference between the unergative and unaccusative predicates: the Benefactive argument must occur clause initially. If it does not, as in (20), the construction becomes ungrammatical.

(20) *t'ərmus-u əster-(ɨn) tə-səbbərə-bb-ət
bottle-DEF A-(ACC) INCH-break.pf.3mS-on-3fO

The Benefactive applicative of unaccusative verbs is also productive. In fact, it occurs with verbs which do not normally take an oblique argument. Consider the verb məššə 'become night'. In isolation, the event encoded by this verb is neutral with respect to affectedness. However, the event can be conceived of as adversely or favorably affecting someone when it is cast in the applicative construction as in (21a) and (21b) respectively:

(21) a. əster-(ɨn) məššə-**bb**-ət
A.-(ACC) night.pf.3mS-**on**-3fO
lit. Aster it became night on her
b. əster-(ɨn) məššə-**ll**-ət
A.-(ACC) night.pf.3mS-**for**-3fO
lit. Aster it became night for her

Predicates which express time, weather, physical and mental states can be involved in the applicative construction. However, there are some constraints which can be explained on semantic and/or pragmatic grounds. Consider the examples in (22).

(22) a. *əster-(ɨn) ɨgr-wə tə-səbbərə-bb-ət
A.-(ACC) leg-her INCH-break.pf.-on-3fO
b. əster ɨgr-wə tə-səbbərə
A. leg-her INCH-break.pf.3mS
lit. Aster her leg broke

Recall that in (19c) the unaccusative verb meaning 'break' can occur in the Benefactive applicative. However, for certain choices of arguments, the construction is not available, as can be seen in (22a). When the event is conceived of as obviously malefactive, such as 'the breaking of one's body part', the applicative is not possible. Instead, a simple predicative relationship between the body part and the unaccusative verb suffices, as in (22b).

Hence, although both unergatives and unaccusatives allow the Benefactive applicative, they exhibit two important differences:

(23)		*Unergative*	*Unaccusative*
	a. Benefactive can occur in a PP	yes	no
	b. Benefactive is clause-initial	no	yes

The interaction between split intransitivity and the applicative construction reveals an interesting partitioning of monadic verbs in a number of other languages. For instance, in several Australian languages only unergative verbs can be involved in the applicative. In Arabana-Wangkangurru, as presented in Austin (1995,

original due to Hercus 1990), there are two causative affixes: ma- and la-. The former is described as encoding 'mediated causation', which, for the present purposes, I consider to be the formal equivalent of the Amharic s-syntax causative as- or the English independent verb 'make'. This affix can attach to both unergative and unaccusative types of verbs.

On the other hand, the affix -la- exhibits a split in intransitivity: when it attaches to unaccusative verbs it derives a causative construction, whereas when it attaches to unergative verbs it derives an applicative. These are shown in (24) and (25) respectively (from Austin 1995):

(24) *Causative* -la-: kaji- 'to turn' kaji-la- 'to turn (it) over'
(25) *Applicative* -la-: wiya- 'to laugh' wiya-la- 'to mock, laugh at'

Thus, to summarize the facts so far, in Amharic the Benefactive applicative construction is available in both unergative and unaccusative predicates. Nevertheless, important differences are shown to pertain between the two classes.

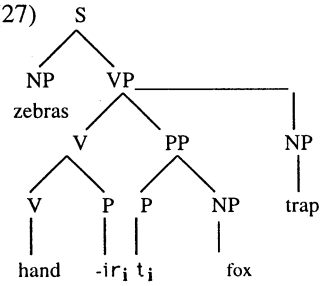
In the remainder of this paper, I will attempt to account for the observed asymmetry between unergative and unaccusative predicates with respect to the Benefactive applicative in Amharic. I will motivate a structural analysis that best accounts for the Amharic facts.

5. Structural account. Let us begin with the incorporation analysis of applicatives as outlined in Baker (1988a). Consider (6) repeated below as (26):

(26) a. *Mbidzi zi-na-**perek**-a msampha **kwa** nkhandwe*
zebras SP-PAST-**hand**-ASP trap **to** fox
the zebras handed the trap to the fox
b. *Mbidzi zi-na-**perek-er**-a nkhandwe msampha*
zebras SP-PAST-**hand**-ASP fox trap
the zebras handed the fox the trap

For Baker, both (26a) and (26b) have an identical D-structure thematic configuration. He argued that the applicative in (26b) is derived when the preposition, or applied affix, incorporates into the verb as in (27).

(27)



The major problem in applying the preposition incorporation (PI) account to the Amharic facts emerges from the co-occurrence of the positional element bə- (or lə-) with the B-complex. Thus, consider the relevant example repeated in (28):

(28) *aster* *bə-ləmmə* *sək'ə-č-(ɨbb-ət)*
A. at-L. laugh.pf.-3fS-(on-3mO)
Aster laughed at Lemma

If the source of -bb- in the verb is the incorporation of the positional bə-, then we would not expect the two to co-occur. In other words, the bə- on the Benefactive NP and the -bb- on the verb should have been in complementary distribution, contrary to fact. Thus, the occurrence of -bb- in the verb could not have been as a result of PI.

It is instructive to note that in a later study, Baker (1992:29) has modified the PI account for some constructions in Chichewa, arguing that certain locative Benefactives are possible without the preposition incorporating into the verb, as in (29b).[2]

(29) a. *Alenje* *a-ku-luk-ir-a* *pa-mchenga* *mikeka*
hunters SP-pres-weave-**appl**-ind **on**-sand mats
the hunters are weaving mats on the beach
b. *Alenje* *a-ku-luk-ir-a* *mikeka* *pa-mchenga*
hunters SP-pres-weave-**appl**-ind mats on-sand
the hunters are weaving mats on the beach

For the Amharic facts, I would like to propose that the B-complex has a different structural status depending on the lexical semantics of the predicate. I would like to argue that the Benefactive is a canonical Path argument in the Lexical Conceptual Structure (LCS) of unergative verbs. I assume that unergative verbs encode an Activity Event-type. I suggest that the Benefactive argument makes the Activity event more complete. To see what I mean by making the event complete, consider the event denoted by verbs such as 'laugh' or 'cry'. When someone laughs or cries, there is often a stimulus for the event. I suggest that the Benefactive argument partially spells out that stimulus.

For unaccusative verbs, I argue that the Benefactive does not make the Achievement Event-type complete. The Benefactive is rather an 'extra' affected argument.

The difference between the Benefactive of the two predicates can be informally highlighted by *wh*-questions. Since the Benefactive elaborates an Activity Event-type it is quite natural to ask, for instance 'who is Mary laughing at?'. On the other hand, as the Benefactive is an extra argument of the Achievement Event-type, it is rather odd, at least in neutral contexts, to ask 'who is affected by the breaking of the glass'.

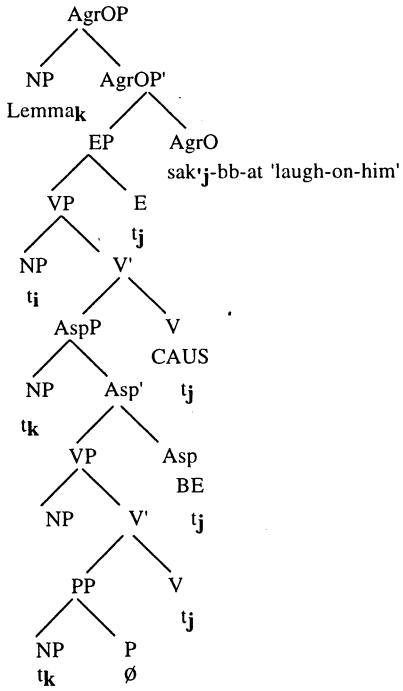
Once the conceptual status of the Benefactive with respect to the predicate is clearly established, the syntactic difference between the two classes of monadic verbs can be accounted for by independently motivated principles of grammar. I suggest that, appealing to a Jackendovian type Conceptual Semantics (cf. Jackendoff 1990), a canonical Path argument is mapped onto a PP. Thus, the

Benefactive argument of unergative verbs occurs in a PP. On the other hand, the Benefactive argument of unaccusatives is a marked Path and is mapped onto an NP.

Suppose that with the unergative predicates, the PP may be generated either with a lexically filled head or an empty head. When P is lexically filled it can assign Case to its argument. Then the B-complex on the verb can be regarded as an optional oblique agreement.

On the other hand, P may be generated as an empty head. I assume that empty heads do not assign Case, nor can they preclude a governing verb from assigning Case (see also Baker 1992). Assuming that in Amharic the unergative verb can assign structural Case, the Benefactive argument gets Case. I assume that the B-complex is generated in AgrO and the Benefactive argument must move to Spec AgrO to get morphological Case. The applicative of unergatives can be represented as in (30), ignoring irrelevant details. (Note that t_i is the trace of the subject):

(30)

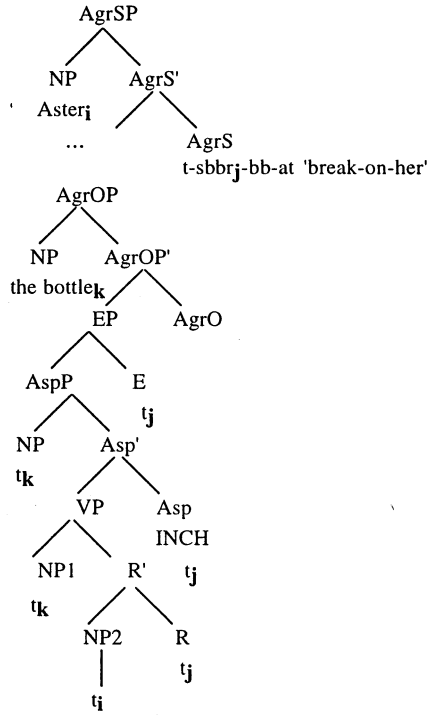


Now, consider the Benefactive applicative of the unaccusative predicate, repeated in (31):

(31) *aster-(ɨn)* *t'armus-u* *tə-səbbərə-*(bb-ət)*
A.-(ACC) bottle-DEF INCH-break.pf.3mS-(on-3fO)
lit. Aster the glass broke on her (she is adversely affected)

The Benefactive argument is generated as an NP complement of the verb root. It cannot get Case from the verb, because, by hypothesis, unaccusatives do not assign structural Case. However, I assume that the Benefactive can be assigned *inherent* Case. It is often assumed (cf. Chomsky 1986) that inherent Case is an optional Case that is assigned only when it is required. I assume that the B-complex is generated as the head of AgrS and the Benefactive argument moves to Spec AgrS to check morphological Case. The derivation is partly schematized in (32):

(32)



Therefore, the distinction between the two predicates is attributed to the structural status of the Benefactive argument which itself follows from the lexical semantics of the verbs.

Incidentally, one may also argue for a 'lexicalist' account where the applicative marker, the -bb- suffix, is lexically inserted into the verb. It is beyond the scope of this paper to address in detail the issue of determining whether a syntactic or a lexicalist account better handles the applicative construction in Amharic. However, I note here that to make a lexical analysis operational one has to assume that the agreement suffix is also lexically inserted. Since for independent reasons I assume that the agreement suffixes are syntactically inserted, the simplest assumption would be to generate the B-complex in the relevant Agr projection.[3]

6. Conclusion. To conclude, the interaction between split intransitivity and the Benefactive applicative construction in Amharic can be accounted for if we assume that the Benefactive argument has a different lexical conceptual status depending on the lexical semantics of the predicate. With unergative verbs it is an argument that makes the event more complete, whereas with unaccusative verbs it is an affected argument.

Notes

* I thank Lisa Travis for comments and suggestions on ideas presented in this paper. I also thank Hilda Koopman for useful questions at the venue of BLS 23rd. Of course, any shortcomings that might be reflected here are entirely mine. The following abbreviations are used in the interlinear gloss of Amharic sentences: ACC 'accusative', m 'masculine', f 'feminine', S 'subject', O 'object', DEF 'definite', CAUS 'causative', INCH 'inchoative', PASS 'passive', pf 'perfect'. Also, the abbreviations EP and AspP refer to 'event phrase' and 'aspect phrase' respectively. For interlinear glosses of examples from other languages, the reader is referred to the original sources.

1. As noted in Baker (1988a), these sentences can be grammatical with a reading that is different from the typical applicative. For instance, (13b) would be grammatical with the reading: 'the hunter lay on the hare'. Note also that Alsina and Mchombo (1990) argue against the claim that beneficiary applicatives cannot be formed from intransitives. However, they implicitly admit that there is an asymmetry between the benefactive applicative of transitives and that of intransitives. They claim: "it is true that beneficiary applicatives based on certain intransitive verbs do not have the full range of interpretations they get when based on transitive verbs" (Alsina and Mchombo 1990:502).

2. See also Baker (1996) for a different analysis of applicative constructions which does not appeal to the standard incorporation account.

3. See Mullen (1986) for a study of Amharic agreement affixes within the framework of Lexical Phonology.

References

Alsina, A. 1993. *Predicate Composition: A Theory of Syntactic Function Alternations*, Ph.D. dissertation, Stanford.

Alsina, A. and S. Mchombo. 1990. 'The Syntax of Applicatives in Chichewa: Problems for a Theta Theoretic Asymmetry', *Natural Language and Linguistic Theory* 8, 493 - 506.

Amberber, M. 1996. *Transitivity Alternations, Event-types and Light Verbs*, Ph.D. dissertation, McGill University.

Austin, P. 1995. 'Causatives and Applicatives in Australian Aboriginal Languages', ms., La Trobe University.

Baker, M. 1988a. *Incorporation, A Theory of Grammatical Function Changing*, Chicago University Press, Chicago.

Baker, M. 1988b. 'Theta Theory and the Syntax of Applicatives in Chichewa', *Natural Language and Linguistic Theory* 6, 353-389.

Baker, M. 1992. 'Thematic Conditions on Syntactic Structures: Evidence from Locative Applicatives', in I. M. Roca (ed.), *Thematic Structure: its Role in Grammar*, Foris, Berlin, pp. 23-46.

Baker, M. 1996. *The Polysynthesis Parameter*, Oxford University Press, Oxford.

Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origins, and Use.* Praeger, New York.

Chung, S. 1976. 'An Object-Creating Rule in Bahasa Indonesia', *Linguistic Inquiry* 7, 1-37.

Gibson, J. 1980. *Clause Union in Chamorro and in Universal Grammar*, Ph.D. dissertation, University of California, San Diego.

Hale, K. and S.J. Keyser. 1993. 'On Argument Structure and The Lexical Expression of Syntactic Relations', in K. Hale and S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvian Bromberger,* The MIT Press, Cambridge, pp. 53-109.

Hercus, L. 1990. 'Arabana-Wangkangurru Grammar', ms., Australian National University.

Jackendoff, R. 1990. *Semantic Structure*, The MIT Press, Cambridge.

Mullen, D. 1986. *Issues in the Morphology and Phonology of Amharic: The Lexical Generation of Pronominal Clitics*, Ph.D. dissertation, University of Ottawa.

Travis, L. (in press). 'The L-syntax/S-syntax Boundary: Evidence from Austronesian', in I. Paul, V. Phillips, L. Travis (eds.), *Formal Issues in Austronesian Linguistics*, Kluwer, Dordrecht.

Discourse salience in Kalenjin inter-clausal syntax*

Gregory D. S. Anderson
University of Chicago

1. Introduction. The Kalenjin languages are spoken in the Western Highlands region of Kenya, and adjacent areas of Uganda by a cattle-rearing population numbering perhaps one million total. Along with the Datooga of Tanzania and the moribund Omotik, they form the Southern Nilotic group of the Nilo-Saharan language phylum.[1] Like various languages of north-eastern Africa, the head-marked Kalenjin verb indexes a variety of argument properties, as well as a range of tense, aspect, and voice categories, etc.[2] Although such processes as Topicalization can move nominal arguments or adjuncts to pre-verbal position in Kalenjin languages, verbs usually appear clause-initially. While the preference for indexing salient, subcategorized animate referents within mono-clausal sentences regardless of their theta-role or grammatical relation is rather common cross-linguistically (see Anderson 1995 for a discussion), the type of referent indexing that holds between clauses in a complex Kalenjin sentence is not encountered that frequently. In Nandi and the closely related Kipsigis, when certain verbs that may take clausal complements are used in the higher clause of a complex sentence and there is an animate referent that is highly salient to the discourse in a lower clause, speakers have the option to maximize the indexing of these salient animate referents, marking them in the higher and lower clauses. In addition to Kalenjin, this 'Copying' formation is also attested in various Algonquian languages. Thus, referent indexing in inter-clausal syntax may be determined not only by the semantic or structural relations of the nominals involved, but by their high degree of discourse salience as well.

2. Copying Constructions in Kalenjin

2.1. Nandi. Nandi (Creider and Tapsubei Creider 1989) is described as a VSO language; thus, verbs are clause-initial, with nominal arguments and adjuncts generally following. In complex sentences this may result in a sequence of two verbs, each bearing an inflectional affix marking the person/number of the SUBJ, even when the SUBJs of the two verbs are co-referential.

(1)

Nandi (Creider and Tapsubei Creider 1989: 145)

á-mácé i:-ke:r (ínye:) kâ:t
[1-'want'] [2-'see'] ['you'] ['house']

'I want that you see the house'

Nandi (Hollis 1909: 112)

ingo'ngo-a-moche a-wei-ke ane, a-ñgêt korok
[Rdpl.CONDIT-1-'want'] [1-'turn'-RFLXV] ['I'] [1-'arise'] ['first of all']

'if I wish to turn over first of all I have to get up'

Nandi (Creider and Tapsubei Creider 1989: 130)

á-mácé a:-ke:r Kiplakat
[1-'want'] [1-'see'] ['Kiplagat']

'I want to see Kiplagat'

As stated above, when a referent in the lower clause is determined by the speaker to be particularly salient, it is possible to copy that referent into the verb form of the higher clause.

(2)

Nandi (Creider and Tapsubei Creider 1989: 145)

á-mác-í:n <ínyê:> i:ke:r <ínyê:> kâ:t
[1-'want'-2OBJ] <'you'> [2-'see'] <'you'> ['house']

'I want you to see the house'

In Nandi, Copying-to-OBJ is possible even when the optional complementizer verb *-le* 'say' is present. The complementizer is opaque to the copying process; that is, rather than being copied onto this word, the copied element skips over it and is indexed directly in the sentence-initial verb.

(3)

Nandi (Creider and Tapsubei Creider 1989: 147)

á-mác-í:n a:-le i:ke:r kâ:t
[1-'want'-2OBJ] [1-COMP] [2-'see'] ['house']

'I want you to see the house'

2.2. Kipsigis. In addition to Nandi, a variety of referent copying formations are also found in the Kalenjin language Kipsigis (Kipsikiis). As in Nandi, verbs are generally clause-initial in Kipsigis.

(4)

Kipsigis (Jake and Odden 1979: 134, 135, 136)

mɔcé Mù:sá à-lápát
['want'] ['Musa'] [1-'run']

'Musa wants that I run'

kà-yây Mù:sá à-tíl pè:ndɔ
[PAST-'make'] ['Musa'] [1-'cut'] ['meat']

'Musa made me cut the meat'

rî:pé Mù:sá à-tíl pè:ndɔ
['watch'] ['Musa'] [1-'cut'] ['meat']

'Musa watches me cut the meat'

As in the closely related Nandi, the person and number properties of the SUBJ are indexed on both verbs even when they are co-referential in Kipsigis as well:

(5)

Kipsigis (Jake and Odden 1979: 135)

ɔ-mɔcé à-lápát
[1-'want'] [1-'run']

'I want to run'

When the SUBJ of the lower clause exhibits a high degree of discourse salience, it may be optionally copied as the OBJ of the higher verb in Kipsigis as well.

(6)

Kipsigis (Jake and Odden 1979: 134, 135, 136)

mɔc-ɔ:n Mù:sá à-lápát
['want'-1OBJ] ['Musa'] [1-'run']

'Musa wants me to run'

kà-yáy-ân Mù:sá à-tíl pè:ndɔ
[PAST-'make'-1OBJ] ['Musa'] [1-'cut'] ['meat']

'Musa made me cut the meat'

rî:p-ɔ:n Mù:sá à-tíl pè:ndɔ
['watch'-1OBJ] ['Musa'] [1-'cut'] ['meat']

'Musa watches me cut the meat'

In addition, a highly salient OBJ of a lower clause may also be copied as the higher verb's OBJ in Kipsigis:

(7)

Kipsigis (Jake and Odden 1979: 136-7)

mɔcé Mù:sá kɔ-tìl-án Kíplàŋàt
['want'] ['Musa'] [3-'cut'-1OBJ] ['Kiplangat']

'Musa wants Kiplangat to cut me'

mɔc-ɔ:n Mù:sá kɔ-tìl-án Kíplàŋàt
['want'-1OBJ] ['Musa'] [3-'cut'-1OBJ] ['Kiplangat']

'Musa wants Kiplangat to cut me'

Note that sentences with multiple applications of Copying-to-OBJ are also attested in Kipsigis; see examples in (8).

(8)

Kipsigis (Jake and Odden 1979: 137-8)

ɔ-mɔcé kɔ-yây Kíplàŋàt kɔ-tìl-ín Mù:sá
[1-'want'] [3-'make'] ['Kiplangat'] [3-'cut'-2OBJ] ['Musa']

'I want that Kiplangat make that Musa cut you'

ɔ-mɔcé kɔ-yày-ín Kíplàŋàt kɔ-tìl-ín Mù:sá
[1-'want'] [3-'make'-2OBJ] ['Kiplangat'] [3-'cut'-2OBJ] ['Musa']

'I want that Kiplangat make Musa cut you'

ɔ-mɔc-í:n kɔ-yày-ín Kíplàŋàt kɔ-tìl-ín Mù:sá
[1-'want'-2OBJ] [3-'make'-2OBJ] ['Kiplangat'] [3-'cut'-2OBJ] ['Musa']

'I want Kiplangat to make Musa cut you'

As Jake and Odden (1979) demonstrate, in these multiply embedded constructions the process seems to apply in a 'cyclic' fashion, i.e. Copying-to-OBJ must spread leftward from the most embedded clause. This is true whether the copied referent acts as the SUBJ or OBJ of the lowest clause.

(9)

Kipsigis (Jake and Odden 1979: 138)

** *ɔ-mɔc-í:n kɔ-yày Kíplàŋàt kɔ-tìl-ín Mù:sá*
[1-'want'-2OBJ] [3-'make'] ['Kiplangat'] [3-'cut'-2OBJ] ['Musa']

'I want Kiplangat to make Musa cut you'

** *ɔ-mɔc-í:n kɔ-yày Mù:sá i:-tìl pè:ndɔ*
[1-'want'-2OBJ] [3-'make'] ['Musa'] [2-'cut'] ['meat']

'I want Musa to make you cut the meat'

Note that Copying-to-OBJ is not permitted in Kipsigis when another, non-copied NP has been 'elevated' into the higher clause (or 'topicalized' in the embedded clause); this is a restriction on having two 'topics' or entities otherwise marked as particularly attention-worthy in one sentence, as immediately pre-verbal position is the favored site for overtly topicalized constituents in Kipsigis clauses, even in complex sentences:

(10)

Kipsigis (Jake and Odden 1979: 138)

***ká:yày-ît Kîplàŋàt kɔ-tìl-ín*
[1.PAST-'make'-2OBJ] ['K'] [3-'cut'-2OBJ]

'I made Kiplangat cut you'

In doubly embedded sentences when the middle verb is intransitive, one encounters an interesting Copying-to-SUBJ operative (11ii), which then in turn 'feeds' Copying-to-OBJ on the highest clause (11iii):

(11)

Kipsigis (Jake and Odden 1979: 148)

i. *ká:-yày kɔtèstâ kɔmàsín Kíplàŋàt*
[1.PAST-'make'] [3-'continue'] [3-'hit'-2OBJ] ['Kiplangat']

'I made Kiplangat's hitting you continue'

ii. *ká:-yày ì:tèstâ kɔmàsín Kíplàŋàt*
[1.PAST-'make'] [2-'continue'] [3-'hit'-2OBJ] ['Kiplangat']

'I made Kiplangat's hitting you continue'

iii. *ká:-yày-ìn ì:tèstâ kɔmàsín Kíplàŋàt*
[1.PAST-'make'-2OBJ] [2-'continue'] [3-'hit'-2OBJ] ['Kiplangat']

'I made Kiplangat's hitting you continue'

Other intransitive verbs show copying-to-SUBJ in Kipsigis, including 'Tough'-movement type predicates. In the first example (12i), the lowest clause OBJ is copied as the SUBJ of the middle clause, while the highest predicate exhibits an interesting long-distance SUBJ-to-SUBJ copying process. Example (12ii) exhibits another long-distance Copying-to-SUBJ process, with the element copied into the highest clause functioning as the lowest clause's OBJ.[3]

(12)

Kipsigis (Jake and Odden 1979: 149, 150)

i. *ì-wûy à-nyɔlú ì:-tìl-án*
[2-'hard'] [1-'necessary'] [2-'cut'-1OBJ]

'it is hard for it to be necessary for you to cut me'

ii. *ì-nyɔlú kɔ-màc-án Kíplàŋàt à-màs-ín*
[2-'necessary'] [3-'want'-1OBJ] ['Kiplangat'] [1-'hit'-2OBJ]

'it is necessary that Kiplangat want me to beat you'

3. Copying-to-OBJ in Algonquian. The Algonquian languages were traditionally spoken in a vast territory in North America stretching from western Canada to the eastern United States. These languages are radically head-marking and polysynthetic, the verb consisting of a large stem-complex with incorporated nouns and adverbs and various inflectional affixes indexing referent properties and tense/aspect/mood categories. Algonquian transitive verb stems are classified according to the animacy of their subcategorized OBJ arguments. As in the Kalenjin languages, salient animate referents in the lower clause of a complex sentence can be copied and thus multiply indexed in various Algonquian languages as well.[4]

3.1. Fox (Mesquakie). The Algonquian language Fox (Mesquakie) is spoken by around 1,000 people at the Mesquakie settlement near the town of Tama, Iowa in the central US. Fox (Dahlstrom 1996-ms) is one of the most archaic Algonquian languages, often exhibiting forms nearly identical to reconstructed Proto-Algonquian (Bloomfield 1946). In Fox clauses that consist of more than just a single verb, constituent order is generally determined by such discourse-based notions as Topic and Focus (Dahlstrom 1993). While Fox transitive verbs are divided according to the animacy of their OBJ argument, verbs subcategorized for a complement clause are usually inflected as Transitive Inanimate, that is, as if the lower clause functions grammatically as an inanimate noun.[5]

(13)

Fox (Mesquakie) (Anderson forthcoming)

nenatawe:neta wi:h=we:ta:se:wiya:ni
[1-'want'-1>0] [FUT='be.warrior'-1/aorist]

'I want to be a warrior'

eh=kehke:netama:ni e:h=kemo:temiwa:tehe aša:haki
[AOR='know'-inan-1>0)/aor] [AOR='steal'-3PL/aor.irrealis]
[Sioux.prox.ANIM.PL]

'I realized that the Sioux must have stolen it'

However, when an animate argument of the lower clause exhibits a high degree of discourse salience, that referent may be indexed as the OBJ in the higher clause. The referent is signaled twice (once in the higher clause and once in the lower clause) and both clauses are potentially free-standing syntactically. As is the case in Kalenjin, it is therefore clear that this is not some kind of 'Raising-to-OBJ' phenomenon, but rather one of copying the referent into the higher clause. When the animate SUBJ of the lower clause is non-coreferential with the SUBJ of the higher clause, the lower clause's SUBJ is copied as the higher clause's OBJ.

(14)

Fox (Mesquakie) (Anderson forthcoming)

a:kwi=ča:h=meko aka:wa:nena:nini wi:h=keteminawiyani
['not'='so'=EMPH] ['want'-1>2/negative] [FUT='bless'-2>1/aorist]

'but I don't want you to bless me'

maneto:wa wi:h=keteminohki ine:nemenokwe:ni
['spirit-prox.sg.anim.] [FUT='bless'-3>2/aor] ['think.thus.of.'-3>2/interrog]

'the spirit must have intended to bless you'

wi:h=to:hki:yani ketene:nemene
[FUT='wake.up.'-2/aor] [2-'want'-1>2/i.i.]

'I want you to wake up'

nekehke:nemekopi e:h=kakano:neti:hena:ni
[1-know-INV-X>(1)/i.i] [AOR=talk.w/-1-2/aor]

'it is known that I talked w/you'

As in the Kipsigis examples above, in some cases it is the OBJ of the lower clause that is copied as the OBJ of the higher one in Fox.[6]

(15)

Fox (Mesquakie) (Anderson forthcoming)

netaka:wa:na:wa=koh wi:h=ne:waki
[1-'want'-DIR(1>)-3/i.i.='you.know'] [FUT='see'-1>3/aorist]

'I do want to see him'

mehto:či=meko nekehke:nema:wa wi:h=komisahiči e:h=ine:nemiči
['like'=EMPH] [1-'know'-DIR-3/i.i.] [FUT='swallow'-3>1/aor] [AOR='intend'-3>1/aor]

'it was like I knew that he intended to swallow me'[7]

In most instances, Fox Copying-to-OBJ constructions are not obligatory, but rather an optional means of highlighting an animate referent exhibiting a high level of discourse salience. In one instance however, Copying-to-OBJ is apparently obligatory. This formation underscores both the discourse-motivated function of the Copying-to-OBJ construction as well as its non-identity with raising constructions from more familiar languages. An animate (non-subcategorized) nominal functioning as topic in the embedded clause can appear only if Copying-to-OBJ operates to index that animate participant in the higher clause as well; otherwise the sentence is ungrammatical (Dahlstrom 1993).

(16)

Fox (Mesquakie) (Anderson forthcoming)

nekehke:nemekwa ni:na e:h=pwa:wi-ke:ko:hi-ašenoniki
[1-'know'-INV-3(>1)/i.i.] [1.TOP] [AOR='not'-'anything'-'disappear'-inan.obv./AOR]

'he knows as for me nothing is missing'

but **kehke:netamwa ni:na e:h=pwa:wi-ke:ko:hi-ašenoniki*
['know'-inan-3/i.i.]

Note also the interaction of embedded topics and copying-to-OBJ in Kipsigis mentioned in 2.2 above.

3.2. Blackfoot. Blackfoot is an Algonquian language spoken in the western parts of the US and Canada by around 9,000 people, or about 60% of the Blackfoot population. Unlike Fox (Mesquakie), Blackfoot is highly divergent within the family, exhibiting great phonological innovations and morphological re-

organizations in comparison to reconstructed Proto-Algonquian and such attested languages as Fox, Cree, or Ojibwa. However, despite the obvious differences between Blackfoot and the more central Algonquian languages, discourse salience also plays a role in the inter-clausal (morpho-)syntax of this language as well. In fact, it was in Blackfoot that the Copying-to-OBJ formations were first described (Frantz 1978).

In terms of the formal mechanisms utilized in the Copying-to-OBJ formations, Blackfoot agrees with Fox in requiring Transitive Animate inflections in the higher clause and multiple indexing of the copied referent. Similar to Kalenjin and Fox, these formations are mostly optional for Blackfoot speakers.[8]

(17)

Blackfoot (Frantz 1978: 90; 92)

nitsíksstatawa noxkówa máxka'po'takssi
[1-'want'[TA]-DIR-3] [1-'son'-3] [3-'might'-'work'-conj.]

'I want my son to work'

kitsíksstato káxkso'kaassi
[2-'want'[TA]-DIR(-1)] [2-'might'-'be.strong'-conj.]

'I want you to be strong'

Also like Fox and Kipsigis, not only lower clause SUBJ but lower clause OBJ can be copied into the higher clause in Blackfoot.

(18)

Blackfoot (Frantz 1978: 99)

kit-aíksim'sstat-o k'áxk-oxk-awaayáki-ook-oo-xsi
[2-'think'-DIR(-1)] [2-'might'-'just'-'hit'-INV-X-conj]

'I'm thinking you might get hit'

Copying-to-OBJ formations are found in other Algonquian languages as well, e.g. Plains Cree (Dahlstrom 1991: 67, 72) *nikiske:yima:w e:=no:hte-sipwe:htet* [1-'know'-DIR-(1>)3] [AOR='want'-'leave'-3/conj] 'I know that he wants to leave' and *ka:=wa:pama:t e:h=kito:we:hkwa:miyit* [AOR='see'-3>obv] [AOR='snore'-obv/conj] 'he_i saw that he_j was asleep and snoring',[9] or Mille Lacs Ojibwa (Nichols 1980: 227) *ninantawenimik ci-maacaayaan* [1-'want'-INV] [PV-'leave'-1/conj] 'he wants me to go'.

4. Summary. The various copying processes examined in the Kalenjin and Algonquian languages above are motivated on the functional level of discourse salience. These are generally optional formations available to the speakers of particular head-marking languages to utilize the morphosyntactic machinery of the

language to maximize the indexing of highly salient referents in complex sentences. These formations copy these highly salient referents from the lower clause and index them in the verb form of the higher clause.[10] Given that patients and syntactic direct objects in general disfavor animate reference (Silverstein 1976; Croft 1990), this type of referent indexing is unexpected. However, the copying processes operative in Kalenjin and Algonquian clearly demonstrate that in inter-clausal syntax, referent indexing may be motivated not only by semantic and structural relations, but by considerations of discourse salience as well.

Notes

* The following abbreviations are used in the interlinear glosses throughout this paper:

anim	animate	AOR	aorist	COMP	complement(izer)
CONDIT	conditional	conj	conjunctive	DIR	direct (theme sign)
EMPH	emphatic	FUT	future	INV	inverse (theme sign)
inan	inanimate	>	acting on	i.i.	independent indicative
OBJ	object	obv	obviative	PL	plural
prox	proximate	PV	preverb	Rdpl	reduplication
RFLXV	reflexive	sg	singular	TA	transitive animate
TOP	topic(alized)	X	unspecified	0	inanimate
1	first person	2	second person	3	third person

1 The sub-grouping of Nilo-Saharan is far from agreed upon, cf. for example the differing theories discussed in Bender (1996). In addition to Nandi and Kipsigis, Kalenjin includes the Keyo (Elgeyo), Kony (Elgon), Marakwet (Markweta), Päkot (Pokot), Sabiny (Sebei, Sabaot), Terik (Nyang'ori) and Tusen languages as well; see Tucker and Bryan (1964) for one possible subgrouping of these languages. The South Nilotic (Rottland 1982) group of languages consists of the Kalenjin cluster, the (nearly) extinct Omotik, and the Datooga of Tanzania. Other Nilotic languages include Dholuo (West Nilotic) and Karimojong (East Nilotic, cf. Vossen 1982).

2 A somewhat contrived, extreme example of this is given in Creider (1989: 40) as the untranslated *ki:-ma:-ko-kas-ta-čin-i:* [DIST.PAST-NEG-PERF-'hear'-TLOC-DAT-PERFV].

3 Similar examples of so-called 'Super-Raising' have apparently been reported in Japanese among other languages (H. Koopman p.c. 1997). Also, various Bantu languages and Moroccan Arabic dialects are said to exhibit phenomena reminiscent of Kalenjin copying processes as well (J. Bresnan p.c. 1997, based on data discussed in dissertations by Carolyn Harford Perez (1985-Wisconsin) and Janet Wager (1983-MIT)).

4 Most of the data from this section has appeared previously in Anderson (1995, 1997).

5 For example this is obvious also in 'Tough'-Movement type formations in Fox (Dahlstrom 1996-ms). Because of this it is impossible to tell whether Copying-to-OBJ has occurred when there is an inanimate referent in the lower clause. Note, however, that inanimate nouns rarely exhibit high degrees of discourse salience in Fox, see Anderson (1997).

6 Generally, when the SUBJs of the two clauses are co-referential.

7 Note that two separate Copying-to-OBJ processes have occurred in this Fox sentence.

8 According to Frantz (1991), for some speakers of Blackfoot, Copying-to-OBJ has become obligatory with the higher verb 'want'.

9 Dahlstrom (1991:70) remarks that for some speakers of Plains Cree, Copying-to-Object formations differ from the corresponding constructions without Copying-to-OBJ by means of a quasi-evidential interpretation.

10 Sometimes the sentences also exhibit an optional 'raising' or movement of an overt pronominal into the verb phrase in the case of Kalenjin.

References:

Anderson, Gregory D. S. 1995. Ditransitives, Possessor Raising, Copying-to-OBJ: 'Animacy' in Morphosyntax. In *CLS 31*, pp. 1-17. Chicago: CLS.

_____ 1997. On 'Animacy Maximization' in Fox. In *International Journal of American Linguistics* vol. 63/2: 227-247.

Bender, M. Lionel. 1996. *The Nilo-Saharan Languages: A Comparative Essay.* LINCOM Handbooks in Linguistics 06. München: LINCOM EUROPA.

Bloomfield, Leonard. 1946. Algonquian. In H. Hoijer (ed.) *Linguistic Structures of Native America.* Viking Funds Publications in Anthropology 6, 85-129.

Creider, Chet A. 1989. *The Syntax of the Nilotic Languages.* Berlin: Dietrich Reimer Verlag.

_____ and Jane Tapsubei Creider. 1989. *A Grammar of Nandi.* Hamburg: Helmut Buske Verlag.

Croft, William. 1990. *Typology and Universals.* Cambridge: Cambridge University Press.

Dahlstrom, Amy. 1991. *Plains Cree Morphosyntax.* New York: Garland Press.

_____ 1993. The Syntax of Discourse Functions in Fox. In *BLS-19S: Syntactic Issues in Native American Languages.* Berkeley: BLS, pp. 11-21.

_____ 1996-ms. *The Morphology and Syntax of Fox (Mesquakie).* Unpublished ms. Chicago.

Frantz, Donald. 1978. Copying from Complements in Blackfoot. In E.-D. Cook & J. Kaye (eds.) *Linguistic Studies of Native Canada.* Vancouver: UBC Press, pp. 89-110.

_____ 1991. *A Grammar of Blackfoot.* Toronto: University of Toronto Press.

Hollis, A. C. 1909. *The Nandi: Their Language and Folklore.* Oxford: Clarendon Press.

Jake, Janice. and David Odden. 1979 Raising in Kipsigis. In *Studies in the Linguistic Sciences* 9/2: 131-56.

Nichols, John. 1980. *Ojibwa Morphology.* Harvard University Ph. D. Dissertation.

Rottland, Franz. 1982. *Die südnilotischen Sprachen.* Berlin: Dietrich Reimer Verlag.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. Dixon (ed.) *Grammatical Categories in Australian Languages*, pp. 112-71. Canberra: Australian Institute of Aboriginal Studies.

Tucker, A. N. and M. A. Bryan. 1964. *Linguistic Analyses: The Non-Bantu Languages of Northeastern Africa*. London: Oxford University Press/ International African Institute.

Vossen, Rainer. 1982. *The Eastern Nilotes. Linguistic and Historical Reconstructions*. Berlin: Dietrich Reimer Verlag.

The emergence of the unmarked pronoun: Chichewa pronominals in Optimality Theory*

Joan Bresnan
Stanford University

In Optimality Theory a grammar consists of a ranking of constraints which are (i) universal and (ii) violable. Languages differ systematically *only* in their rankings of these constraints (Prince and Smolensky 1993). The latter is a powerful theoretical principle which plays a central role in the explanatory scope of OT (Smolensky 1996a,b), its learnability (Tesar and Smolensky 1996) and its consequences for linguistic typology (e.g. Legendre, Raymond, and Smolensky 1993). It is sometimes referred to as *richness of the base.*

According to the principle of richness of the base, systematic differences in the lexical inventories of languages cannot simply be derived from language-particular constraints on lexical features or morphology. All such differences must derive from the rerankings of universal constraints. From the perspective of generative syntax, however, this consequence initially seems implausible, even absurd: after all, it has now been almost universally accepted that much of syntax derives from the lexicon, but the lexicon itself has been regarded as the residual core of what cannot be predicted. In defence of this view it is often observed that the inventory of forms present in each language reflects a contingent and individual path of historical change and areal contact. Previous OT syntax work on deriving the lexicon (e.g. Grimshaw 1995 on empty *do*, Legendre, Smolensky, and Wilson 1995 on resumptive pronouns, Grimshaw and Samek-Lodovici 1995 and Samek-Lodovici 1996 on null and expletive pronouns, and Grimshaw 1996 on Romance clitics) does not explicitly address the issues of contingency and markedness taken up here.

While the contingency of the lexicon is inescapable, both phonologists and functional linguists have recognized that linguistic inventories also reflect universal patterns of markedness and are often functionally motivated by perceptual and cognitive constraints. I will argue in support of this conclusion by showing how different inventories of personal pronouns across languages may be formally derived by the prioritizing of motivated constraints in Optimality Theory. The contingency of the lexicon—exemplified by accidental lexical gaps—then acts as a simple filter on the harmonic ordering derived by the general theory.

In what follows I will make three simplifying assumptions. First, I will assume without argument that elements which function as personal pronouns are not structurally uniform across languages, but show formal variation, as schematized in (1). The range of structures available to pronominal arguments includes the null structure (for zero or null pronouns), affixal structure on a head (for morphologically bound pronouns, also called 'pronominal in-

flections'), the structure of clitics (syntactically positioned but phonologically dependent), the structure of weak or atonic pronouns, and of (ordinary) pronouns, which can bear primary sentence accents.

(1) **Range of personal pronominal forms:**
Zero Bound Clitic Weak Pronoun

This assumption is in accordance with longstanding typologically oriented work within functional syntax (e.g. Givón 1976, 1983, 1984, 1990, 1995, Nichols 1986, Van Valin 1996) and lexical functional grammar (e.g. Mohanan 1982, Simpson 1983, 1991, Kameyama 1985, Bresnan and Mchombo 1986, 1987, Andrews 1990, Austin and Bresnan 1996, Bresnan 1995, 1996b), as well as with recent work within Optimality Theoretic syntax (Grimshaw and Samek-Lodovici 1995, Samek-Lodovici 1996). On this conception of pronominal elements, what universally characterizes a pronoun are its referential role and functions, not its syntactic category.

Second, for purposes of this initial study I will further simplify the problem by considering only the three types of pronominal forms shown in (2):

(2) **Range of pronominal forms to be derived:**
Zero Bound Pronoun

For concreteness, I will take the pronominal inventory of Chicheŵa, which includes both morphologically bound and free pronouns (Bresnan and Mchombo 1986, 1987), as the target to be derived. And third, although Chicheŵa subject inflections are markers of grammatical agreement as well as pronominality (Bresnan and Mchombo 1986, 1987), space limitations preclude an analysis of agreement and its relation to pronominal inflection here.

Marked and unmarked pronominal forms. Our goal, then, is to derive the pronominal inventory of Chicheŵa—both the analytic and the synthetic forms of pronouns—from the ranking of universal constraints within Optimality Theory. Let us begin with the reasonable assumption that we can identify personal pronouns crosslinguistically by their semantic, information structural, and morphosyntactic properties. Semantically, they have variable reference and minimal descriptive content; in information structure they may be specialized for reference to topical elements (Givón 1976, 1983, 1984, 1990: 916ff); morphologically they usually distinguish the classificatory dimensions of person (allowing for participant deixis and inclusion/exclusion relations among participants), number (singular, dual, trial/paucal, and plural), and gender (classifications into kinds) (Givón 1984: 354–5). We can abbreviate these three types of properties by the features PRO, TOP, and AGR in (3). Not all pronouns need have all these features, but these are the types of

features that identify personal pronouns crosslinguistically, and in terms of which universal optimality-theoretic constraints on personal pronouns can be stated.

(3) Crosslinguistic properties of personal pronouns:
PRO — variable referentiality
TOP — topic-anaphoricity
AGR — classification by person, number, gender

Bound and free personal pronouns can be represented in a language independent way using these feature types, as illustrated in part by (4):

(4) Feature types of bound and free personal pronouns:

$$\text{Bound: } \begin{bmatrix} \text{TOP} \\ \text{PRO} \\ \text{AGR} \end{bmatrix} \quad \text{Free: } \begin{bmatrix} \text{PRO} \\ \text{AGR} \end{bmatrix}$$

(4) represents bound pronominals as universally specialized for topic anaphoricity, and free syntactic pronouns as unmarked for this property.

The morphologically bound pronouns of Chichewa are in fact specialized for topic-anaphoric functions, as documented by Bresnan and Mchombo 1986, 1987. They are used for anaphora to a discourse topic, a crosslinguistically general pattern (Givón 1976, 1983, 1984, 1990; Lambrecht 1981, 1994).

(5) Discourse topics (Bresnan and Mchombo 1987: 768):

(a) *Fîsi anadyá ḿkângo. Á-tá-ú-dya, anapítá ku San Francîsco.*
hyena ate lion(3) he-serial-it(3)-eat he-went to S.F.
'The hyena ate the lion. Having eaten it, he went to S.F.'

(b) *Fîsi anadyá ḿkângo. Á-tá-dyá íwo anapítá ku San Francísco.*
hyena ate lion(3) he-serial-eat it(3) he-went to S.F.
'The hyena ate the lion. Having eaten it (something other than the lion), he went to S.F.'

In contrast to the synthetic pronominal in (5a), the analytic pronoun in (5b) is interpreted as referring to topics not mentioned in the previous sentence. Thus the example (5b) is bizarre, disconnected as a discourse. Within sentences, the bound pronominals are used for resumption in relative clauses and clefts, and for coreference with syntactically dislocated topic constituents, as illustrated in (6) and (7):

(6) Dislocated topics (Bresnan and Mchombo 1987: 769):

(a) *mkángó uwu fîsi a-ná-(ú)-dy-a.*
lion(3) this hyena SM-past-OM(3)-eat-indic
'This lion, the hyena ate it.'

(b)?* *mkángó uwu fîsi a-ná-dy-á (ĭwo)*
lion(3) this hyena SM-past-eat-indic it(3)
'This lion, the hyena ate it.'

(7) Resumptive relativization (Bresnan and Mchombo 1987: 769):

(a) *Ndi-ku-lír-ír-a mkángó u-méné fîsi á-ná-(ú)-dy-a.*
I-pres-cry-appl-indic lion(3) 3-rel hyena SM-past-OM(3)-eat-indic
'I'm crying for the lion that the hyena ate.'

(b)?* *Ndi-ku-lír-ír-a mkángó u-méné fîsi á-ná-dy-á (íwo)*
I-pres-cry-appl-indic lion(3) 3-rel hyena SM-past-eat-indic it(3)
'I'm crying for the lion that the hyena ate.'

The free pronominals are excluded from the syntactic and discourse environments in which a corresponding bound pronominal can be used; instead, they serve for introducing new topics or for contrastive focus.

These facts are consistent with the proposed language-independent analysis in (4), but they do not explain why the bound form is represented as more specialized in its functions than the free form. In fact, phonologically reduced pronominal forms such as bound pronominals are often taken to be the unmarked referent coding devices (Givón 1983, 1990: 916ff, 1995: 50; Comrie 1996); yet (4) represents them as marked for the property of topic anaphoricity (TOP). Moreover, the free pronominal form of Chichewa, as we have seen in the above examples, appears to be equally specialized in its non-topic-anaphoric uses. What, then, is the motivation for treating the free pronoun as unmarked for the topic-anaphoric property rather than taking it to be marked for an opposite property?

The reduced forms are indeed unmarked in the sense of having fewer morphemes or less phonological content. However, they are not unmarked in the sense of being the forms used under neutralization of oppositions. These two senses of unmarkedness are clearly distinguished in German, where *unmarkiert* refers to the value of a morphosyntactic category or feature under neutralization of oppositions, and *merkmallos* refers to the element in the paradigm having fewest morphemes or least phonological material (Bernard Comrie, p.c. June 30, 1996). The neutralization interpretation is the sense of morphosyntatic unmarkedness used in Jakobson's analysis of the Russian verb

system: "If Category I announces the existence of A, then Category II does not announce the existence of A, i.e. it does not state whether A is present or not. The general meaning of the unmarked Category II, as compared to the marked Category I, is restricted to the lack of 'A-signalization'" (Jakobson 1931[1984]: 1). Jakobson (1931[1984]: 1–2) gives the example of a Russian feminine gender noun *oslíca* 'she-ass' being the marked category used only for a female animal of the species, where the corresponding masculine gender noun *osël* 'donkey' is used for animals of both sexes. However, in a specific context of contrast the female meaning may be cancelled, leaving only the male meaning: *èto oslíca?* 'Is it a she-ass?' —*nét, osël* 'no, a donkey'.

As Bresnan and Mchombo 1986, 1987 show, the morphologically bound pronominal forms in Chicheŵa are indeed specialized for the topic-anaphoric use, while the free syntactic pronoun is a more general form: where a bound pronominal form is lacking, the free pronoun takes on the syntactic and discourse functions of the bound form, filling gaps in the morphological paradigm. This constitutes a classic case of markedness in Jakobson's 1931 sense: depending on context, the unmarked (neutral) form can be used either inclusively, subsuming the marked form, or exclusively, in opposition to the marked form.

The crucial evidence is that the restrictions on the use of the free pronouns appear only where a contrasting synthetic form exists (Bresnan and Mchombo 1987: 768–75). Thus, verbs have an optional pronominal object inflection, and the independent object pronoun in the VP is contrastive, as we see in (5)–(7). Similarly, Chicheŵa has a bound pronominal form for the preposition meaning 'with, by':

Chicheŵa (Bresnan and Mchombo 1987: 869):

(8) a. *ndí* *íye*
with/by her/him (class 1)

b. *năye* < **na + ĭye*
with/by+her/him (cl 1) with/by her/him (cl 1)

(9) a. *ndí* *íwo*
with/by it (class 3)

b. *năwo* < **na + ĭwo*
with/by+it (cl 3) with/by it (cl 3)

These contrast in the same way as the verbal object pronominals, as illustrated in (10a,b):

(10) a. *mkángó uwu ndi-na-pít-á* (*nawó*) *ku msika*
lion(3) this I-rm.pst-go-indic with-it(3) to market
'This lion, I went with it to market.'

b.?* *mkángó usú ndi-na-pít-á* (*ndí íwó*) *ku msika*
lion(3) this I-rm.pst-go-indic with it(3) to market
'This lion, I went with it to market.'

Significantly, in contexts where a bound pronominal form is lacking, the free pronoun takes on the communicative functions reserved for the synthetic forms elsewhere. For example, in contrast to the preposition *ndí* 'with, by', which has contracted pronominal counterparts (8)–(9), the preposition *kwá* 'to' occurs only with free pronouns:

(11) a. *kwá íyo*
to him (class 3)

b. **kwăyo* < *kwa + ĭyo*
to+him (cl 3) to him (cl 3)

With *kwá*, the uses of the independent pronoun subsume the uses of the contrasting bound and free pronominals elsewhere. Examples showing the free pronoun taking on the functions of the bound pronominals are given in (12), showing coreference with dislocated topics, and in (13), showing anaphora to a discourse topic.

(12) *mfúmú iyi ndi-ká-kú-neněz-a kwá* (*íyo*)
chief(3) this I-go-you-tell.on-indic to him(3)
'This chief, I'm going to tell on you to him.'

(13) *ndikufúná kuónána ndí ḿkángó wănu; mu-nga-ndí-téngere kwá* (*íwo*)*?*
I-want to-meet with lion your you-could-me-take to it
'I want to meet your lion. Could you take me to it?'

The overall picture, then, is that the morphologically bound pronominals are specialized forms reserved for topic anaphoric uses, while the free pronouns are general, neutral forms. As in Jakobson's 1931 example of the she-ass and the donkey, the free pronoun is the unmarked form: it subsumes the meaning of the marked form (the bound pronoun), but in contexts of contrast it takes on the opposite meaning. Or, to put the situation in terms of the concept of paradigm, the free syntactic pronouns can be seen to fill the functional gaps in the morphological paradigms of bound pronominal forms.

In this analysis privative features have been used to represent pronominal content. The feature TOP, for example, stands for a privative or monovalent

feature, which has only a single (the 'marked') value.[1] Such features give rise to benign ('permanent', 'inherent', or 'trivial') underspecification in the sense of Steriade 1995. With this type of representation, the meaning or function of an underspecified pronoun is not fully determinate from its featural characterization alone (cf. Frisch 1996, Reiss 1997), but depends on its relation to other pronominal elements in the paradigm. Thus our representations provide a good formal model of the Jakobsonian conception of morphosyntactic markedness, which as we see from the example of the she-ass and the donkey, allows for precisely this ambivalence in the unmarked form. The meaning or function of the unmarked pronoun depends not on its inherent features alone, but on its relation of dynamic competition with other members of the pronominal paradigm.

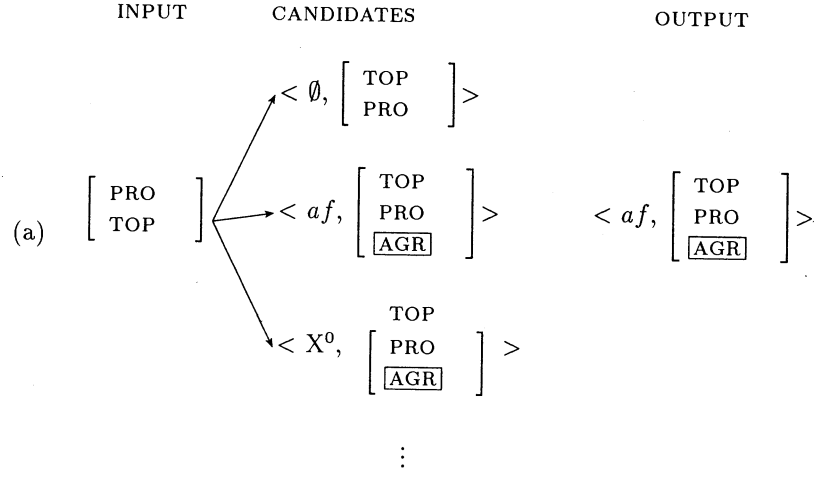
The theoretical framework. Within OT morphosyntax, then, the universal content of personal pronominals (which will be the 'input') will consist of all possible combinations of the pronominal feature types in (3), represented as feature matrices. The universal candidate set of structural analyses of pronouns will include bound and free pronominals as in (4); these are formally representable as pairings of structural analyses (such as a morphological affix *af* or a syntactic category X^0) with the functional content of pronouns represented by a feature matrix. See (14).

(14) Candidate pronoun types as structure-content pairs:

Bound: $< af, \begin{bmatrix} \text{TOP} \\ \text{PRO} \\ \text{AGR} \end{bmatrix} >$ Free: $< X^0, \begin{bmatrix} \text{PRO} \\ \text{AGR} \end{bmatrix} >$

Thus each candidate is a structural analysis (whether morphological or syntactic) of specified pronominal content. Which of the ways of structurally analyzing pronouns will appear in the inventory of a given language depends on how the candidates are harmonically ordered by the language. The harmonic ordering is induced by the strict dominance ranking of universal constraints. One candidate is more harmonic than another if it better satisfies the top ranked constraint on which the two forms differ (Grimshaw 1995, Smolensky 1996c). Crucially, the candidates need not be perfect analyses of the input; as illustrated in (15), they may overparse or underparse the input pronominal content. (Overparsing is marked with a box, underparsing by a feature outside of the matrix.)

(15) **Optimality Theory**



(b) GEN: INPUT→ CANDIDATES

(c) EVAL: CANDIDATES → OUTPUT

Are there well-defined GEN and EVAL functions that meet the specifications we have just set out for (15)? GEN must satisfy two fundamental requirements of OT: (i) the universality of the input implied by 'richness of the base' and (ii) the recoverability of the input from the output, implied by the 'containment' or 'correspondence' theories of the input-output relation (Prince and Smolensky 1993, McCarthy and Prince 1995). Because 'richness of the base' implies that the input must be universal, the syntactic GEN cannot simply be defined as mapping a set of language-particular 'lexical heads' or morphemes onto structural forms. A more abstract and crosslinguistically invariant characterization of the input is required. Because the recoverability of the input from the output is fundamental to the learnability of OT (Tesar and Smolensky 1996), the input must either be contained in the output or must be identifiable from the output by a correspondence. Hence the candidate set cannot simply consist of syntactic forms (such as strings of morphemes parsed into phrase structure trees) alone.

Both of these requirements can met by defining GEN formally as an LFG (as proposed in Bresnan 1996a and Choi 1996). This provides a mathematically well-defined correspondence between feature structures (representing language-independent content) and constituent structures (representing the variety of surface forms). The universal input can be modelled by sets of

f-structures, which provide an abstract and form-independent characterization of morphosyntactic content. The candidate set can consist of pairs of a c-structure and its corresponding f-structure, which may be matched to the input f-structure by correspondence. This definition of GEN also satisfies the basic intuition shared by many that the input to an OT syntax should provide a semantic interpretation for candidate forms: the f-structures of LFG were originally proposed as schematic, grammaticalized representations of semantic interpretations (Kaplan and Bresnan 1982[1995]), and recent work within formal semantics has validated this conception by showing how f-structures can be read as underspecified semantic structures, either Quasi-Logical Forms or Underspecified Discourse Representation Structures (Genabith and Crouch 1996).

On this conception of GEN, then, the input simply represents language-independent 'content' to be expressed with varying fidelity by the candidate forms, which carry with them their own interpretations of that content. The input f-structure corresponds to and is recoverable from the f-structure in the output pair.[2]

EVAL, as we have already discussed, consists of the following:

(16) EVAL

(i) A universal Constraint Set; constraints conflict and are violable.

(ii) A language-particular strict dominance ranking of the Constraint Set.

(iii) An algorithm for harmonic ordering: The optimal/most harmonic/-least marked candidate (= the output for a given input) is one that best satisfies the top ranked constraint on which it differs from its competitors (Grimshaw 1995, Smolensky 1996c).

The existence of an appropriate EVAL, then, reduces to the discovery of universal constraints whose ranking generates the desired inventories of pronominal forms. We further require that these constraints be *motivated.* The constraints are our next topic.

The constraints. To derive the personal pronominal inventories of English and Chichewa, we can use the relative ranking of a structural markedness constraint on pronominal candidates and the constraint(s) of faithfulness to the input.

(17) Constraints:

(a) ∅TOP (Topic is unexpressed): TOP ⊃ ∅

(b) FAITH (Faithfulness to the input): $\text{PARSE}^{\text{FEAT}}$

The faithfulness constraint (17b) is violated when a feature of the input, such as TOP, PRO, or AGR, is not analyzed by a candidate. In the present framework, this means that a violation occurs when the feature matrix of a candidate lacks the designated feature value present in the input.[3] The motivation for faithfulness in this framework is that it ensures the expressibility of the input content (Edward Flemming, p.c.).

The structural markedness constraint (17a) asserts that the least marked analysis of the topic-anaphoric property is the absence of any morphosyntactic expression at all. It is thus a constraint on the formal complexity of expression of the topic—a constraint on *merkmalhaft* forms. In one version or another, the generalization that the least marked expression of the topical element is no expression at all has been widely adopted. Givón (1984, 1990: 917) refers to it under the name 'referential iconicity'; see also Kameyama 1985. Haiman (1985b: ch. 3) regards it as an economy constraint which allows the most familiar and predictable material to be omitted (cf. Givón 1985). Recent examples include Grimshaw and Samek-Lodovici's 1995 constraint DROPTOPIC, and Van Valin's 1996 scalar representation of the relative markedness of referential coding devices with zero pronominals at the most topical extreme.

Within the present framework, we interpret the constraint (17a) as follows: we are given a universal set of candidates that we represent formally as pairs of a pronominal feature matrix and a structural analysis, whether as a bound affix ('Bound'), head of a syntactic category such as X^0 ('Free'), or some other morphosyntactic form; constraint (17a) assesses a mark to any candidate whose feature matrix contains the TOP property and whose structural analysis is nonempty. The intuition is that if a pronoun is specialized for topic anaphoricity, its unmarked expression is empty, null, zero. This will penalize the bound pronominal form compared to the free pronoun, which is unspecialized or neutral for topic anaphoricity, and it will also penalize the bound pronominal form compared to the null or zero pronoun, to which we will come directly.

(18) Interpretation of constraint (17a):

	∅TOP
Bound: [PRO, TOP, AGR]	*
Free: [PRO, AGR]	

If a language gives priority to this constraint ∅TOP over the faithfulness constraint $\text{PARSE}^{\text{TOP}}$, an instance of (17b), the result will be that violations of the zero topic constraint are worse than violations of faithfulness: in other words, it is worse to express the topic property morphosyntactically than to represent it unfaithfully in a candidate structural analysis. Since this is

true for any input (combination of pronominal content features), the marked topical pronominal inflections will be absent in such a language (all else being equal). Only the neutral free pronouns will occur in the inventory. English is such a language. (The ranking of the constraints is indicated by their left-to-right order in the tableaux columns.)

(19) Ranking for English:

(a)

	Input: [PRO, TOP]	∅TOP	FAITH
	Bound: [PRO, TOP, AGR]	*!	
⇒	Free: [PRO, AGR]		*

(b)

	Input: [PRO]	∅TOP	FAITH
	Bound: [PRO, TOP, AGR]	*!	
⇒	Free: [PRO, AGR]		

Conversely, if a language gives priority to the faithfulness constraint over the structural markedness constraint ∅TOP, it will include both bound and free pronominal forms in its inventory. For topic-anaphoric inputs, the bound pronominal will be more harmonic than the free pronoun; for non-topic-anaphoric inputs, the free pronoun will be more harmonic. Chichewa is such a language:

(20) Ranking for Chichewa:

(a)

	Input: [PRO, TOP]	FAITH	∅TOP
⇒	Bound: [PRO, TOP, AGR]		*
	Free: [PRO, AGR]	*!	

(b)

	Input: [PRO]	FAITH	∅TOP
	Bound: [PRO, TOP, AGR]		*!
⇒	Free: [PRO, AGR]		

Because of the principle that languages differ systematically *only* in their rankings of the universal constraint set, this (partial) theory makes the typological prediction that there are languages like English with free pronouns only and no bound pronominals, and languages like Chichewa with both free and bound pronominals, but no languages having only bound pronominals and lacking free pronouns. To the extent that this prediction is borne out, it provides evidence for our hypothesis that the free syntactic pronoun is the unmarked pronominal form (that is, the neutral, *unmarkiert*, form):

(21) **Markedness relation among bound and free pronoun inventories:**
Free pronouns only (English)
Both free and bound pronouns (Chichewa)
Bound pronouns only (none)

Thus far, however, we have artificially restricted the candidate set by excluding zero pronouns/null anaphors. We may formally analyze zero pronouns as the absence of structural analysis of topical pronominal content; in other words, zero pronouns lack any exponence at all (as in Mohanan 1982, Kameyama 1985, Simpson 1991). (Bound morphological inflections with pronominal content are analysed as pronominal inflections rather than zero pronominals (see Bresnan and Mchombo 1987, Austin and Bresnan 1996 for references). Where the other candidates (14) pair a feature matrix representing pronominal content with a morphosyntactic analysis (either bound morphology or a syntactic category), the zero or null pronominal pairs a pronominal feature matrix with nothing—no morphological or syntactic structure:

(22) Candidate personal pronoun types:

Zero: $< \emptyset, \begin{bmatrix} \text{PRO} \\ \text{TOP} \end{bmatrix} >$ Bound: $< af, \begin{bmatrix} \text{TOP} \\ \text{PRO} \\ \text{AGR} \end{bmatrix} >$

Free: $< \text{X}^0, \begin{bmatrix} \text{PRO} \\ \text{AGR} \end{bmatrix} >$

If nothing further were said, the zero pronoun would always be the optimal expression for topic-anaphoric pronominal content. Regardless of the ranking of our constraints (17), it would receive a higher harmonic ordering than either bound or free pronouns, as we see in (23). (The constraints are unranked in (23), as indicated by the absence of a column line seperating them.)

(23)

Input: [PRO, TOP]	∅TOP	FAITH
⇒ Zero: [PRO, TOP]		
Bound: [PRO, TOP, AGR]	*!	
Free: [PRO, AGR]		*!

There are indeed many languages that have zero pronouns and lack morphological bound pronominals or agreement morphology, e.g. Chinese, Japanese, Malayalam (Mohanan 1982) and Jiwarli (Austin and Bresnan 1996). But Chichewa is not among them. How can the different inventories of Chichewa and these languages be derived?

There is one salient difference between null or zero pronouns and morphologically bound pronominals. That is that zero pronouns have no intrinsic specification for the classificatory properties of person, number, and gender (AGR), which are morphologically distinguished in overt pronominals, both bound and free. In Chinese and Japanese, which lack verbal agreement morphology, zero pronouns can be used for various persons and numbers. The same is true in Malayalam (Mohanan p.c., November 11, 1996). In Jiwarli, Austin and Bresnan (1996: 248–50) give examples of the zero pronoun used for third person singular object, third person dual subject, first person singular subject, first person plural subject, and second person subject; Jiwarli, too, has no agreement morphology. In Warlpiri, the Auxiliary registers agreement for subject and object, but as Simpson 1991 shows, in Warlpiri sentences with nominal main predicates, the Auxiliary is optional. In such Auxiliary-less sentences, the zero pronoun is not restricted in person or number (Austin and Bresnan 1996: 241–2; Simpson 1991: 141–3).

We can therefore explain the absence of zero pronouns in some languages by means of a universal constraint stating that pronominals have the referentially classificatory properties denoted by AGR, as in (24c). This constraint can be compared to a structural constraint on feature cooccurrence in phonology, such as [voice] $\supset$ [sonorant], which plays a role in deriving markedness relations in phonological inventories (Prince and Smolensky 1993: ch. 9; Smolensky 1996b). The functional motivation for the present constraint could be that pronouns (in the unmarked case) bear classificatory features to aid in reference tracking, which would reduce the search space of possibilities introduced by completely unrestricted variable reference (Haiman 1985b: pp. 190–1).

(24) Constraints:

(a) $\emptyset$TOP (Topic is unexpressed): TOP $\supset$ $\emptyset$

(b) FAITH (Faithfulness): PARSE^{FEAT}

(c) PROAGR (Pronouns classify for AGR): PRO $\supset$ AGR

In languages like English and Chichewa, which lack zero pronouns, this constraint will dominate the zero topic constraint (24a); zero pronoun languages like Chinese, Japanese, Jiwarli, and Malayalam will have the reverse ranking.

The table in (25) shows how these constraints are interpreted with respect to our three pronominal forms:

(25) Interpretation of constraints:

		∅TOP	PROAGR
Zero:	[PRO, TOP]		*
Bound:	[PRO, TOP, AGR]	*	
Free:	[PRO, AGR]		

It is clear that the two markedness constraints conflict and disagree only on the bound and zero pronominal forms. Now whenever PROAGR dominates ∅TOP, the bound pronoun will be more harmonic than a zero pronoun. The inventories of pronominals admitted under the three rankings consistent with PROAGR ≫ ∅TOP will therefore reduce to (21). In contrast, when ∅TOP dominates PROAGR, the null pronoun will be more harmonic than the bound pronoun. Whether the null pronoun is more optimal than the free pronoun in this case depends on the relation of ∅TOP to faithfulness. The three rankings consistent with ∅TOP ≫ PROAGR yield the inventories in (26):

(26) **Markedness relation among null and free pronoun inventories:**
Free pronouns only (English)
Both free and null pronouns (Jiwarli)
Null pronouns only (none)

The constraints proposed here not only suffice to derive the pronominal inventories of head-marking languages like Chichewa, the typology they generate by rerankings explains the crosslinguistic fact that no languages contain zero pronouns or bound pronouns without also containing free pronouns. The free pronoun is the least marked pronominal form crosslinguistically. Let us now turn to the language internal distribution of pronominal forms in Chichewa, to see how the same theory also explains the emergence of the unmarked pronoun, which was originally observed by Bresnan and Mchombo 1986, 1987.

The emergence of the unmarked pronoun. The present theory predicts a general complementarity between the bound and free pronominal forms in Chichewa: the bound forms are optimal for topical input, the free forms are optimal elsewhere. This happens because the markedness of bound pronominals is submerged by the higher ranked constraint of faithfulness to the input topicality. However, in Optimality Theory a form is grammatical not when it perfectly satisfies all constraints under a given ranking, but only when it better satisfies them than its competitors. Thus in conditions where the faithfulness difference between bound and free pronominals is neutralized or

overridden, the relative unmarkedness of the free pronoun will emerge. This is what happens with pronominal objects of prepositions in Chichewa.[4] It is an instance of the 'emergence of the unmarked' (McCarthy and Prince 1994).

In many head-marking languages, pronominal inflections may appear on all heads, including prepositions or postpositions. But Chichewa has a very small set of prepositions (*ndí* 'with, by' used for instrumentals, comitatives, passive agents, and inanimate causees, *mpâka* 'until, up to', *kwá* 'to' used for datives and animate causees); *ndí* alone has an alternant for bound pronominal forms (Sam Mchombo p.c.). This appears to be a contingent property of the Chichewa lexicon, which is not derivable from morphosyntactic principles. How can we account for such a contingency within the OT framework?

In Optimality Theory the morphosyntactic inventory of a language, modelled here as pairings of structural types (e.g. $\emptyset$, *af*, X^0) with grammatical content (e.g. [PRO, AGR, TOP]), is derived by the constraint ranking. The role of the lexicon is to pair this abstractly characterized inventory with phonological representations. Thus the lexicon does not tell us what the inventory of pronominal forms of a language is; it only tells us how they are pronounced. In this way, too, we can understand how to characterize the contingency of the lexicon, such as the existence of accidental lexical gaps (Bresnan 1996a). These are elements of the inventory which are admitted by the constraint ranking, but for which there happens not to exist a pronunciation (as suggested by Edward Flemming p.c.). Thus the presence of bound pronominal inflections in Chichewa is a systematic property of the language, derived by constraint ranking; the absence of bound pronominal forms for two of the three prepositions of the language is an unsystematic property which we treat by means of lexical gaps. If we assume that inventory elements normally must be phonologically realized to be used, then the existence of accidental gaps forces consideration of competing, realizable, candidates. ('Normally' refers to those inventory elements that are paired with nonnull morphosyntactic forms such as *af* or X^0; let us call these 'expressed' inventory elements.)

The constraint that expressed inventory elements must be lexically paired with phonological realizations is stated as LEX in (27):

(27) LEX:
Expressed inventory elements must be lexically paired with phonological realizations.

We then explain the emergence of the unmarked pronoun as object of a preposition:

(28) **Emergence of the unmarked pronoun:**

	Input: ['to'(x),[PRO,TOP]$_x$]	LEX	PROAGR	FAITH	∅TOP
a	kwá+Bound ['to'(x), ...]	*!			*
b	kwá Null ['to'(x), ...]		*!		
⇒ c	kwá Free ['to'(x), ...]			*	

The bound pronoun (candidate a) fails LEX because it happens to have no pronunciation in the Chichewa lexicon; the null pronoun (candidate b) lacks agreement features. The free pronoun (candidate c) fails to parse the TOP input, but this faithfulness violation is less important than the preceding violations. Contrast this situation with the preposition *ndí* 'with, by' which has a lexically available ('pronunciable') allomorph that contracts with pronouns:

(29)

	Input: ['with'(x),[PRO,TOP]$_x$]	LEX	PROAGR	FAITH	∅TOP
⇒ a	na-+Bound ['with'(x), ...]				*
b	ndí Null ['with'(x), ...]		*!		
c	ndí Free ['with'(x), ...]			*!	

The candidates in (29) will be among the infinite set of candidates for the input in (28), and vice versa. But these candidates will incur additional marks for unfaithfulness to the input PRED value 'to(x)', as illustrated in (30):

(30) **Emergence of the unmarked pronoun (continued):**

	Input: ['to'(x),[PRO,TOP]$_x$]	LEX	PROAGR	FAITH	∅TOP
a	kwá+Bound ['to'(x), ...]	*!			*
b	kwá Null ['to'(x), ...]		*!		
⇒ c	kwá Free ['to'(x), ...]			*	
d	na-+Bound ['with'(x), ...]			*	*!
e	ndí Null ['with'(x), ...]		*!	*	
f	ndí Free ['with'(x), ...]			**!	

Again the unmarkedness of the free pronoun (c) in comparison to the bound pronominal form (d) emerges, this time from the cancellation of the higher ranking faithfulness violations.

If this theory is correct, then what appear merely to be accidental gaps in the distributional patterns of Chichewa pronominals are actually a window into the universal markedness relations among pronominal inventories across languages. Even more interestingly, we see the same kind of markedness structures that Optimality Theory has explained so successfully in phonology appearing in the domain of morphosyntax.

Notes.

*I am grateful to Bernard Comrie, Edward Flemming, Mirjam Fried, Chris Manning, Sam Mchombo, Scott Myers, and Peter Sells for invaluable discussion of the issues of markedness, Optimality Theory, and Chichewa morphosyntax addressed herein. I alone am responsible for errors.

[1]For cases where one value of an equipollent feature does not appear to be universally the 'marked' value, we may use sets of privative features (Steriade 1995, Frisch 1996), whose values are inherently incompatible. For example, a pair of privative features (e.g. TOP and FOC), could replace a binary equipollent feature (e.g. [± NEW] as in Choi 1996); the fact that a single element cannot simultaneously have both properties TOP and FOC would follow from pragmatic considerations (as Bresnan and Mchombo 1987 argue) rather than the formal opposition of ± values.

[2]Specifically, the information about 'overparsing' and 'underparsing' shown in (15) is inferrable from the candidates (14) together with the marks they incur in violation of constraints on faithfulness to the input. See below.

[3]Violations of this constraint—called MAX in the correspondence theory of input-output relations (McCarthy and Prince 1995)—account for 'underparsing'. An 'overparsing' violation occurs when a feature of the candidate matrix has no correspondent in the input; this constraint is called FILL or DEP, and is not discussed in the present study.

[4]Other instances of noncomplementarity include subject pronominals in Chichewa, which involve obligatory grammatical agreement. Agreement and related phenomena such as pronominal 'doubling' are not discussed here. For the LFG theory, see Bresnan and Mchombo 1987, Andrews 1990, Bresnan 1996b, Börjars, Chapman, and Vincent 1996 and the references cited therein.

References

Andrews, Avery D. 1990. Unification and morphological blocking. *Natural Language & Linguistic Theory* 8.507–57.

Austin, Peter and Joan Bresnan. 1996. Nonconfigurationality in Australian aboriginal languages. *Natural Language & Linguistic Theory* 14.215–68.

Barbosa, Pilar, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (eds.). To appear. *Is the Best Good Enough? Proceedings from the Workshop on Optimality in Syntax, MIT, May 19–21, 1995.* Cambridge, MA: The MIT Press and MIT Working Papers in Linguistics.

Börjars, Kersti, Nigel Vincent, and Carol Chapman. 1996. Paradigms, periphrases and pronominal inflection: a feature-based account. *The Yearbook of Morphology 1996*, 1–26. Dordrecht: Kluwer.

Bresnan, Joan (ed) 1982. *The Mental Representation of Grammatical Relations.* Cambridge: The MIT Press.

Bresnan, Joan. 1995. Morphology competes with syntax: explaining typological variation in weak crossover effects. In Barbosa, et al. (eds).

Bresnan, Joan. 1996a. Optimal syntax: Notes on projection, heads, and optimality. Draft available on-line, Stanford University: http://www-csli.stanford.edu/users/bresnan/download.html.

Bresnan, Joan. 1996b. Lexical-Functional Syntax. Stanford, CA: Stanford University Department of Linguistics MS.

Bresnan, Joan. In prep. The emergence of the unmarked pronoun II. Paper to be presented at the Hopkins Optimality Theory Workshop, Baltimore, Maryland, May 11, 1997.

Bresnan, Joan and Sam A. Mchombo. 1986. Grammatical and anaphoric agreement. *CLS 22:2. Papers from the Parasession on Pragmatics and Grammatical Theory*, 278–97. Chicago.

Bresnan, Joan and Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63.741–82.

Choi, Hye-Won. 1996. *Optimizing Structure in Context: Scrambling and Information Structure.* Stanford, CA: Stanford University Ph.D. dissertation.

Comrie, Bernard. 1996. Form and function. Conference on Future Developments in Linguistic Theory. Max-Planck-Institut für Psycholinguistik, Nijmegen, 28-30 June 1996.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen (eds). 1995. *Formal Issues in Lexical-Functional Grammar.* Stanford, CA: CSLI Publications.

Frisch, Stefan. 1996. *Similarity and Frequency in Phonology.* Evanston, IL: Northwestern University Ph.D. dissertation.

Genabith, Josef van and Richard Crouch. 1996. F-structures, QLFs and UDRSs. Proceedings of the First LFG Conference, Rank Xerox, Grenoble, France, August 26-28, 1996. Available online, Stanford University: http://www-csli.stanford.edu/publications/LFG/lfg1.html.

Givón, Talmy. 1976. Topic, pronoun and agreement. In *Subject and Topic*, ed. by Charles N. Li, 149–88. New York: Academic Press.

Givón, Talmy. 1983. Introduction. *Topic Continuity in Discourse*, ed. by Talmy Givón, 5–41. Amsterdam: Benjamins.

Givón, Talmy. 1984, 1990. *Syntax. A Functional-Typological Introduction.* Vols. I and II. Amsterdam: John Benjamins Publishing Company.

Givón, Talmy. 1985. Iconicity, isomorphism, and non-arbitrary coding in syntax. In Haiman (ed.)

Givón, Talmy. 1995. *Functionalism and Grammar.* Amsterdam: Benjamins.

Grimshaw, Jane. 1995. Projection, heads, and optimality. *Linguistic Inquiry.* To appear.

Grimshaw, Jane. 1996. The best clitic and the best place to put it. Stanford Workshop on Optimality Theory, Stanford, December 6–8, 1996.

Grimshaw, Jane and Vieri Samek-Lodovici. 1995. Optimal subjects and subject universals. In Barbosa et al. (eds.).

Haiman, John. 1985a. *Natural Syntax. Iconicity and Erosion.* Cambridge: Cambridge University Press.

Haiman, John (ed.) 1985b. *Iconicity in Syntax.* Amsterdam: John Benjamins Publishing Company.

Jakobson, Roman. 1931[1984]. Structure of the Russian verb. *Roman Jakobson. Russian and Slavic Grammar. Studies 1931–1981*, ed. by Linda R. Waugh and Morris Halle, pp. 1–14. Berlin: Mouton Publishers.

Kameyama, Megumi. 1985. *Zero Anaphora: The Case of Japanese.* Stanford, CA: Stanford University Ph.D. dissertation.

Kaplan, Ronald M. and Joan Bresnan. 1982[1995]. Lexical-functional grammar: a formal system for grammatical representation. In Bresnan, ed. Reprinted in Dalrymple et al. (eds).

Lambrecht, Knud. 1981. *Topic, Anti-Topic and Verb Agreement in Non-Standard French.* Amsterdam: Benjamins.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form. Topic, Focus, and the Mental Representations of Discourse Referents.* Cambridge: Cambridge University Press.

Legendre, Géraldine, William Raymond, and Paul Smolensky. 1993. An Optimality-Theoretic typology of case and grammatical voice systems. *BLS* 19.464–78.

Legendre, Géraldine, Paul Smolensky, and Colin Wilson. 1995. When is less more? Faithfulness and minimal links in *wh*-chains. In Barbosa et al. (eds.).

McCarthy, John J. and Alan S. Prince. 1994. The emergence of the unmarked. Optimality in prosodic morphology. *NELS 24*.333–79. Amherst, MA: University of Massachusetts GLSA.

McCarthy, John J. and Alan S. Prince. 1995. Faithfulness and reduplicative identity. *Papers in Optimality Theory*, ed. by J. Beckman, L. W. Dickey, and S. Urbanczyk. University of Massachusetts Occasional Papers 18. Amherst, MA: University of Massachusetts GLSA.

Mohanan, K. P. 1982. Grammatical relations and clause structure in Malayalam. In Bresnan (ed.), 504–89.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62.56–119.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar.* RuCCS Technical Report #2. Piscateway, NJ: Rutgers University Center for Cognitive Science.

Reiss, Charles. 1997. Unifying the interpretation of structural descriptions. *WCCFL* 15.225–39. Stanford: CSLI Publications.

Samek-Lodovici, Vieri. 1996. *Constraints on Subjects. An Optimality Theoretic Analysis.* New Brunswick, NJ: Rutgers University Ph.D. dissertation.

Simpson, Jane. 1991. *Warlpiri Morpho-Syntax: A Lexicalist Approach.* Dordrecht, Holland: D. Reidel Publishing Company.

Smolensky, Paul. 1996a. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27.720–731.

Smolensky, Paul. 1996b. The initial state and 'richness of the base' in Optimality Theory. *Technical Report* JHU-CogSci-96-4, Department of Cognitive Science, Johns Hopkins University.

Smolensky, Paul. 1996c. Generalizing optimization in OT: a competence theory of grammar 'use'. Stanford Workshop on Optimality Theory, Stanford, December 6–8, 1996.

Steriade, Donca. 1995. Underspecification and markedness. In *Handbook of Phonological Theory*, ed. by John Goldsmith, pp. 114–74. Oxford: Blackwell.

Tesar, Bruce and Paul Smolensky. 1996. The Learnability of Optimality Theory: An Algorithm and Some Basic Complexity Results. *Technical Report*, Cognitive Science Department, Johns Hopkins University, Baltimore, Maryland and Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, New Jersey.

Van Valin, Robert. 1996. Assumptions and principles of RRG. Handout. Conference on Future Developments in Linguistic Theory, Max-Planck-Institut für Psycholinguistik, Nijmegen, June 28, 1996.

Predicate clefting in Kisi

G. Tucker Childs
Portland State University

1.0. Introduction. This paper examines the focus construction of Kisi, an Atlantic language (Niger-Congo) spoken by some half a million people primarily in Guinea but also in nearby Sierra Leone and Liberia. The data come from work done in 1983-84 on the southern dialect spoken in the Foya area of Upper Lofa County, Liberia. Of particular interest is the presence of what has been known in the literature as "predicate clefting", e.g., DeGraff 1996. Its interactions and complementarity with negation, an inherently focusing construction (Marchese 1983), evince some complexity. Despite some superficial similarity, however, substantial syntactic differences exist. More similarities exist in comparison to relativization and question formation, two other constructions involving movement.

The paper begins by examining the formal features of focus, presenting the most highly grammaticized construction, "Focus with *ní*", as one of several ways an item can be focused. When a constituent is focused with *ní*, the item is fronted and the particle *ní* appears at the end of the clause in which the fronted item appears. No phonetic trace of the item appears at the place of extraction. This contrasts with the usual and expected placement of the focus marker more immediately adjacent to the item of focus, e.g., Bambara (Creissels 1991:336), Miya (Schuh 1997:137).

If the focused item is the subject (the initial position; Kisi has a "mixed" word order of SVO and SAuxOV), the only surface evidence for a focused (non-pronominal) subject is the final focus particle; pronouns, however, always show a change in case. All nominal items can be focused, including adpositional phrases and nominalized adjectives. Of a different nature, however, is what happens when verbs are focused, in what has been called the "predicate clefting" construction. A nominalized form of the verb appears in initial position while the inflected form remains *in situ*.

The approach taken by this paper is an historical-typological one attempting to resolve issues in the syntax of a West African language. Work on the phonology and morphology of Kisi shows a language falling into no single typological category. The language has a "mixed" noun class system involving both prefixing and suffixing (Childs 1983); the prosody similarly shows two types of prominence, as it perhaps changes from tone to accent (Childs 1989). The explanation for at least the last set of facts is prolonged and intimate contact with speakers of Mande languages (Childs 1995). What follows is the first analysis of a purely syntactic construction, with comments drawn from language typology and from comparative linguistics.

The implications of the Kisi facts for proposals on basic word order in Kisi and in Niger-Congo are then considered using Koopman 1983 as the organizing basis. In her discussion of several Kru languages, Koopman provides guidance as to how the "mixed" word order of Kisi should be analyzed (synchronically), and the likely historical source of the synchronic situation. Although the Kisi facts are substantially different from those found in Kru, the conclusion here is that Kisi, just as the Kru languages she discusses, is also an SOV language.

2.0. Focus constructions in Kisi. There are several formally different ways of focusing in Kisi, but in the following discussion I concentrate on the syntax of the most highly grammaticized means. The syntax of focus is relatively straightforward. When items are focused, they are fronted to initial position, with the particle *nɪ́* appearing at the end of the clause. The basic generalization is that only nominal elements can be focused. Adjectives and verbs, then, can be focused only when nominalized; items that cannot be focused are verbs, ideophones, adjectives, adverbs, adpositions, conjunctions, and other function words.

2.1. Nominal constituents. All types of nominal phrases can be fronted. In fact, a Kisi clause can consist only of a noun and *nɪ́* (as in the (a) example in (1)) although the copula *cò* may also be used, as in (b).

(1) a. *yá* *nɪ́* 'It's me.'
me FOC

b. *yá* *cò* *nɪ́*
me be FOC

I give several examples of subject focus in (2). The (a) sentence shows a proper noun being focused; (b), a pronoun; (c), a modified noun; and (d), a relativized noun.

(2) a. *wàŋgó* *có* *wàná* *sòŋmgbɔ́ɔ́ŋndó* *nɪ́*
Wango be person stingy FOC
'Wango is a stingy person.'

b. *yá* *kàándɪ́á* *ndú* *nɪ́*
I whip him FOC
'I (myself) whipped him.'

c. *yòŋ* *wɔ̀ɔ́ŋndó* *kéléŋá* *ndú* *yùùwìá* *nɪ́*
thing bad hurry him age-PL FOC
'Disreputable behavior has caused him to age prematurely.'

d. *nyè* *[wànà* *mándàŋ]-ó* *có* *còŋgùláŋ* *nɪ́*
thing people careful-REL be gossips FOC
'What one has to be wary of is gossips.'

Noun phrases in object position can equally well be focused.

(3) a. *mààlóŋ ó có cùùcúúwó nɪ́*
rice he AUX sow FOC
'It's the rice he's sowing.'

b. *wànàà pèèlaŋ pɪ́lá ŋ̀ cà núà nɪ́*
people swagger one you see thus FOC
'These are just swaggerers that you see.'

Adpositional phrases can also be focused with *nɪ́*.

(4) a. *lé kɔ̀láŋ ŋ̀ hùŋ nɪ́*
for going you come FOC
'Did you come just to leave?'

b. *ò bɔ́lɔ́ɔ́ bèŋgú ndá cɔ̌l yá ndú nɪ́*
to banana under they bury me him FOC
'Under the banana tree they buried him for me.'

2.2. Nominalized constituents. Adjectives can be focused only if they are preceded by the pronoun of the noun they modify, the same form used when they appear as predicate adjectives. The first example shows such a construction with the *o*-class subject pronoun *ò*; the second features an adjective preceded by the *ŋ*-class pronoun *ŋ̀*. Note that both of these adjectives appear in copular constructions, i.e., involving the copula *co*. Adjectives cannot be focused when they directly modify a noun.

(5) a. *ò hùŋnǎŋ kó ó có nɪ̂ŋ nɪ́*
PRO come indeed he be now FOC
'He's really ready to come now.'

b. *ŋ̀ fùlùlǎŋ mààlóŋ mûŋ cò nɪ́*
PRO parboil rice this be FOC
'Is this rice parboiled?'

Adjectives have this form in predicate-adjective constructions, as illustrated in (6), the non-focused form of (5a).

(6) *ò* *có* *nɪ̂ŋ* *ò* *húŋnŋ*
PRO be now PRO come
'He's now prepared to come.'

That it is only nominal constituents that can be focused is even more dramatically demonstrated with finite verbs. The morphosyntax of the focus construction is superficially quite different for verbs: the finite form of the verb is left behind in its original spot as a non-finite form of the verb is fronted to the focus position, as seen in (7). Verb copying with front focusing is quite common in West Africa, e.g., Yoruba, Akan, Temne (Gilman 1986:39) and in Caribbean creoles. Koopman 1983 discusses the same phenomenon in several Kru languages, where it is called "predicate clefting", and Lefebvre 1992 discusses it in Fɔ̀n and other, related languages, as well as in Haitian Creole. I give rather non-idiomatic glosses for such constructions to clarify the syntax of the construction.

(7) a. *pùɛ́ŋ-ndáŋ* *yá* *púɛ́ŋ* *nɪ́*
forget-NOM I forgetFOC
'It's forgetting that I did.'
b. *yòù-wó* *yá* *yóú* *ndú* *nɪ́*
lend-NOM I lend him FOC
'It's lending to him I did.'
c. *kpùwà-á* *ò* *kpúwá* *yá* *ndú* *ó* *bà* *nɪ́*
grab-NOM he grab me him to hand FOC
'It's grabbing he did to me.'

That it is the nominalized form that is fronted is shown by the different nominalizing suffixes ("NOM"), realized differently for each verb in (7), respectively, *-láŋ* (phonetically realized as [-ndáŋ]) (a), *-ó* [wó] in (b), and *-ó* [-á] in (c). These forms can be compared to the finite forms (without the suffix) appearing later in each sentence. Note also that it is not the full VP that is fronted, only the nominalized form in the verb with no complements.

In several tenses Kisi verbs consist of two parts, the first of which can be identified as an auxiliary, and the second as something like a participle, identical in form to the nominalized form of the verb. In such compound constructions the auxiliary carries all of the information as to tense. The position of the object is now between the auxiliary and the non-finite verb, thus contrasting with the normal SVO order when there is no auxiliary. The meaning of these compound tenses is most often something like 'progressive' (there is a 'future' or 'irrealis' as well), not surprising in light of the fact that cross-linguistically the progressive is the most widely-cited aspectual context for OV constructions (e.g., Fabb 1992:5).

(8) SVO *ɪ́ béɪ́ ndú* 'I hurt him.'
I hurt him
SAuxOV *ɪ́ có ndú bèìyó* 'I am hurting him.'
I AUX him hurt

Kru languages also display this split word order, and on the basis of several movement rules, the OV order has been identified as basic (Koopman 1983). The Kisi verb movement rule, however, does not pattern in exactly the same way.

When the non-finite part of a compound verb is focused, there is no change in the verb and no verb form is left behind (represented by the "*Ø*" in the Kisi line of (9)). Here the focused verb form is *pìsúltáŋ* (second clause). The finite auxiliary *wa* (*wé* when negated) immediately precedes the *Ø*.

(9) *à wé cìèé lé pìsúltáŋ ndá wà Ø nɪ́*
they AUX fight NEG play they AUX FOC
'They weren't fighting, it's playing they were doing.'
Cf. *ndá wà pìsúltáŋ*
they AUX play
'They were playing.'

It is clear that it is the inflection that causes a different treatment of the lexical verb, i.e., the fact that the verb is now nominal.

The example in (10) shows that no verbal arguments are fronted with the nominalized form. The object *bùŋgàŋ mùŋ* remains in place after the auxiliary.

(10) *fèfèñáá ó có bùŋgàŋ mùŋ nɪ́*
staying PRO AUX portions these FOC
'He's lingering for (to get) those portions (of food).'

This set of facts supports the constituency of non-subject arguments with the finite verb rather than with the lexical (non-finite) verb. (The phonology of such constructions, namely, cliticization, also support this contention.) They also suggest a relatively weak bond between the non-finite verb and the rest of the verb phrase.

The focus particle *nɪ́* can be used with the copula (cf. other examples in (1)).

(11) *tàmbà béɛ́ kpá có béɛ́ wánà wɔ́ɔ́ŋndó nɪ́*
Tamba even truly COP even person bad FOC
'It is truly Tamba himself that is a bad person.'

bùùlùùláŋ kpósŏŋ lá có ó kóbó lá láŋ nɪ́
holes many PRO COP to can PRO those FOC
'There are all sorts of holes in those cans.'

Focus constructions also provide evidence for evaluating the integrity of the subject-finite verb unit, one which seems close on the basis of tonal evidence. When finite verbs are focused with *nɪ́*, a non-finite form is fronted without the subject, and the finite form of the verb remains in place. In all cases the subject pronoun remains behind, just as do objects. In (12a) the first word *cùlùkàá* shows that it is non-finite by its tone pattern; it has a high tone only on the final nominalizing suffix *-á* while the finite form *cùlúká* has the LH pattern of the Perfect (and no suffix). Similar facts are displayed in (12b).

(12) a. *cùlùkàá fɔ́ndàá hóò cùlúká téɪ́ nɪ́*
slippery place this slippery IDPH FOC
'This place is exceedingly slippery.'
b. *wòó ó wó yá nɪ́*
fear he fear me FOC
'It's afraid that he is of me.'

Thus the focus construction when used for verbs shows the centrality of the <u>inflected</u> verb to the clause and the integrity of the bonds with its arguments.

As was noted above, the subject pronoun can be extracted and focused, just as can any other noun phrase. When subject pronouns, e.g., *ɪ́*"I' in (13), are focused with *nɪ́*, they change their case to the objective (*yá*).

(13) *ɪ́ káándɪ́ándú* — *yá kàándɪ́ándú nɪ́*
I whip him — I whip him Foc
'I whipped him.' — 'It's me that whipped him.'

The fact was noted above that when subjects are focused with *nɪ́*, they are linked more closely to the verb than subjects focused without *nɪ́*.

3.0. Focus and negation. In a pragmatic sense the negative and focus particles are quite similar. Negation has inherent focus comparable to that produced by the focus particle, and there is some implied contrastiveness or negation in the focus construction. Such shared functionality is illustrated by the pair of sentences in (14a). The second example shows the same complementarity with a verb.

(14) Pragmatic similarity of *lé* and *nɪ́*

a. *tàndǎŋ nɪ́* — pubic-hair FOC — 'It's pubic hair.'
ò có má tándà lé — it be PRO pubic-hair NEG — 'It's not pubic hair.'

b. *à fèèyáá lé wàŋndá wàà nɪ́*
they brush NEG people plow FOC
'They weren't brushing; the people were plowing.'

The two particles can co-occur, but not if focusing on the same constituent, as might be suggested by the example in (15). In the first clause the focus is on the object *kákáá* 'stinginess', and in the second on *nyè kòŋ* 'that thing', which refers to the entire preceding clause (the reason why no one wants to work for Tamba). The focus particle *nɪ́* appears after or outside the Negative marker *lé* since the focused item functions as topic rather than as a syntactic constituent of the intervening clause.

(15) Clauses with *nɪ́* and *lé*

kákáá tàmbà nɔ́ téɪ́ nɪ́
stingy Tamba have much FOC
nyè kòŋ wànáwálɪ́ wɪ́ù ndú ɔ́ bà lé nɪ́
thing that laborers stay him to hand NEG FOC
'Tamba is so incredibly stingy that no one stays working with him.'

What is vastly different between the two is the syntactic object of focus. Negation affects the verb or INFL, while focus is restricted to noun phrases. That negation is intimately involved with the verb is clearly shown by the fact that it is doubly marked. In addition to the clause-final particle, verbs undergo segmental and tonal changes. Furthermore, if a full S is to be focused, it must have a complementizer, such as *mèé* or *màà* in the examples below.

(16) Complement clauses focused

màà sìèléɪ́yómál kpùndɔ́ɔ́ nɪ́
COMP laziness drop trouble FOC
'It's that his laziness caused the trouble.'

mèɛ́ náá ŋ́ sí´m núáà yɛ́ bɛ́ɛ́ nɪ´ ŋ́ kúà mì ŋ́ cúá
COMP us we stand thus PRT indeed FOC we go and we take
'It's because we indeed stand this way that we must go and steal.'

Thus *nɪ´* can focus only nominal constituents; this fact is confirmed by its comparison to the negative particle *lɛ́*, which focuses non-nominal elements.

Although a negated clause cannot be focused (without a complementizer), a focused construction can be negated. This involves a rather circuitous means involving a cleft construction much like that found in English.

When a focused construction is negated, the product is bi-clausal (the (a) sentence in (17)), the new order differing considerably from its affirmative counterpart (b). This construction still relies on fronting for focus but now the fronted element appears with a full negated clause. Simple negation of the focused element, the (c) sentence, is ungrammatical.

(17) a. *ò có sàà lɛ́ ò có lɛ́*
PRO be Saa NEG PRO be NEG
'It's not Saa.'
b. *sàà ò cò nɪ´*
Saa (PRO be) FOC
'It's Saa.'
c. * *sàà ò cò nɪ´ lɛ́* (* *sàà ò cò lɛ́ nɪ´*)

In a naturally occurring sentence such as (18), the subject is focused and negated. Since *nɪ´* clauses cannot be directly negated, once again a doubly negated construction is used and *nɪ´*disappears, where only one item, the focused item, is logically negated. The first clause negates the item of focus; the second clause represents the clause from which it was extracted, albeit also negated. The noun *wèlâŋndáŋ* is the subject of the second clause, its pronoun *lá* remaining behind.

(18) *ò có wèlâŋndáŋ ndâŋ lɛ́ lá hîŋ náá hɔ̀llá lɛ́*
it be resemblances these NEG PRO spin us face NEG
'It's not these resemblances that confuse us.'

If (18) were simply focused, the order would be as in (19).

(19) *wèlâŋndáŋ ndâŋ lá hîŋ náá hɔ̀llá nɪ´*
resemblances these they spin us face FOC
'It's these similarities that confuse us.'

If (18) were simply negated, the order would be as in (20).

(20) *wɛ̀lâŋndáŋ ndâŋ lá hîŋ náá hɔ̀llá lé*
resemblances these PRO spin us face NEG
'These similarities don't confuse us.'

The example in (21) shows a focused object being negated. Here the object of *sɔ̀ɲóó* (*cáŋ-ndɔ̀* 'his crying') is removed and fronted, just as in *nɪ́* clauses, but instead of appearing isolated in a noun phrase, it now appears in a full clause with the copula *co* and the negative *lé*.

(21) *ò có cáŋ ndɔ̀ lé ndá cò sɔ̀ɲóó lé*
it be crying his NEG they AUX take-out NEG
'They are not holding a wake for him.'

In the (a) example of (22), a focused adjective is negated. The nominalized adjective *ó bîŋgɪ́ɪ́náá* would normally appear right after the copula *co*, as in sentence (b). If focused with *nɪ́*, it would appear at the beginning of the clause, as in (c). That it is possible to negate the adjective without it being focused is shown in the (d) sentence.

(22) a. *ò có ó bɪ̂ŋgɪ́ɪ́náálé ó có lé*
he be he short NEG he be NEG
'It's not very short that he is.'
b. *ò cò ò bɪ̂ŋgɪ́ɪ́náá*
he be he short
'He is very short.'
c. *ó bɪ̂ŋgɪ́ɪ́náá ò cò nɪ́*
he short he be FOC
'It's very short he is.'
d. *ò có ó bɪ̂ŋgɪ́ɪ́náálé*
he be he short NEG
'He is not very short.'

Double negation with focus is found elsewhere in Kisi's genetic group, e.g., in closely related and geographically nearby Temne, where similar structures have been used to motivate a multiclausal analysis of focus constructions (Hutchinson 1989:9), the same analysis that can be applied to Kisi.

Focus with *nɪ́* is clearly an NP-movement rule which, despite its functional similarity to negation, exhibits dramatic structural differences. Negation is verbal in its orientation while focus is a nominal process.

4.0. Basic word order. In this section I discuss the reasons for seeing basic word order in Kisi as SOV in much the same way that SOV word order has been seen as basic for several Kru languages and for Niger-Congo in general.

The general consensus is that Proto-Niger-Congo had (S)OV word order. For example, "every branch of Niger-Congo shows morphological evidence of the safest kind ... in support of the hypothesis that an older OV syntax must be indeed reconstructed for the family as a whole" (Givón 1979:221). Kay Williamson reaches the same conclusion: Proto-Niger-Congo was strictly SOV (with serial verbs) (Williamson 1986). These observations would recommend a bias towards seeing the mixed system of Kisi as being basically SOV, especially in light of arguments for Kru languages presented in Koopman 1983.

In an extensive look at two Kru languages of Ivory Coast (Gbadi and Vata), Koopman 1983 identifies facts similar to Kisi's and uses them to motivate a rule of verb movement and to reconstruct the proto-language as SOV. Just as with Kisi the languages she discusses have a mixed SVO / SAuxOV system, which state has been considered by others to be "unstable" but her analysis establishes considerable underlying regularity (Koopman 1983:99) and harmony with the Niger-Congo pattern.

Language-internal arguments include the following. (Adjectives are a non-projecting lexical category (vs. Kisi)).

(23) Some arguments for basic OV word order in Kru (Koopman 1983:41-99)

1) Word order in tensed clauses with no auxiliary are "exceptional"; all other clause types exhibit complement-verb word order.
2) Many processes require adjacency between V and certain elements, where the basic position of the verb must be last in VP.
3) Projections of other lexical categories, N and P (adjectives are non-projecting), both exhibit complement-head order.

Koopman's motivation for seeing SVO order as derived comes from Case Theory, the verb moving into the Aux spot just when there is no Aux segmentally present, placing "the verb in the INFL node of a tensed clause" (Koopman 1983:98). In more technical terms, she sees the verb movement rule as a way of "saving" a base-generated structure that violates the Case filter, much as NP movement does, e.g., for passive constructions such as $[_{NP}\ \ ]$ *was killed John.* Such similarities to NP movement (there are others) moves her to call such movement "NP-type of V-movement" (Koopman 1983:139-141). In (24) I give some more specific examples of Koopman's evidence for her analysis. Each point is followed by a comment on the same phenomena in Kisi.

(24) Evidence for basic OV word order in Kisi

1. When the verb appears in non-final position, it can be explained as a result of movement:
   a. NP-type of V-movement: Repairing violations of the Case Filter (no case assigner for subject).
   b. Parallels between predicate clefting and WH-movement.

Kisi: Exact parallels in both cases. In addition Kisi verbs are nominalized when fronted (and show no verbal inflections when final).
2. Gerunds, infinitives, and other verbal constructions are verb-final. Higher arguments precede, e.g., patient, lower, e.g., locative, follow.
Kisi: The same facts obtain in Kisi.
3. Other lexical categories with maximal projections are also head final.
Kisi: No other Kisi lexical categories are (exclusively) head final. Nouns are always head-initial; prepositions can appear before, after, and around their objects, although head-initial seems to be the basic position (Smallwood 1996).
4. PPs (which are head-initial) cannot be related to NPs, as is done in many West African languages.
Kisi: PPs in final position (PostPs) can be transparently related, but not PrePs.
5. Adjacency arguments: Particle verb constructions, idioms.
Kisi: not applicable or not known.

The one striking fact to emerge from this comparison is the contrast in headedness between Kru and Kisi lexical projections. At the present time I have no explanation for the fact except that we would probably want to interpret it as an illustration of the fact that headedness may be a separate setting for word categories. Above the word the language shows a general tendency for final marking, i.e., the focus particle, the marking of relativization, etc., but there is also a preference for marking INFL, i.e., the question particle.

Nonetheless I believe this comparison of Kisi with the Kru languages as analyzed by Koopman furnishes insights into the basic word order of Kisi and suggests a greater attachment of the language to the Niger-Congo stock than has been previously thought. It may turn out that many of the other SVO languages within Atlantic may be analyzed in this way.

References

Childs, G. T. 1983. Noun class affix renewal in Southern West Atlantic. In *Current Approaches to African Linguistics 2*, ed. J. Kaye, H. Koopman, D. Sportiche, A. Dugas, pp. 17-29. Dordrecht, Holland and Cinnaminson, NJ: Foris.

Childs, G. T. 1989. The extra-high tone of Kisi: just another tone or a new system of prominence? In *Current Approaches to African Linguistics 5*, ed. P. Newman, R. D. Botne, pp. 141-55. Dordrecht, Holland and Cinnaminson, NJ: Foris.

Childs, G. T. 1995. Tone and accent in the Atlantic languages: an evolutionary perspective. In *The Complete Linguist. Papers in Memory of Patrick J. Dickens*, ed. A. Traill, R. Vossen, M. Biesele, pp. 195-215. Köln: Rüdiger Köppe Verlag.

Creissels, D. 1991. *Description des langues négro-africaines et théorie syntaxique* Grenoble: Éditions littéraires et linguistiques de l'Université Stendhal - Grenoble 3, E.L.L.U.G.

DeGraff, M. F. 1996. Creole languages and parameter-setting: A case study using Haitian Creole and the pro-drop parameter. In *Creole Languages and Language Acquisition*, ed. H. Wekker, pp. 65-105. Berlin and New York: Mouton de Gruyter.

Fabb, N. 1992. Reduplication and object movement in Ewe and Fon. *Journal of African Languages and Linguistics* 13:1-40.

Gilman, C. 1986. African areal characteristics: Sprachbund, not substrate? *Journal of pidgin and creole languages* 1:33-50.

Givón, T. 1979. Language typology in Africa: a critical review. *Journal of African Languages and Linguistics* 1:199-224.

Hutchinson, L. G. 1989. Emphasis and mood in Temne. *Manuscript, Department of Linguistics, University of Minnesota.*

Koopman, H. 1983. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht, Holland and Cinnaminson, NJ: Foris.

Lefebvre, C. 1992. Towards a typology of predicate cleft languages. *The Journal of West African Languages*. 22:53-62.

Marchese, L. 1983. On assertive focus and the inherent focus nature of negatives and imperatives: evidence from Kru. *Journal of African Languages and Linguistics*. 5:115-129.

Schuh, R. G. 1997. *A Grammar of Miya* Berkeley and Los Angeles: University of California Press.

Smallwood, C. 1996. The structure of Kisi: syntax assessment. Manuscript, Department of Linguistics, University of Toronto.

Williamson, K. 1986. Niger-Congo: SVO or SOV? *Journal of West African Languages* 16:61-82.

The auxiliarization of *re* 'say' in Setswana

Denis Creissels
Laboratoire Dynamique Du Langage (CNRS & Université Lumière Lyon 2)

1. *Re* 'say', a verb with several syntactic properties unique among Setswana verbs. The Setswana verb **re** [rɪ̀] 'say'[1] is in many respects unique. Morphologically, **re** has the possibilities of variation characterizing the regular verbs of Setswana, but in these variations, it shows several irregularities: in its tonal behavior (**re** generally has the tonal behavior of a lexically toneless verb, but in the perfect tense, it may optionally behave like a lexically H-toned verb or like a lexically toneless verb), in its final vowel (in the perfect tense, it has the regular ending **-ile**, but in the other tenses, it invariably ends with **-e** [ɪ], whereas regular verbs end with **-e** [ɪ] in certain tenses only and take an ending **-a** [a] or **-e** [ɛ] in other tenses)[2] and in the way its initial consonant alternates (in Setswana, **r** normally alternates with **th** in contexts triggering the alternation known as 'consonant strengthening', but the **r** of **re** alternates with **t**, and this alternation occurs in contexts in which the initial consonant of regular verbs is not strengthened). Syntactically, **re** combines with a subject NP in the same way as other Setswana verbs do, but in its relationship with its other arguments, it has properties that distinguish it from any other Setswana verb.

Re is the only Setswana verb that can be directly followed by a complement clause showing the same form as an independent declarative or hortative sentence – ex. (1); the other verbs of saying, thinking, etc. require the use of a complementizer which is in fact the infinitive or a participial form of the verb **re** – ex. (2).

(1a) **Ba tlaa tla kamoso** / bátɬáátɬà kámɔ̃̀sɔ́ / They will come tomorrow

(1b) **Ba rile [ba tlaa tla kamoso]** / bárìlé bátɬáátɬà kámɔ̃̀sɔ́ / They said they would come tomorrow / **ba rile** [bárìlé] is the perfect of **re** with a SM of cl 2

(2a) **O ya ko Kanye** / ɔ̃́jà kó kàɲɛ́ / He is going to Kanye

(2b) **O re boleletse [gore [o ya ko Kanye]]** / ɔ̃́rìbɔ̃́lélétsí χɔ̃́rì ɔ̃́jà kó kàɲɛ́ / He told us that he was going to Kanye / **o re boleletse** [ɔ̃́rìbɔ̃́lélétsí] is the perfect of the applicative form of **bolela** 'say, report' with a SM of cl 1 and an OM of 1st ps pl; the complementizer **gore** is morphologically the infinitive of **re** 'say'

Re cannot take a NP as its complement: apart from clauses and ideophones, its only possible complements are **eng?** 'what?', **jaana** 'like this' or **jalo** 'like that'.

Re is also the only Setswana verb that freely combines with the so-called ideophones; the combination of **re** with an ideophone constitutes a kind of compound syntactically equivalent to a verb form, the grammatical elements of a verb form (subject marker, object marker, TAM markers) being all attached to **re**, whereas the lexical meaning and the argument structure of such a compound are entirely determined by the ideophone, as in ex. (3).

In addition to that, **re** is monosemous only when followed by a clausal complement; when followed by **eng?** 'what?', **jaana** 'like this' or **jalo** 'like that', depending on the context, its subject may represent not only a person saying something, but also a person to whom something happens – ex. (4).

Re presents also the following anomaly: it is employed in the -a consecutive[3] with the meaning normally carried by the present – ex (5).

(3) **O ne a tsaya mmidi a o re goro! fa fatshe** / ɷ́nè àtsájá m̀mídí àɷ́rí χɷ̀rɷ̀ fá fàtshí / He then took the maize and poured it out on the ground / **goro** [χɷ̀rɷ̀] is an ideophone denoting the pouring out of solid contents; **a o re** [àɷ́rí] is the a-consecutive of **re** with a SM of cl 1 and an OM of cl 3 representing **mmidi** 'maize'

(4) **Ba rile eng?** / bárìlé ì̀ŋ́ / What did they say?, or What happened to them? (depending on the context) / **ba rile** [bárìlé] is the perfect of **re** with a SM of cl 2

(5) **Ngaka ya re eng?** / ŋàkà jàrì ì̀ŋ́ / What is the doctor saying? / **ya re** [jàrì] is the a-consecutive of **re** with a SM of cl 9, i.e. a form which would normally be expected to mean 'and then (s)he$_{cl.9}$ said', and not '(s)he$_{cl.9}$ is saying'

The aim of this paper is to analyze a construction characterized by the occurrence of grammaticalized forms of **re** whose precise status calls for a discussion.

2. Presentation of the data analyzed in this paper. Setswana sentences carrying a meaning identical to that of biclausal English sentences 'When S_1, S_2' may follow a pattern that can be provisionally described as a sequence of three positions labeled X, Y and Z:
– position X is occupied by an inflected form of the verb **re**;
– in the English translation, Y corresponds to clause S_1 in a construction 'When S_1, S_2'; its predicate is in a so-called participial form, i.e. in a form typically employed as the predicate of dependent clauses of time;
– in the English translation, Z corresponds to clause S_2 in a construction 'When S_1, S_2', with however an important difference: the Setswana equivalent of an English sentence with a negative clause S_2 may be a sentence with a positive clause in position Z, the negation being then expressed by the inflected form of **re** in position X – ex. (7), (10), (13), (15) & (17).

Morphologically, the inflected form of **re** in position X includes a SM, but this SM has no meaning of its own, since it may equally be an invariable SM of class 9 or the copy of the SM of the following verb, without any difference in the meaning. It is interesting to observe that this SM is tonally irregular. In a number of tenses, Setswana has a tonal contrast between low-toned SMs in the 1st or in the 2nd person and high-toned SMs in the third person[4]; in these tenses, the invariable SM of class 9 included in the inflected form of **re** in position X in the sentence pattern analyzed here may optionally have a L tone (and this is the only context in which this irregularity is observed). We shall return to this later.

Ex. (6) to (17) illustrate this sentence pattern.[5]

(6) [$_X$ **E rile**] [$_Y$ **ba bangwe ba ya kwa masimong**] [$_Z$ **ene a sala mo gae**] / **ìrílè** bá bàŋwí bájá kwá màsímɷ́ŋ̀ ènɛ́ **àsálà** mó χâì / When the others went to the field, (s)he remained at home / **e rile** [ìrílè] is the perfect positive of **re** with an expl SM of cl 9, and **a sala** [àsálà] is the a-consecutive of **sala** 'remain' with a SM of cl 1

(7) [X **Ga e a re**] [Y **ba bangwe ba ya kwa masimong**] [Z **ene a sala mo gae**] / **χàíàrí** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ɛ̀nɛ́ **àsálà** mó χáì / When the others went to the field, (s)he did not remain at home / **ga e a re** [χàíàrí] is the perfect negative of **re** with an expl SM of cl 9, and **a sala** [àsálà] is the a-consecutive of **sala** 'remain' with a SM of cl 1

(8) [X **Ya re**] [Y **ba bangwe ba ya kwa masimong**] [Z **ene a sala mo gae**] / **jàrì** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ɛ̀nɛ́ **àsálà** mó χáì / ... And when the others went to the field, (s)he remained at home / **ya re** [jàrì] is the a-consecutive of **re** with an expl SM of cl 9, and **a sala** [àsálà] is the a-consecutive of **sala** 'remain' with a SM of cl 1

(9) [X **E tlaa re**] [Y **ba bangwe ba ya kwa masimong**] [Z **ene a sale mo gae**] / **ìtɬààrì** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ɛ̀nɛ́ **àsálì** mó χáì / When the others go to the field, (s)he will remain at home / **e tlaa re** [ìtɬààrì] is the future positive of **re** with an expl SM of cl 9, and **a sale** [àsálì] is the e-consecutive[6] of **sala** 'remain' with a SM of cl 1

(10) [X **E tlaa se ke e re**] [Y **ba bangwe ba ya kwa masimong**] [Z **ene a sale mo gae**] / **ìtɬààsìkí ìrí** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ɛ̀nɛ́ **àsálì** mó χáì / When the others go to the field, (s)he will not remain at home / **e tlaa se ke e re** [ìtɬààsìkí ìrí] is the future negative of **re** with an expl SM of cl 9, and **a sale** [àsálì] is the e-consecutive of **sala** 'remain' with a SM of cl 1

(11) [X **E re**] [Y **ba bangwe ba ya kwa masimong**] [Z **ene a sale mo gae**] / **ìrí** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ɛ̀nɛ́ **àsálì** mó χáì / ... And when the others go to the field, (s)he usually remains at home / **e re** [ìrí] is the e-consecutive of **re** with an expl SM of cl 9, and **a sale** [àsálì] is the e-consecutive of **sala** 'remain' with a SM of cl 1

(12) [X **E (a) re**] [Y **ba bangwe ba ya kwa masimong**] [Z **ene a sale mo gae**] / **ì(à)rì** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ɛ̀nɛ́ **àsálì** mó χáì / When the others go to the field, (s)he remains at home / **e (a) re** [ì(à)rì] is the present positive[7] of **re** with an expl SM of cl 9, and **a sale** [àsálì] is the e-consecutive of **sala** 'remain' with a SM of cl 1

(13) [X **Ga e re**] [Y **ba bangwe ba ya kwa masimong**] [Z **ene a sale mo gae**] / **χàírí** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ɛ̀nɛ́ **àsálì** mó χáì / When the others go to the field, (s)he does not remain at home / **ga e re** [χàírí] is the present negative of **re** with an expl SM of cl 9, and **a sale** [àsálì] is the e-consecutive of **sala** 'remain' with a SM of cl 1

(14) [X **E ka re**] [Y **ba bangwe ba ya kwa masimong**] [Z **ene a sala mo gae**] / **ìkárì** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ɛ̀nɛ́ **àsálà** mó χáì / When the others go to the field, (s)he may remain at home / **e ka re** [ìkárì] is the potential positive of **re** with an expl SM of cl 9, and **a sala** [àsálà] is the a-consecutive of **sala** 'remain' with a SM of cl 1

(15) [$_X$ **E ka se re**] [$_Y$ **ba bangwe ba ya kwa masimong**] [$_Z$ **ene a sala mo gae**] / **ìkásìrí** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ènɛ́ **àsálà** mó χáì / When the others go to the field, (s)he may not remain at home / **e ka se re** [ìkásìrí] is the potential negative of **re** with an expl SM of cl 9, and **a sala** [àsálà] is the a-consecutive of **sala** 'remain' with a SM of cl 1

(16) [$_X$ **E re**] [$_Y$ **ba bangwe ba ya kwa masimong**] [$_Z$ **ene a sale mo gae**] / **írì** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ènɛ́ **ásálè** mó χáì / When the others go to the field, (s)he must remain at home / **e re** [írì] is the subjunctive positive of **re** with an expl SM of cl 9, and **a sale** [ásálè] is the subjunctive of **sala** 'remain' with a SM of cl 1

(17) [$_X$ **E se re**] [$_Y$ **ba bangwe ba ya kwa masimong**] [$_Z$ **ene a sale mo gae**] / **ísìrí** bá bàŋwí bájá kwá màsímɔ́ŋ̀ ènɛ́ **ásálè** mó χáì / When the others go to the field, (s)he must not remain at home / **e se re** [ísìrí] is the subjunctive negative of **re** with an expl SM of cl 9, and **a sale** [ásálè] is the subjunctive of **sala** 'remain' with a SM of cl 1

The position occupied by the inflected forms of **re** in this construction is apparently identical to that of **when** in the English translation, which suggests that they should be analyzed as taking on, in relation to the clause in position Y, the function of complementizer. But this analysis faces serious objections, and the precise status or **re** in this construction is not easy to establish. The way Cole treats it in his reference grammar of Setswana[8] clearly reveals the difficulty he had in making a decision on this issue: he describes this use of **re** in the chapter devoted to what he calls 'special verb tenses', which implies analyzing the forms of **re** in this construction as fulfilling the auxiliary function, but at the same time he indicates that they "behave very much like conjunctives" and that they "constitute the main verb".

The analysis put forward in this paper is that, in the present state of Setswana, the status of the inflected form of **re** occupying position X in this sentence pattern cannot be established in an entirely satisfying way in a strictly synchronic framework; the inflected form of **re** in position X must generally be analyzed as an auxiliary, but its behavior differs in some respects from that of the most typical Setswana auxiliaries, and we shall examine a variant of this construction in which the inflected form of **re** in position X does not behave as an auxiliary anymore and can be reanalyzed as a complementizer introducing a subordinate clause of time.

3. Evidence against analyzing the inflected form of *re* occurring in initial position in the sentence pattern illustrated by ex. (6) to (17) as a complementizer introducing a subordinate clause of time. In this section I examine the possibility of an analysis according to which the main predicate is the verb of the clause in position Z and X is the complementizer introducing the embedded clause Y, and I present observations showing that, if we leave aside a variant of this construction to which we will return in section 8, the analysis of the inflected form of **re** in position X as a complementizer of the clause in position Y must be abandoned.

The main evidence against this analysis is that position Y in such a sequence cannot be left empty, but is not necessarily occupied by a dependent clause of time with a participial predicate; it may also be occupied by an infinitive, a prepositional phrase, a noun phrase, an adverb with a temporal meaning, or a word meaning 'perhaps' – ex. (18) to (22).

(18) [X **E (a) re**] [Y **go lwala**] [Z **ene a sale mo gae**] / ì(à)rì χɔ̀lwálá ɛ̀nɛ́ àsálì mó χâì / When (s)he is sick, (s)he remains at home / **e (a) re** [ì(à)rì] is the present of **re** with an expl SM of cl 9, **a sale** [àsálì] is the e-consecutive of **sala** 'remain' with a SM of cl 1, and **go lwala** is the infinitive of **lwala** 'be sick'

(19) [X **E rile**] [Y **fa morago ga foo**] [Z **ene a sala mo gae**] / ìrílè fá mɔ̀ráχɔ̀ χáfóɔ́ ɛ̀nɛ́ àsálà mó χâì / After that, (s)he remained at home / **e rile** [ìrílè] is the perfect of **re** with an expl SM of cl 9, **a sala** [àsálà] is the a-consecutive of **sala** 'remain' with a SM of cl 1, and **fa morago ga foo** is a prepositional phrase whose literal meaning is 'in back of that'

(20) [X **E tlaa re**] [Y **kamoso**] [Z **ene a sale mo gae**] / ìtɬààrì kámɔ̀sɔ́ ɛ̀nɛ́ àsálì mó χâì / Tomorrow (s)he will remain at home / **e tlaa re** [ìtɬààrì] is the future of **re** with an expl SM of cl 9, **a sale** [àsálì] is the e-consecutive of **sala** 'remain' with a SM of cl 1, and **kamoso** is the Setswana equivalent of English 'tomorrow'

(21) [X **E rile**] [Y **gongwe**] [Z **ene a sala mo gae**] / ìrílè χɔ̀ŋwì ɛ̀nɛ́ àsálà mó χâì / Perhaps (s)he remained at home / **e rile** [ìrílè] is the perfect of **re** with an expl SM of cl 9, **a sala** [àsálà] is the a-consecutive of **sala** 'remain' with a SM of cl 1, and **gongwe** is a Setswana equivalent of English 'perhaps'

(22) [X **E tlaa re**] [Y **kgotsa**] [Z **ene a sale mo gae**] / ìtɬààrì qhɔ̀tsà ɛ̀nɛ́ àsálì mó χâì / Perhaps (s)he will remain at home / **e tlaa re** [ìtɬààrì] is the future of **re** with an expl SM of cl 9, **a sale** [àsálì] is the e-consecutive of **sala** 'remain' with a SM of cl 1, and **kgotsa** is another Setswana equivalent of English 'perhaps'

4. Evidence against considering that the verb of the clause occupying position Z in the sentence pattern illustrated by ex. (6) to (17) constitutes by itself the main predicate. Let us now turn to observations showing that the verb of the clause in position Z cannot be considered as constituting by itself the main predicate of this construction, and that the only possible analyses are that the main predicate in this construction is, either the inflected form of **re** in position X, or the combination of the inflected form of **re** in position X and of the verb of the clause in position Z.

4.1. In the sentence pattern illustrated by examples (6) to (17), the inflected form of **re** in position X shows the full range of tense variations characteristic of verbs fulfilling the predicate function in independent clauses. By contrast, the predicate of the clause in position Z is in a dependent form (a-consecutive, e-consecutive or subjunctive), and its tense variations are determined by those of **re**, as shown in the following chart:

inflected form of **re** in position X	verb of the clause occupying position Z
present	e-consecutive
future	e-consecutive
e-consecutive	e-consecutive
perfect	a-consecutive
potential	a-consecutive
a-consecutive	a-consecutive
subjunctive	subjunctive

In other words, the modal/temporal value assigned to the sentence as a whole is entirely determined by the tense form of **re**.

It may also be observed that the relationship between the tense of **re** and the tense of the verb fulfilling the predicate function in the clause occupying position Z is identical to that observed between the tense of the auxiliary and the tense of the 'main verb' in a number of compound verb forms.

4.2. Another important observation concerns the behavior of sentences of the type analyzed here when put in contexts requiring a special form of the predicate. In Setswana, the main predicate of a relativized clause must be in a so-called relative form (characterized by the ending **-ng**), and various other syntactic contexts require the use of a so-called participial form of the predicate – ex. (23). In the same contexts, in cases when the predicate is a compound verb form, the auxiliary shows the morphological particularities characteristic of relative or participial verb forms, whereas the main verb does not vary – ex. (24).

(23a) **Molelo o gotsitswe mo tlung** / mɔ̀lìlɔ̀ ɔ́χótsítswè mó tɬúŋ̀ / The fire has been lit in the house / **o gotsitswe** [ɔ́χótsítswé] is the perfect positive of the passive form of **gotsa** 'light' with a SM of cl 3

(23b) **Molelo o [o gotsitsweng mo tlung] o sa ntse o tuka** / mɔ̀lìlɔ̀ ó ɔ́χótsítswèŋ́ mó tɬúŋ̀ ɔ́sántsì ɔ́túkà / The fire that has been lit in the house is still burning / **o gotsitsweng** [ɔ́χótsítswèŋ́] is the relative form of the perfect positive of the passive form of **gotsa** 'light' with a SM of cl 3

(24a) **Molelo o ne o gotsitswe mo tlung** / mɔ̀lìlɔ̀ ɔ́nè ɔ́χótsítswé mó tɬúŋ̀ / The fire had been lit in the house / **o ne o gotsitswe** [ɔ́χótsítswé] is the pluperfect positive of the passive form of **gotsa** 'light' with a SM of cl 3

(24b) **Molelo o [o neng o gotsitswe mo tlung] o ne o sa ntse o tuka** / mɔ̀lìlɔ̀ ó ɔ́nèŋ́ ɔ́χótsítswé mó tɬúŋ̀ ɔ́nè ɔ́sántsì ɔ́túkà / The fire that had been lit in the house was still burning / **o neng o gotsitswe** [ɔ́nèŋ́ ɔ́χótsítswé] is relative form of the pluperfect positive of the passive form of **gotsa** 'light' with a SM of cl 3 – note that the morpheme characteristic of relative verb forms **-ng** is suffixed to the auxiliary

Given this regularity, it is important to observe that, when sentences of the type analyzed here are relativized or put into contexts requiring a participial form of the main predicate of the embedded sentence, the inflected form of **re** in position X occurs in a relative form or in a participial form, whereas the verb of the clause in position Z does not undergo any modification – ex. (25).

(25a) [$_X$ **E rile**] [$_Y$ **go robalwa**] [$_Z$ **molelo wa gotsiwa mo tlung**] / ìrílè χɔ̀rɔ̀bàlwà mɔ̀lìlɔ̀ wàχótsíwà mó tɬúŋ̀ / When people went to bed, the fire was lit in the house / **e rile** [ìrílè] is the perfect positive of **re** with an expl SM of cl 9, **go robalwa** [χɔ̀rɔ̀bàlwà] is the infinitive of the passive form of **robala** 'lie down and sleep', **wa gotsiwa** [wàχótsíwà] is the a-consecutive of the passive form of **gotsa** 'light' with a SM of cl 3

(25b) **Molelo o [[$_X$ e rileng] [$_Y$ go robalwa] [$_Z$ wa gotsiwa mo tlung]] o sa ntse o tuka** / mɔ̀lìlɔ̀ ó ɪ́rílèŋ̀ χɔ́rɔ̀bàlwà wàχɔ́tsíwà mó tɬúŋ̀ ɔ́sáńtsì ɔ́túkà / The fire that was lit in the house when people went to bed is still burning / **e rileng** [ɪ́rílèŋ̀] is the relative form of the perfect positive of **re** with an expl SM of cl 9, **go robalwa** [χɔ̀rɔ̀bàlwà] is the infinitive of the passive form of **robala** 'lie down and sleep', **wa gotsiwa** [wàχɔ́tsíwà] is the a-consecutive of the passive form of **gotsa** 'light' with a SM of cl 3

In other words, the inflected form of **re** in position X behaves in this respect as if it were, either the main predicate of the sentence, or the first element of a compound main predicate. Any other analysis would be in contradiction with this aspect of the behavior of **re**.

4.3. From the observations presented in **4.1.** and **4.2.** it follows that **re** in the construction we are examining maintains more verbal properties and shows a closer relationship to the verb of the clause in position Z than we could have expected by only taking into account its position at the very beginning of this construction and the usual translation of this construction into English. One must agree with Cole's statement that "Though these forms are translated idiomatically by 'when', 'whenever', 'if', etc., in English, it is clear that fundamentally they express the idea of happening". In other words, non-idiomatic translations such as the following ones would certainly be closer to the Setswana sentences quoted in ex. (6) to (17) than those given above, both in structure and in meaning:

(6) It happened that, the others going to the field, (s)he remained at home
(7) It did not happen that, the others going to the field, (s)he remained at home
(8) ... And it happened that, the others going to the field, (s)he remained at home
(9) It will happen that, the others going to the field, (s)he will remain at home
(10) It will not happen that, the others going to the field, (s)he will remain at home
(11) ... And it will happen that, the others going to the field, (s)he will remain at home
(12) It happens that, the others going to the field, (s)he remains at home
(13) It does not happen that, the others going to the field, (s)he remains at home
(14) It may happen that, the others going to the field, (s)he remains at home
(15) It may not happen that, the others going to the field, (s)he remains at home
(16) It must happen that, the others going to the field, (s)he remains at home
(17) It must not happen that, the others going to the field, (s)he remains at home

However, the fact that such translations are undoubtedly relatively close to the structure of the Setswana sentences does not imply that they reflect every detail of the structure of their Setswana equivalents, and the question of the precise status of **re** in this constructions remains to be discussed.

5. Evidence against considering that the inflected form of *re* occurring in initial position in the sentence pattern illustrated by ex. (6) to (17) constitutes by itself the main predicate. We must therefore go on to discuss the analysis according to which, in the sentence pattern illustrated by ex. (6) to (17), the inflected form of **re** in position X is the exact equivalent of English **happen** in 'It happened that ...', i.e. the main predicate, the clause in position Z being then analyzed as its complement.

In fact, it is not very difficult to convince oneself that this analysis too must be rejected: the morphosyntactic properties of the inflected form of **re** occupying

position X in the Setswana construction analyzed here are different from those of regular Setswana verbs introducing clausal complements, and in particular from those manifested by **re** itself in other constructions where its behavior is clearly the behavior of a verb fulfilling the predicate function.

At first sight, it seems possible to relate the construction illustrated by ex. (6) to (17) to the construction in which **re** has a subject representing a person concerned by an event, as in ex. (26).

(26) **Ngwana o rile eng?** / ŋwàná ɷ́rìlé ɩ̀ŋ́ / What happened to the child?[9] / **o rile** [ɷ́rìlé] is the perfect of **re** with a SM of cl 1, **eng** is the interrogative 'what?'

Unfortunately, the use of **re** analyzed here cannot be considered, at least from the synchronic point of view, as a mere subjectless (or 'impersonal') variant of the use of **re** illustrated by ex. (26), since according to the productive rules of Setswana grammar, subjectless (or 'impersonal') verb forms take an expletive SM of class 17 (and not of class 9). Ex. (27) illustrates the subjectless construction that regularly corresponds to ex. (26).

(27) **Go rile eng mo ngwaneng?** / χɷ́rìlé ɩ̀ŋ́ mó ŋwànéŋ̀ / What happened to the child? / **go rile** [χɷ́rìlé] is the perfect of **re** with an expl SM of cl 17, **eng** is the interrogative 'what?', **mo** is a preposition, and **ngwaneng** is the locative form of **ngwana** 'child'

The use of the SM of class 9 as an expletive SM in the construction analyzed here makes it impossible to analyze this construction, within the limits of present-day Setswana grammar, as involving a subjectless use of the verb **re** fulfilling the predicate function. Moreover, it has already been observed that in this construction (and only in this construction), **re** has a toneless SM of class 9 in tenses in which the SM of class 9 is normally high-toned, and this provides further evidence against the analysis of **re** as constituting by itself the main predicate of sentences such as ex. (6) to (17). The point is that in Proto-Bantu, the tenses in question had a low-toned (or toneless) SM in the 1st person, in the 2nd person, in class 1 and in class 9, and a high-toned SM in the other classes. In Setswana, the SMs of class 1 and 9 in the tenses in question became high-toned by analogy with the other classes. The fact that **re** maintains a toneless SM in the construction analyzed here is certainly an indication that the forms of the verb **re** occurring in this construction had already undergone a process of decategorialization at the time when this change occurred.

Another argument against considering the clause occupying position Z in the construction analyzed here as the complement of **re** is that, when **re** as a verb meaning 'say' has a clausal complement, the verb fulfilling the predicate function in the complement clause occurs in tenses characteristic of independent clauses, and the choice of its tense is independent from the choice of the tense of **re**. By contrast, the possible combinations of tenses observed in the construction illustrated by ex. (6) to (17) are identical with those observed in the consecutive construction, illustrated by ex. (28) to (32).

(28) **Ke ile Gauteng ka etela ditsala tsa me** / kɩ̀ɩ̀lé χàútéŋ̀ kàɛ̀tɛ̀là dìtsálà tsámɩ̀ / I went to Jo'burg and visited my friends / the first verb – **ke ile** 'I went'– is in the perfect, the second one – **ka etela** – is in the a-consecutive

(29) **Nka ya Gauteng ka etela ditsala tsa me** / ŋ̀kájá χàútéŋ̀ kàɛ̀tɛ̀là dìtsálà tsámì / I can go to Jo'burg and visit my friends / the first verb – **nka ya** 'I can go' – is in the potential, the second one – **ka etela** – is in the a-consecutive, i.e. in the same form as in ex. (28)

(30) **Ke ya Gauteng ke etele ditsala tsa me** / kìjà χàútéŋ̀ kìètélí dìtsálà tsámì / I (usually) go to Jo'burg and visit my friends / the first verb – **ke ya** 'I go' – is in the present, the second one – **ke etele** – is in the e-consecutive

(31) **Ke tlaa ya Gauteng ke etele ditsala tsa me** / kìtɬààjà χàútéŋ̀ kìètélí dìtsálà tsámì / I shall go to Jo'burg and visit my friends / the first verb – **ke tlaa ya** – is in the future, the second one – **ke etele** – is in the e-consecutive, i.e. in the same form as in ex. (30)

(32) **O ye Gauteng o etele ditsala tsa gago** / ɷ́jɛ́ χàútéŋ̀ ɷ́ɛ́tɛ̀lɛ́ dìtsálà tsáχáχɷ̀ / You should go to Jo'burg and visit your friends / both verbs – **o ye** and **o etele** – are in the subjunctive

6. Evidence against analyzing the sentence pattern illustrated by ex. (6) to (17) as a consecutive construction. In fact, if the inflected form of **re** in position X and the verb of the clause in position Z were to be analyzed as distinct predicates, the only analysis compatible with the morphological data would be to consider that they constitute a consecutive construction. However, syntactically, the construction illustrated by ex. (6) to (17) has in some respects properties that sharply contrast with those of the consecutive construction, so that its analysis as a consecutive construction must be rejected too.

A first observation providing evidence against this analysis is that the first member of a true consecutive construction is a complete clause capable of standing alone as a declarative or hortative sentence, as can be seen by comparing ex. (28) to (32) above with ex. (33) to (37).

(33) **Ke ile Gauteng** / kìilé χàútéŋ̀ / I went to Jo'burg

(34) **Nka ya Gauteng** / ŋ̀kájá χàútéŋ̀ / I can go to Jo'burg

(35) **Ke ya Gauteng** / kìjà χàútéŋ̀ / I go to Jo'burg

(36) **Ke tlaa ya Gauteng** / kìtɬààjà χàútéŋ̀ / I shall go to Jo'burg

(37) **O ye Gauteng** / ɷ́jɛ́ χàútéŋ̀ / You should go to Jo'burg

By contrast, X alone or X + Y (without Z) are not possible Setswana sentences – ex. (38) to (42).

(38) *** E rile** *** E rile ba bangwe ba ya kwa masimong**

(39) *** E (a) re** *** E (a) re ba bangwe ba ya kwa masimong**

(40) *** E tlaa re** *** E tlaa re ba bangwe ba ya kwa masimong**

(41) *** E ka re** *** E ka re ba bangwe ba ya kwa masimong**

(42) *** Ya re** *** Ya re ba bangwe ba ya kwa masimong**

Therefore, the clause occupying position Z in this sentence pattern cannot be analyzed (by itself or in combination with Y) as the second member of a consecutive construction, in spite of the fact that the tense variations of its verb are identical to those of verbs fulfilling the predicate function in the non-initial clauses of a consecutive construction.

The way negation functions in the sentence pattern illustrated by examples (6) to (17) provides additional evidence that this analysis must be rejected. Ex. (7), (10), (13), (15) & (17) show that the inflected form of **re** in position X may be in the negative form, and that the clause in position Z is within the scope of a negation morphologically affecting the inflected form of **re** in position X. By contrast, in a true consecutive construction, a non-initial clause cannot be within the scope of a negation morphologically affecting the verb of the first clause. For example, if we want to express something like 'I did not go to Jo'burg and I did not go to Pretoria either', both verbs must be in the negative form – ex. (43). A consecutive construction with the first verb in the negative form and the second one in the positive form is usual only with a connective between the two verbs – ex. (44), but, even in the absence of a connective, the only possible interpretation is that the second verb is outside the scope of the negation of the first one.

(43) **Ga ke a ya Gauteng ka se ka ka ya Tshwane** / χàkíàjá χàútéŋ̀ kàsìká kájà tshwánì / I did not go to Jo'burg and I did not go to Pretoria either / the first verb – **ga ke a ya** – is in the perfect negative, the second one – **ka se ka ka ya** – is in the a-consecutive negative

(44) **Ga ke a ya Gauteng mme ka ya Tshwane** / χàkíàjá χàútéŋ̀ m̀mí kájà tshwánì / I did not go to Jo'burg but I went to Pretoria / the first verb – **ga ke a ya** – is in the perfect negative, the second one – **ka ya** – is in the a-consecutive positive

7. Conclusion: the inflected form of *re* occurring in initial position in the sentence pattern analyzed here is an auxiliary, but the position it occupies is not the normal position of Setswana auxiliaries. As regards negation, the relationship between the inflected form of **re** in position X and the verb of the clause in position Z has the same properties as the relationship between the auxiliary and the main verb in a number of compound verb forms which, from the diachronic point of view, are probably frozen consecutive constructions.

More generally, the conclusion we can draw from the preceding discussion is that the analysis according to which the inflected form of **re** in position X and the verb of the clause in position Z constitute a single predicate is the only one that does not raise important difficulties. In other words, what seems at first sight to be the verb form assuming the predicate function in the clause in position Z is best analyzed as constituting only the second part of a compound verb form whose first part (i.e. the auxiliary) is the inflected form of **re** in position X.

According to this analysis, the simple tenses of the Setswana verb have compound variants characterized by the use of **re** as the auxiliary. Semantically, these compound forms emphasize the idea of happening. For example:

simple tenses		compound variants with the auxiliary **re**
o (a) sala	(s)he remains	**e (a) re ... a sale**
o tlaa sala	(s)he will remain	**e tlaa re ... a sale**
o setse	(s)he remained	**e rile ... a sala**
o ka sala	(s)he may remain	**e ka re ... a sale**

This analysis is not entirely satisfying, since the behavior of **re** is different from that of typical Setswana auxiliaries in at least two respects: the SM of typical Setswana auxiliaries necessarily varies in agreement with the subject (whereas **re** may have an invariable SM of class 9, the agreement with the subject being then manifest in the main verb only), and typical Setswana auxiliaries usually occur immediately before the main verb and never precede the subject NP (whereas the auxiliary **re** precedes the subject NP and is necessarily separated from it (and from the main verb) by the constituent in position Y).

However, there are in Setswana other cases of constructions which must be analyzed as compound verb forms in spite of the fact that they raise similar problems: the auxiliary **sa le** 'it happened long ago that ...' may optionally have a SM identical with the SM of the main verb, an invariable SM of class 9, or even no SM at all, and the auxiliary **tswa**, in a construction meaning that a situation has continued from a given point in the past until now, is necessarily separated from the main verb by a constituent referring to that point of time, as in ex. (45).

(45) **O tswa bongwaneng a rata dijo tse** / ɷ́tswá bɷ́ŋwànéŋ̀ árátá dìʤɔ́ tsè / (S)he has liked this food since (s)he was a child / **bongwaneng**, the locative form of **bongwana** 'childhood', is inserted between the two parts of the compound verb form **o tswa ... a rata** – literally '(s)he comes from ... liking'

8. The possibility of a subsequent reanalysis. In the preceding sections, I tried to show that inflected forms of the verb **re** that at first sight seem to fulfill the function of complementizer introducing a subordinate clause of time S_1 in a construction '**re** S_1 S_2' are best analyzed, in the present state of Setswana, as the first element of a compound verb form fulfilling the function of predicate of clause S_2. A crucial point in this analysis is that, in this construction, the constituent immediately following the inflected form of the verb **re** is not necessarily a clause: it may be a noun phrase, a prepositional phrase or an adverb. However, in Setswana texts, the occurrences of the auxiliary **re** immediately followed by a subordinate clause of time are far more numerous than the occurrences of the auxiliary **re** followed by another type of constituent. In other words, **re** as an auxiliary emphasizing the idea of happening occurs mainly in a context making it possible to reanalyze it as a complementizer introducing a subordinate clause, and it is interesting to note that some observations suggest that such a process is beginning to develop in Setswana.

In section **4**, I insisted on the fact that the verb fulfilling the predicate function in the clause occupying position Z occurs in a dependent form, and that its form is entirely determined by the inflected form of **re** in position X. However, this is not always true. Setswana speakers sometimes use sentences almost identical to those following the pattern analyzed in the preceding sections, but in which the verb fulfilling the predicate function in the clause occupying position Z is in an

independent form. For example, instead of (12) (repeated here as (46)), with **sala** in the e-consecutive, it is possible to find (47), with **sala** in the indicative present.

(46) **E (a) re ba bangwe ba ya kwa masimong ene a sale mo gae /** ì(à)rì bá bàŋwí bájá kwá masímɔ̀ŋ ɛ̀nɛ́ àsálì mó χáì / When the others go to the field, (s)he remains at home / **a sale** [àsálì] is the e-consecutive of **sala** 'remain'

(47) **E (a) re ba bangwe ba ya kwa masimong ene o sala mo gae /** ì(à)rì bá bàŋwí bájá kwá masímɔ̀ŋ ɛ̀nɛ́ ɔ́sálà mó χáì / When the others go to the field, (s)he remains at home / **o sala** [ɔ́sálà] is the indicative present of **sala** 'remain'

This construction, in which the relationship between the two elements of a compound verb form is broken, does not seem very frequent, and when questioned about it, Setswana speakers often seem rather reluctant to accept it. But it does exist, and its existence is already mentioned in Cole 1955. This means that the grammaticalization of **re** into an auxiliary emphasizing the sense of happening could well be followed by a reanalysis of this auxiliary into a complementizer introducing temporal clauses.

---

[1] Setswana has two other verbs corresponding in certain contexts to English 'say' or 'tell': **bolela** 'say', 'tell', 'report' (as in **O mpoleletse dilo tse di kgatlhang** 'He told me interesting things') and **bua** 'say', 'tell', 'speak' (as in **Bua nnete** 'Tell the truth'). By contrast with **re**, they are morphologically regular, and their syntactic behavior is that of ordinary transitive verbs.

[2] **Itse** 'know' and **lere** 'bring' show the same irregularity.

[3] The a-consecutive is a tense typically used in sequences of clauses referring to sequences of past events or to sequences of conditional events – see ex. (28) & (29) below.

[4] See Creissels & al., forthcoming.

[5] Note that, in the absence of a wider context, the written form of the sentence is sometimes ambiguous, and that a precise transcription of vowels and of tone is necessary in order to identify the form of **re** in position X and/or the form of the verb fulfilling the predicate function in Z.

[6] The e-consecutive is a tense typically used in sequences of clauses referring to sequences of future events or to sequences of habitual events – see ex. (30) & (31) below.

[7] In this construction, the 'short' form and the 'long' form of the present of **re** are interchangeable. This is an irregularity, since in principle, the choice between these two forms expresses distinctions in the discursive structure of the utterance – see Creissels 1996. It is interesting to observe that several auxiliaries show the same irregularity.

[8] See Cole 1955, p. 302.

[9] This sentence may equally be interpreted as 'What did the child say?'; the choice between these two interpretations entirely depends on the context.

REFERENCES

COLE D. T., 1955, *An introduction to Tswana grammar*, Longman, Cape Town.

CREISSELS D., 1996, Disjunctive and conjunctive verb forms in Setswana, *South African Journal of African Languages,* vol. 16 n° 4, p. 109-115.

CREISSELS D., CHEBANNE A. M. & NKHWA H. W., forthcoming, *Tonal Morphology of the Setswana Verb*, LINCOM Studies in African Linguistics.

Unifying predicate cleft constructions

Hilda Koopman
UCLA

1. The problem. Like many West African languages, Vata, a Niger Congo language of the Kru family, has a particular verbal focus construction, sometimes referred to as the *predicate cleft* construction. This construction involves contrastive focus on V or a predicate: a V is understood as contrasting with some verb implicit in the discourse: [1]

(1) **pā ǹ ká mɛ́ pá ā** (Vata) (Koopman, 1984)
throw you FUT it throw Q
'Are you going to THROW it' (*throw* as opposed to *roll*)

(2) **pā ń ká̄ mɛ́ pá**
throw I will it throw
'I will throw it'

The following properties characterize the contrastive verb focus construction in Vata.
Morphology: the clause contains two copies of the verb. The verb in initial position carries special morphology associated with the construction (realized in Vata as a 'construction' tone). The verb in the clause looks and acts like any regular V.
Order: The contrastively focused verb occurs in clause initial position. Omission of the focused verb yields a regular sentence without focus.
Dependency: The dependency between the focused verb and the copy obeys the same locality as manner and reason adjuncts (Koopman 1984, Koopman and Sportiche 1986).

A very similar construction involving contrastive focus on V, is found in Nweh, a Grassfield Bantu language spoken in Cameroon (Nkemnji 1995)[2]:

(3) à kɛ̀ʔ **ncù** ká **cǔ** (Nweh)
s/he P boil crab boil
'She BOILED the crab'(as opposed to frying it.

(4) à kɛ̀ʔ **ncù** ká **cǔ** lɛ̌ (Nweh)
s/he P boil crab boil (Q)
'Did she BOIL the crab (as opposed to frying it)

As in Vata, the clause contains two copies of the same V. The *leftmost* verb in Nweh has the form and distribution characteristic of Vs in clauses without verbal

focusing. The *rightmost* verb carries particular verbal morphology (a tonal prefix and suffix and a segmental suffix).[3]

Apart from linear order, there is a further difference between Vata and Nweh which concerns cooccurrence restrictions of wh-phrases and focused verbs. In Vata, a focused V cannot cooccur with any wh-phrase, regardless of whether the wh-phrase is a subject, an object, or an adjunct (Koopman 1984):

(5) a. *pā àlɔ̋ ɔ̀ ká mɛ̋ **pá** lá (Vata)
throw who he-R FUT it throw wh

b. * àlɔ̋ pā ɔ̀ ká mɛ̋ **pá** la (Vata)
who throw he-R FUT it throw wh

In Nweh subject wh-phrases can cooccur with predicate cleft (6), but non subject wh-phrases cannot (7):

(6) àwɔ́ kɛ̀ʔ ǹjuɔ̀ bɛ́ jùɔ́ lɛ́
who P1 n-buy fufu buy Q
'Who BOUGHT the fufu?' (as opposed to who sold the fufu)

(7) *àtɛ̀m kɛ̀ʔ ǹjuɔ̀ kɔ́ jùɔ́ lɛ́
Atem P1 n-buy what buy Q
'What did Atem BUY' (as opposed to sell)

The data above raise the questions that I will try to answer in this paper:

(8) How should one account for the difference in linear order?
(9) How should one account for the different cooccurrence restrictions of focused verb and wh-phrases.

I will present an analysis of the predicate cleft construction, and argue for an optimally simple analysis of the crosslinguistic variation which derives both the differences in word order and the differences in cooccurrence restrictions from a common underlying structure.

2. Theoretical assumptions. The theoretical assumptions below are 'minimalist' in spirit, but differ in the general shape of the theory. The overall picture is closest to the work of Sportiche 1993, Kayne 1994, Rizzi 1995, Cinque 1996.

- Syntactic structures are Binary Branching structures, obeying X-bar theory. Whether the properties of X-bar theory can be derived (Kayne, 1994, Chomsky, 1995) is of no concern to the present paper.
- Each feature projects. This is what Sportiche (1996) calls the 'atomization' of syntactic structures. The one-projection-per-feature theory is a logical continuation of work in the eighties on the architecture of clauses, DPs, APs and PPs.
- All languages are underlyingly identical (Universal Base Hypothesis).
    - There is no head initial head final parameter: all languages are Spec head complement underlyingly (Kayne, 1994). For the purposes of this paper it is sufficient that there are no underlyingly mixed languages.
    - There is no crosslinguistic difference in hierarchical structure (Sportiche 1993 1995, Cinque 1996, Koopman, 1996). Language variation cannot be attributed to different hierarchical locations of projections associated with the same semantic interpretation. This rules out analyzing the difference between Vata and Nweh in terms of a different location of Focus (say high focus versus low focus).
    - There is no Procrastination: everything must move overtly. Crosslinguistic differences do not derive by overt or covert movement, but by movement of different sized constituants (see Koopman, 1996, for more diccussion as well as the present paper). Movement (copy and deletion) is of usual kind:
        - head movement (left adjunction only, no base generated morphology)
        - XP movement (leftward only to designated Spec positions, no adjunctions).

It is important to keep in mind that *there is much more XP movement (pied-piping)* than we are used to in standard analyses of say English, with big parts of sentences, and sentences themselves moving around (Sportiche 1993, Koopman 1995, 1996, Kayne 1994, Nkemnji 1995). Much of this pied-piping is transparent in African languages both within DPs (final determiners, quantifiers, demonstratives etc) and within the clause (final negation, final question particles) (see in particular Nkemnji 1995).

- Movement obeys locality. Head movement obeys the Head Movement Constraint, and XP movement can only reach a local Spec. Head movement extends the domain of movement, and brings the next local Spec in the local domain (the Head Constraint of Van Riemsdijk 1978, The Government Corollary of Baker 1988, Equidistance of Chomsky 1991). Locality is 'wired-in', and not subject to Economy.

Movement takes place for licensing purposes, either for the familiar morphological reasons, or for semantic reasons (scope). In addition, I assume that movement can

be forced because of a principle that I have called the *PPA* (*Principle of Projection Activation* (Koopman 1996))

(10) Principle of Projection Activation (Koopman 1996) (PPA)
A Projection is interpretable iff it is associated with lexical material at some stage in the derivation.

The PPA prevents representations with truly empty projections (where neither Spec, nor head contains a lexical item or a trace) and forces movement. A translation of the PPA into the standard Minimalist terminology comes close to (11):

(11) functional heads are strong.

- Overt material must be linearized. I assume that the distribution of overt lexical items over these huge universal structures is determined by some version of the *LCA* (Linear Correspondence Axiom (Kayne 1994)). In Koopman, 1996, I modify the LCA and show that this modification yields the doubly filled C filter[4] .

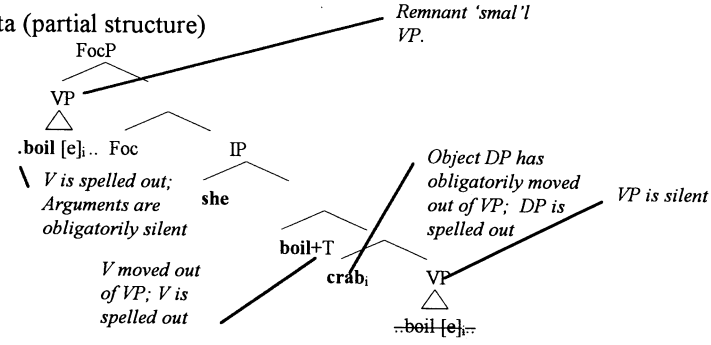
(12) Modified LCA has as consequence that no Spec and head position can simultaneously contain overt lexical material.

3. The analysis of predicate cleft. The verbal focus construction in Vata and Nweh receives the same contrastive focus interpretation (which Larson and Lefebvre (1991) analyze as quantification of the event). This construction never yields an emphatic reading, (*he DID want to..* as opposed *to he did NOT want to...*), and cannot be used with individual level predicates. Since the focused verb occurs in a particular position in the clause, I will assume that it is 'associated'(in a sense to be made precise below) with the Focus Projection (FocP). Since the same semantic interpretation arises, I will assume that it is associated with the same FocP in both Vata and Nweh.

3.1. Predicate cleft in Vata. In Vata, the focused V appears at the left edge of the sentence, pointing to a head initial FocP (FocP>IP). The focused verb is 'associated' with the FocP, which implies that the focused verb is either in the Spec position of the FocP, or in the head position. In Koopman 1984, I argued that the focused verb moved to COMP via head movement (at that point basically the only available analysis). I called this type of head movement the wh-type of head movement (A' head movement), because it behaved like phrasal A' movement, and not like V to I movement which I called the A type head movement. With the subsequent development of the 'middle' field, an XP analysis of the predicate cleft construction has become feasible and desirable: instead of

head movement, predicate cleft involves XP movement of a 'small' VP containing nothing but V to Spec, FocP. This analysis immediately accounts for the A' properties of the construction: the predicate cleft construction patterns with XP movement, because it *is* XP movement. An XP analysis of predicate cleft makes the distinction between two types of head movement unnecessary, a welcome result. Finally, the XP analysis finds emperical support: some adverbs and aspectual markers may optionally accompany the focused V (Koopman 1984). There is no evidence that these should be analyzed as forming a complex head with the focused V. Since arguments and small clause predicates may never accompany the focused V, it must be the case that all arguments and complex predicates *must* obligatorily vacate the VP[5]. As I argue elsewhere on independent grounds, arguments and predicates must always be licensed in specific landing site positions outside of the minimal VP. The following annotated tree illustrates the derivation for Vata (English words are used for convenience).

(13) Vata (partial structure)



The resulting sentence contains two overt copies of the same V, each carrying different morphology. This is an old and well known problem of this construction that requires a new explanation. Under a head movement analysis, as in Koopman 1984, it was the spell-out of the V in the clause that required an explanation. Under a remnant movement analysis, it must be explained why the focused verb cannot be silent. Let us briefly consider the 'spell-out' problem taking into account the hierarchical relations. The V within the VP does not c-command the V in I, and therefore does not form a V chain with it. The V within the clause is thus spelled out for the same reason any V in the head of a chain position is. What is unexplained is why the focused V cannot be silent. The V moves outside of the VP to get tense morphology, just as arguments move out of the VP, prior to movement of the VP to Spec FocP. The V must be spelled out within the VP, (and within IP) but the copies of the arguments that are contained in the preposed VP cannot be spelled out and can only be spelled out in the IP. I will assume that spell out of V is forced by recoverability (after all, if the focused verb were silent, nothing would signal verbal focusing), and that spelling out of arguments is

prohibited in the absence of a local licenser. The morphology associated with the verbal focus plays a crucial role in that it makes the spell out of the focused verb possible.

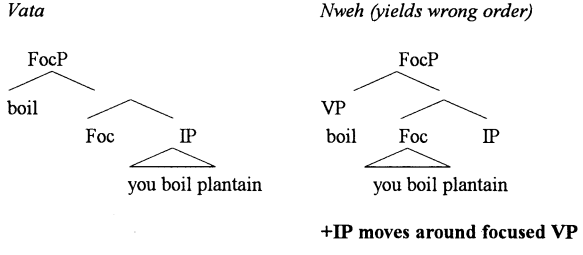
3.2. Predicate cleft in Nweh. Vata and Nweh have the same verbal focus construction, with the same meaning, and therefore involve the same underlying hierarchical structure. The languages differ however as to the position in which the clefted predicate (=small VP) appears. In Vata it surfaces at the left edge of the sentence, pointing to a hierarchical order FocP>IP, in Nweh it surfaces somewhere toward the right edge. This is depicted in (14), which also includes information about the morphological structure of the clefted predicate:

(14) *Vata*: [V+tone] DP T ...V_1... (Q)
V-foc morphology V_1

Nweh DP T Vf DP... [focus tone+ V +segment] (Q)
V_1 *[focus morphology - V- focus morphology]*

The FocP in Nweh cannot be *underlyingly* head final, because of the assumption that there are no underlyingly head final languages (see section 0). The surface order in Nweh must therefore be derived by some leftward movement. At the surface it looks as if Nweh is using a low FocP, and Vata a high FocP. However, since by assumption there is no crosslinguistic difference in hierarchical order (this is really a 'minimalist' assumption) and since the construction in Nweh and Vata yield the same interpretations, FocP must be higher than IP in Nweh as well. Thus:

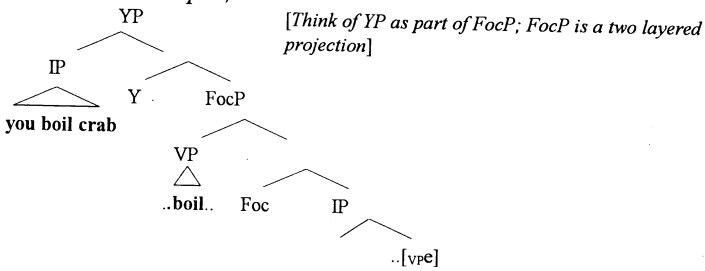
(15)



IP movement around the focused VP in Nweh raises the question of the landing site for IP. IP cannot land in Spec, FocP, because it hosts the clefted VP. Since it precedes the clefted V, it must be in the Spec of some higher position, YP. YP itself must be lower than Q, because the entire complement of Q precedes Q and Q scopes over FocP. Hence Q>Y>Foc. Although I will continue to label this projection YP, it is probably part of the focus projection, which should thus be viewed as a two layered projection. The YP possibly plays a role in pseudocleft constructions: *what John boiled is a crab*. The predicate cleft construction in

Vata would be comparable to a cleft construction; the predicate cleft construction in Nweh to a pseudocleft construction.

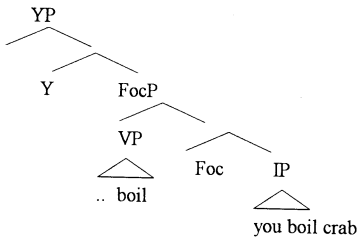
(16) *Nweh: IP moves to Spec, YP*



The movement of IP to Spec, FocP obeys locality. If Foc moves to Y both Spec, FocP and Spec, YP are equidistant to IP. Empirical evidence for Foc to Y head movement consists of a low tone preceding the focus constituent. Thus, Spec, YP and Spec FocP are 'active' in Nweh.

3.3. Crosslinguistic variation involving YP. If there is a YP above FocP in Nweh, this projection must be present in Vata as well (there is no crosslinguistic variation in structure, see section 0). Since the PPA requires that all projections be activated by lexical material (i.e. all functional projections are strong) the question arises how the YP is licensed in Vata. There is no indication of any head preceding the focused verb in Vata. It must therefore be the case that the Spec of YP is filled. I assume that the entire FocP moves to Spec, YP in Vata, thus giving the appearance of the head initial character of the FocP.

(17)

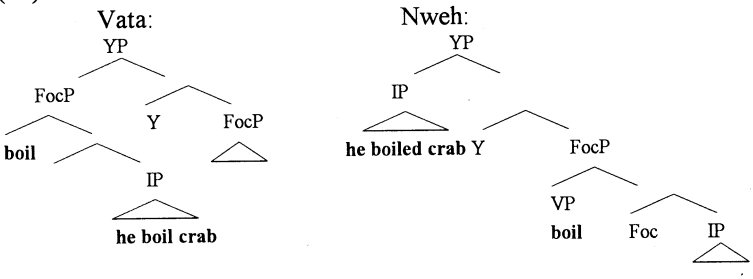


In both Vata and Nweh, Spec, YP contains a constituent, but the size of this constituent varies:

(18) in Vata: FocP is in Spec, YP
in Nweh: IP is in Spec, YP

The difference between Vata and Nweh is not a structural difference, nor a difference involving head initial/head final character of a particular head, nor a difference involving covert versus overt movement. *The difference lies in the size of the constituent that occupies YP*, with the entire FocP in YP Vata, and the IP (the complement of Foc) in Nweh. This is depicted in the following structure:

(19) Derived structures in Vata and Nweh: snapshot of YP.



It is easy to see that this analysis yields the different linear orders of Vata and Nweh from a common structure. I show in the next section that it does more: the incompatibility of wh-phrases and predicate cleft in Vata, and the compatibility of subject wh and predicate cleft in Nweh fall out from the derived structures in (19).

4. Deriving cooccurrence restrictions from necessary structural properties. Predicate cleft and wh-phrases are always incompatible in Vata. In Nweh, subject wh-phrases can cooccur with predicate cleft, but no other wh-phrases can. The incompatibility of focused verbs and wh-phrases in Vata was accounted for quite simply in earlier versions of the theory with a single landing site position for A' moved elements (COMP). Complementary distribution followed from competition for the same landing site. This solution obviously cannot work for Nweh since the cooccurrence of wh-phrases and predicate cleft is configurationally determined. Intuitively speaking, subject wh-phrases are able to reach the wh-position, but object wh-phrases are not, and this is precisely what the structures give us, as I will show below.

In the one projection per feature theory, these cooccurrence restrictions must be derived in a different way. For a similar problem arising in Italian, Rizzi 1995 proposes that the incompatibility of focus and wh-phrases follows from the fact that *wh-phrases are inherently focused*. This type of explanation predicts that wh-phrases and focus can *never* cooccur, and runs into trouble because Nweh subject wh-phrases and focused verbs can cooccur. It is unlikely that subject wh-phrases, and object wh-phrases receive a different focus interpretation, and we are dealing with contrastive focusing on V throughout. The Nweh data suggest a structural explanation which should have the effect that subject wh-phrases are able to reach the wh-position, but object wh-phrases not. I will now argue that the

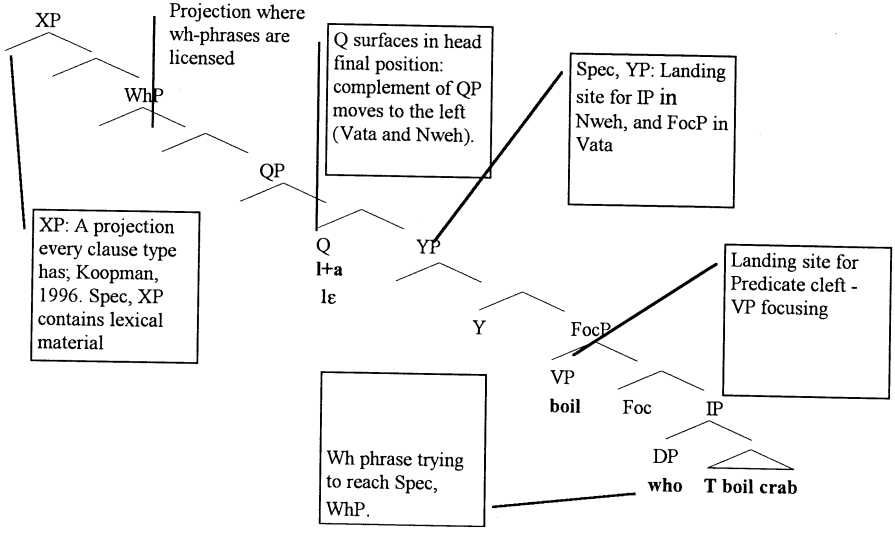
structures in (19) exactly yield this effect. What must be explained is the following:

(20) a. Wh-phrases cannot cooccur with predicate cleft in Vata
b. Subject wh-phrases can cooccur with predicate cleft in Nweh
c. Object wh-phrases cannot cooccur with predicate cleft in Nweh

In Koopman (1996), I have argued that wh-questions consist of a Wh projection, where wh-phrases are licensed and a Q projection, with Wh>Q. This yields the structure Wh>Q>YP>FocP. Wh-question formation involves the appearance of a sentence final matrix question particle in both languages (*la* in Vata, *lɛ* in Nweh) indicating leftward movement of the complement of Q.

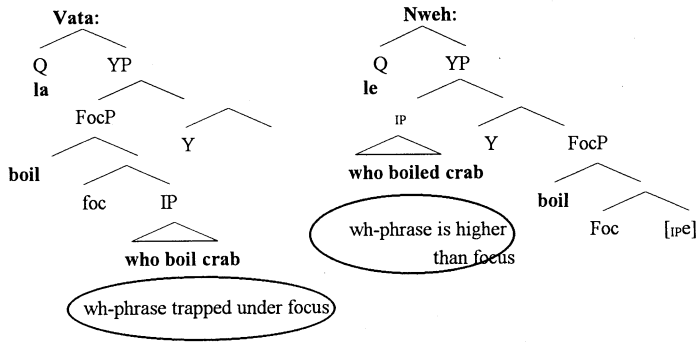
(21) Snapshot of a cooccurring predicate cleft and subject wh-phrase in Vata and Nweh:

snaphot taken at the point in the derivation where arguments and predicates have scrambled out of VP, VP has moved to Spec, FocP. The wh-phrase is in subject position of IP, and needs to reach Spec, Wh.



The structures in (21) are going to diverge, when movement to YP takes place (FocP moves to YP in Vata), IP moves around FocP to Spec, YP in Nweh, yielding the structures below:

(22) Snapshot of derived YPs in Vata and Nweh (lexical items boldfaced):



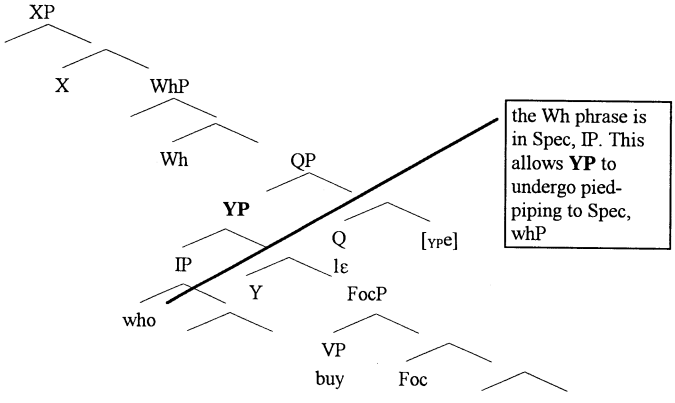
The cooccurrence restrictions fall out from these structures, as I will show in more detail below. In a nutshell:

- Wh-phrases in Vata can never cooccur with predicate cleft, because the wh-phrase will be unable to reach Spec, WhP. (section 4.1.)
- Subject wh phrases in Nweh can cooccur with FocP, because the movement of IP around FocP brings the wh-phrase in the local realm of the Spec, wh. (section 4.2.)
- Non subject wh-phrases in Nweh are also moved around focus, yet cannot coocur with focused Vs. The question why they cannot cooccur with predicate cleft cannot be answered in the same way as in Vata, since the wh-phrase in Nweh is no longer trapped under focus (section 4.3.)

4.1. The non-occurrence of predicate cleft and wh-phrases in Vata. As shown in (22) the wh-phrase in the predicate cleft constructions in Vata will always be trapped under focus. In order for a licit wh-interpretation to arise, the wh-phrase must move to Spec, WhP. But in order to do so, the wh-phrase must cross an intervening A'position, yielding a locality violation. The wh-phrase cannot trigger pied-piping of the entire FocP complement, because it is not in the right structural configuration to trigger pied-piping. It follows that predicate cleft and wh-phrases are incompatible in Vata: the wh-phrase always remains trapped under the FocP and can never reach the WhP in this configuration.

4.2. The cooccurrence of subject wh-phrases and predicate cleft in Nweh. Let us look at the next stage in the derivation in (22), when the complement of Q has raised to Spec, QP.

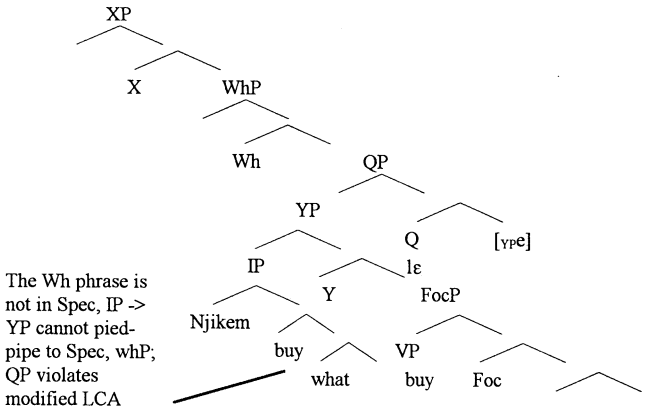
(23) Nweh: subject wh-phrases cooccurring with predicate cleft



The wh-phrase in the IP occurs in the Spec of the Spec position. This is a well-known pied-piping configuration (cf. *whose brother's picture did you take*), allowing it to pied-pipe the YP to Spec, whP. (For arguments that English subject extraction involves pied-piping of the entire clause, see Koopman 1996). This structure can be linearized without any problems, because no projection contains lexical material in both Spec and head position simultaneously. Note that the option of subextracting the wh-phrase is blocked by the modified LCA (which derives the doubly filled C filter, see section 2). If the subject were to extract, there would not be enough space for linearization: the QP projection would contain lexical material in both the head and the Spec position. movement, and thus violate the doubly filled C filter. Subject wh-phrases and predicate cleft can thus cooccur in Nweh because of the movement of IP around the focused constituent, and the particualr positon the wh-phrase occupies within the IP which allows for pied-piping of the bigger constituent.

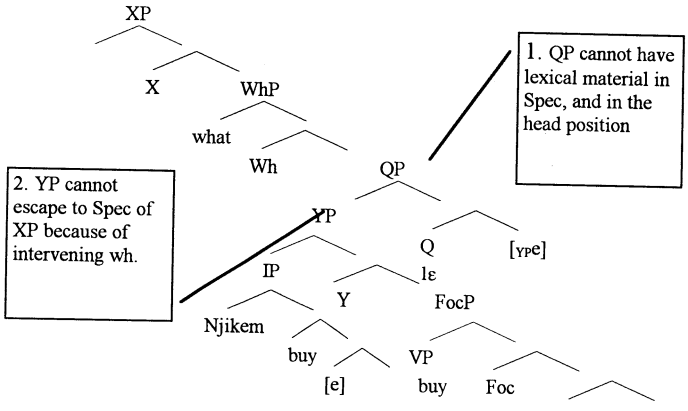
4.3. The non cooccurrence of wh-phrases and predicate cleft in Nweh. Let us consider the structure of a cooccurring predicate cleft and non-subject wh-phrase in Nweh at the point in the derivation where YP (containing the wh-phrase) has inverted with Q (moved to Spec, QP):

(24) Nweh: object wh-phrases cannot cooccur with predicate cleft



No grammatical sentence results from this structure. Two questions arise: why is pied-piping of the entire YP not possible (as it is for subjects), and why cannot the wh-phrase move to Spec, WhP on its own: one certainly must allow for movement of wh-objects! Non-subject wh-phrases are not in a pied-piping configuration (they are not in the Spec position of the entire constituent). The YP therefore cannot undergo pied-piping, and the surface string that would be derived by pied-piping is ungrammatical (**njikem boil what boil* (cf. example (7)). If the wh-phrase is too deeply embedded to trigger pied-piping, why cannot it extract by itself to Spec, WhP? This certainly seems to be a normal configuration for object extraction: it crucially differs from the Vata configuration, in that there is no intervening A' position. Suppose then that the wh-phrase is able reach Spec, whP in this configuration, and let us see if the resulting structure can be excluded on other grounds:

(25) Movement of wh-phrase to Spec, WhP:



The structure will be ruled out by the doubly filled C filter, if nothing else happens: both Spec, QP and Q contain overt lexical material. This accounts for the illformedness of (26):

(26) *akO njikEm kE? njuO juO lE *(without further inversion)*
what Njikem P1 N-buy buy Q

It is interesting that this structure improves substantially if it is embedded in an overt cleft construction (which seems to create the additional space for the realization of lexical material)

(27) ? akO m azea njikEm kE? njuO juO lE
what Foc Rel Atem P1 n-buy buy Q
What is it that Atem BOUGHT

The structure in (25) is blocked as well if further inversion takes place, showing that such inversion (IP movement) must be impossible as well in this structure (presumable because of the intervening wh-phrase).

(28) *njikEm kE? njuO juO kO lE *(with further inversion)*
Njikem P1 bought buy what Q

5. Conclusion. In this paper, I have provided strong support for a unified analysis of the predicate cleft construction in Vata and Nweh. A common hierarchical structure can be assumed to underly Vata and Nweh. The difference in surface order reduces to a difference in the size of the constituent that occupies a

particular Spec position: in Nweh IP moves around FocP to YP, whereas in Vata, the entire FocP occurs in YP. This analysis not only gets the different linear orders, but also, quite suprizingly, yields a simple explanation of the different cooccurrence restrictions of wh-phrases and focus Vs. This explanation uses non controversial assumptions about locality, a conservative assumption about pied-piping configurations (a wh-phrase can pied-pipe a constituent iff it is 'associated' with the Spec of that constituent), and (restricted) appeal to the generalized doubly filled C filter. My analysis does not appeal to a head initial head final parameter (such an analysis in fact would not allow the same explanation of the cooccurrence restrictions), nor to a different hierarchical organization (the analysis shows that it is not *necessary* to assume a different hierarchical struture, it does not show that a different hierarchical structure cannot be assumed). All movements are overt and no appeal is therefore necessary to covert movement, nor to the strong weak distinction of functional categories.

FOOTNOTES

---

[1] This particular way of expressing contrastive focus on V is not wide spread typologically. It is found in many West African languages of the Kwa family (for instance Yoruba, Fongbe, Ewe, Abe) and the Kru families (Vata, Gbadi, ...), in Caribbean Creoles (Saramaccan, Sranan, Haitian, Jamaican...). The Vata data are based on my fieldwork, discussed and analyzed in Koopman (1984) and Koopman and Sportiche (1986).

[2] The Nweh data are based on Nkemnji (1995) and data gathered during the UCLA fieldmethods class on Nweh, UCLA spring and winter quarter, 1996. Thanks go to the participants of the class, Michael Nkemnji, Tonia Androutsopoulou, Edward Garrett, Matt Gordon, Catherine Crosswhite, Javier Guttierez, Peter Hallman, Chai-Shune Hsu, and Matt Pearson, as well as to Manuel Espanol-Echevarría, Anna Szabolcsi, Ed Stabler, Andrew Simpson, and Dominique Sportiche. A computerized data base on Nweh is available on request. Working papers are in preparation.

[3] Because of space limitations, I will not be able to go into the dependency between the two verbs in Nweh. It is difficult to show that the relation can be non-local, since many clause types can contain focus. Straightforward island violations are observed for subject islands and purposive islands.

[4] Maria Rita Manzini (personal communication) points out that the same conclusion can be reached in a particular version of the Minimalist Program: if each feature projects, then each projection will have exactly one feature to be checked. This can be achieved either by head movement or by XP movement. This might not be sufficient however. There are configurations with an overt head, where some constituent must still move to the Spec of that projection. Final question particles for example, are overt, but still trigger pied-piping of their complement. At the point of linearization, the projection no longer contains lexical

material in both Spec and head position. One could say alternatively that the pied-piping is triggered to satisfy the features of yet another projection.

[5] The verbal focus construction thus represents a case of *remnant* movement (cf. Den Besten and Webelhuth 1990), i.e. a case in which extraction out of a constituent is followed by subsequent movement of that constituent to some higher position.

References.

Baker, M. (1988) *Incorporation: A Theory of Grammatical Function Changing*, Chicago: University of Chicago Press.

Cinque, G. (1997) *Adverbs and Functional Heads. A cross-linguistic perspective*. To be published by Oxford Univeristy Press.

Chomsky, N. (1995) *The Minimalist Program*, MIT Press.

Kayne, R. (1995) *The Antisymmetry of Syntax*, MIT Press.

Koopman, H. (1984) *The syntax of Verbs: from Verb movement rules in the Kru languages to Universal Grammar*, Foris Publications.

Koopman, H (1996) 'The Spec Head Configuration' in Edward Garrett and Felicia Lee (eds.) *Syntax at Sunset, UCLA Working papers in Syntax and semantics*. Number 1.

Koopman, H. and D. Sportiche (1986) 'A Note on Long Extraction in Vata', *NLLT 4*, 357-374.

Larson, R. and C Lefebvre (1991) *Predicate Clefting in Haitian Creole*, in *Proceedings of NELS 21* .

Nkemnji, Michael (1995) *Heavy pied-piping in Nweh*, UCLA PhD dissertation. UCLA dissertation series.

Riemsdijk, H. van (1978) *A Case Study in Syntactic Markedness: the Binding Nature ofPrepositional Phrases*. Lisse: The Peter de Ridder Press.

Rizzi, L. (1995) 'On the Fine structure of the Left Periphery', ms. Université de Genève.

Sportiche, D. (1995) 'Sketch of a Reductionist Approach to Syntactic Variations and Dependencies', in H. Campos and P. Kempchinsky eds., *Evolution and Revolution in Romance Syntax*, Georgetown University Press.

Sportiche, D. (1994), 'Adjuncts and Adjunctions', ms. UCLA.

Sportiche, D. (1996), The Atomization and Partitioning of Syntactic Structures, ms, UCLA.

Parameters of meaning in the spatial structure of temporal semantics: an investigation of Wolof lexicon and grammar*

Kevin Ezra Moore
University of California, Berkeley.

1. Introduction. This paper is about a particular schema of motion (depicted in Diagram 1) and how it is extended to temporal use in Wolof, a Niger-Congo language spoken in Senegal and Gambia. In addition to treating temporal semantics, this investigation reveals interesting descriptive facts about Wolof and contributes to the study of crosslinguistic lexicalization patterns in words and phrases that denote translational motion (i.e., motion from one location to another).

We will be primarily concerned with issues involving a particular cognitive strategy for extending the use of spatial vocabulary to the realm of time. This strategy, which we will call the *Moving Time* metaphor, is common crosslinguistically (cf. Traugott 1975, 1978), and is inherently of interest to the study of mind and language. A comparison of the way this metaphor manifests itself in Wolof and English suggests that for the most part it remains constant relative to lexical and grammatical phenomena in the two languages. This paper investigates how lexical and grammatical structures in Wolof interact with the metaphor. An appreciation of these issues allows us to make interesting observations about the Wolof language, about how English and Wolof are similar and how they are different, and about how metaphor works in language.

The investigation is situated in the framework of the theory of conceptual metaphor as developed by George Lakoff and his associates (e.g., Lakoff 1993, Lakoff and Johnson 1980). A metaphor in this theory is a structured ensemble of correspondences between two domains, or kinds of experience. One kind of experience, the Source, is held to play a role in the linguistic and conceptual structure of the other kind of experience, the Target. For the data we are considering, the Source experience is an experience of motion and the Target experience is an experience of time.

The Moving Time metaphor[1] can be exemplified for both Wolof and English with example 2a, *Kirismas mungiy ñów* 'Christmas is coming'. In this example, Christmas is talked about as if it were a thing, like the train in example 1a. [2]

(1) a. *saxaar gaangiy* <u>*ñów*</u>
train the:PRSTV:DUR <u>come</u>
'The train is coming.'

b. *saxaar gi* <u>*agsi*</u> *na*
train the <u>arrive</u> 3PERF
'The train has arrived.'

(2) a. *kirismas mungiy* <u>*ñów.*</u>
Christmas 3PRSTV:DUR <u>come</u>
'Christmas is coming.'

b. *noor* <u>*agsi*</u> *na.*
dry.season <u>arrive</u> 3PERF
'The dry season has arrived.'

Parts (a) and (b) of Diagram 1 are a schematic representation of the scenario -- e.g. a train moving to a destination -- that gives rise to the Source experience for the metaphor exemplified in 2. (The full three stages depicted in the diagram will become relevant later on when we examine verbs of "Passing." All of the predicators that we will examine have to do with the scenario in Diagram 1.)

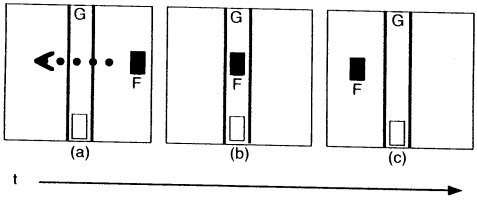


Diagram 1.

In the diagram, F stands for Figure, which is the entity whose location is in question. G stands for Ground, the entity with respect to which the location of the Figure is determined. Typically, the Figure is moving relative to a stationary Ground. The Ground is demarcated in the diagram by bold vertical lines. The white rectangle represents an entity, e.g. the speaker, located in the Ground region.

We now return to the discussion of the Source expression exemplified in 1a and the metaphorical (i.e., Target) expression in 2a. In the first stage of the Source scenario, represented by box (a) in the diagram, the Figure is located away from the Ground. The Figure then moves to the Ground, where we see it in the second stage, (b). In example 1a, the Ground is the location of the speaker and the train is moving toward this Ground. In the metaphor, a change in location of an object like the train corresponds to Christmas's change of status from future to present, as in 2a. From now on, I will use the word *Ego* to refer to the role played by the speaker in this example. More generally, the word *Ego* will denote the person who is having the experience of motion or time in question.

Metaphors are described in terms of specific correspondences in which elements of the Source *map onto* elements of the Target. Using this terminology, the Moving Time metaphor can be described as follows.

A moving thing (F) maps onto a time, for example Christmas in 2a. Ego's location (G) maps onto Ego's present moment. F's arrival at G maps onto the occurrence of the time. Continued motion, as in box (c), corresponds to a change from present to past.

Let me mention at this point that whereas space consists of three dimensions, the conceptualizations of time studied in this paper treat time as one dimensional. The diagram depicts the Source experience; in order to get a mental picture of the temporal phenomena that correspond to the diagram, one needs to imagine it in a single dimension.

There are several metaphors for time in Wolof. In this paper I will have opportunity to discuss only Moving Time.

2. Familiar patterns: coming, arriving, and passing. The English and Wolof versions of the examples in 2 are remarkably similar. This is presumably due to the universal availability of the Moving Time metaphor in human cognition

plus the high degree of similarity of the lexical semantics of Wolof *ñów* and *agsi* with the English *come* and *arrive*.[3]

In the examples in 3, some interesting differences between Wolof and English begin to show up. After a brief characterization of the semantics of the verbs involved, we will see what some of these differences are.

The verbs *romb* 'go by', *weesu* 'go beyond' and *jàll* 'get past' can all be used to denote variations on the schematic event depicted in Diagram 1, (a), (b), and (c). In the denotation of *romb* 'go by', the Figure simply goes by the Ground, without interacting with it. This is exemplified in 3a.

(3) a. *romb na bunt bi/mburngël mi/tool yi/ doj wi*
go.by 3PERF door the/tunnel the/fields the/rock the
'S/he went by the door/tunnel/fields/rock.' (Not through or over it/them)

In the case of *jàll* 'get past', the Ground is a barrier, boundary, or passage. *Jàll* highlights the interaction of the Figure with the Ground. Example 3b exemplifies the fact that *jàll* forces the interpretation that its Ground is a barrier, boundary, or passage.

(3) b. *jàll na basaŋ gi*
get.past 3PERF mat the
'#S/he went across/over the mat.' 'S/he got past the mat.' [The mat is construed as a barrier or a boundary.]

What *weesu* 'go beyond' denotes is the passage of the Figure from one to the other of two contrasting regions. This is exemplified in 3c, which is a proverb.

(3) c. *yéeg du weesu xob.*
climbing DUR:NEG go.beyond leaf
'You don't climb beyond the leaves.' [Cissé et al. 1982:61]

The idea of the proverb is that activities have their natural limits beyond which they should not be continued. Here we are not concerned with how the proverb is understood but rather with the Source scenario it depicts -- the area of a tree with leaves and an area beyond without leaves. The Ground of *weesu* demarcates the boundary between two regions -- one in which climbing should be done and another in which climbing should not be done.

The meanings of *jàll* and *weesu* are based on the same core schema of motion as those of *romb*. What distinguishes *jàll* and *weesu* is that their semantics are elaborated in key ways having to do with the purposes and expectations of the participants in the motion scenario and/or the speech event. The semantics of *romb*, by contrast, is essentially the core schematic semantics of motion.

In some dialects of Wolof, it is *weesu* 'go beyond' and *jàll* 'get past' that are conventionally extended to temporal uses by the Moving Time metaphor, and not *romb* 'go by'. The reason for this is that the Source semantics of *weesu* and *jàll* are better suited to temporal concepts than are the semantics of *romb*. *Weesu* and *jàll*, each in its own way, construe the Figure exiting one region and entering another.

An important kind of case (exemplified in 4a below for both Wolof and English) in which people talk about temporal experience with expressions of Passing is that in which what is at issue is whether a particular temporal Figure is in the future,

present, or past, where these are construed metaphorically as regions. What is important is the metaphorical region the figure is in, more than the metaphorical movement relative to the ground *per se*. It is this idea of *temporal location* that *weesu* and *jàll* are particularly well suited to express.

(4) a. *moo* *sàggan* *ba* *gàpp* *bi* *weesu*
3SUBJECT.FOCUS be.negligent to.the.point.of deadline the go.beyond
'He was so negligent that the deadline passed.'
[Fal et al. 1990:189 (my translation)]

Following is an example of both *weesu* and *jàll* in a temporal use. (In order to give the reader a feel for the Wolof structure, I sometimes provide a word-for-word translation in double quotes, preceding the free translation in single quotes.)

(4) b. *koor-gi* *weesu/jàll* *na*
Ramadan go.beyond/get.past 3PERF
"Ramadan has gone beyond/gotten past." 'Ramadan is over.' (Ramadan is the Muslim holy month of fasting.)

In examples like 4b above, all of the three native speakers consulted accept *weesu* and *jàll*, while only one accepts *romb* in place of *weesu* or *jàll*. This is evidence that i) the schematic semantics that *weesu*, *jàll* , and *romb* share is exploitable for use with the Moving Time metaphor, and ii) *weesu* and *jàll* have additional components of meaning that make them more appropriate for use with the Moving Time metaphor than *romb*.

Of the three verbs, *weesu* is the most thoroughly conventionalized in temporal uses. This observation is based on informant judgments and is supported by a survey of example sentences in Fal et al. 1990, a Wolof-French dictionary.[4] Excluding compounds, I found eighteen examples of *weesu*, eight of *jàll* and nine of *romb*. Of the eighteen *weesu* examples, eleven instantiated metaphors of temporal location, and seven of those instantiated the Moving Time metaphor. The search did not uncover any expressions of temporal location involving *jàll* or *romb*. The reason that *weesu* is the most thoroughly conventionalized in temporal uses is presumably that it is the one whose Source denotation is the most directly concerned with what kind of region the Figure is in, where "kind of region" is understood in terms of the purposes and expectations of a participant in the motion scenario. This makes *weesu* appropriate for talking about past, present, and future when they are construed metaphorically as regions, and what is significant about these regions has to do with whether Ego can interact with what is in them. The point here is that an adequate understanding of how language construes temporal experience in terms of movement requires the investigation of issues that go beyond a schematic analysis of movements and trajectories.

This kind of lexical analysis is important in three ways: First, we need it in order to have a complete description of Wolof. Second, the kinds of distinctions encoded by *jàll* 'get past' and *weesu* 'go beyond' have not received much previous attention in studies of predicators of motion.[5] Third, the analysis sheds light on questions of what elements of Source semantics are appropriate to a given Target experience, showing something about how the purposes and expectations of participants in Source experiences map onto Target experiences in certain metaphors.

3. Unfamiliar patterns.

3.1. *Fekk* 'become co-located with'. The Moving Time expressions that we have just examined are variations on patterns that are familiar from English. In the examples in 5, with the verb *fekk* 'become co-located with', we see some expressions that differ from the English pattern. The Source scenario for the *fekk* examples is depicted in Diagram 1, (a) and (b). The movement scenario for *fekk* consists of the same overall configuration as that for *ñów* 'come' and *agsi* 'arrive'; the essential difference is that *fekk* highlights the fact that what the Figure becomes co-located with is a preexisting configuration involving an explicitly denoted entity and its location. Thus, in the case of *fekk* the white rectangle in the diagram denotes a secondary figure (e.g. the bowl in 5a), which functions as part of the Ground at the level of the clause.

(5) a. *sàmba fekk* *na*
Samba <u>become.co-located.with</u> 3PERF

bool ba ca waañ wa
bowl the PREP kitchen the
"Samba became co-located with the bowl in the kitchen." 'Samba came across the bowl in the kitchen.' [*Sàmba* is a personal name.]

b. *bal ba fekk* *na ma ca kër ga*
ball the <u>become.co-located.with</u> 3PERF me PREP home the
"The ball became co-located with me at home." "The ball came across me at home." 'I was at home when the ball came flying in through the window.'

Fekk, 'to become co-located with', does not have a natural-sounding translation in English. While *fekk* is very similar to English *find*, there are crucial differences. The most important of these for our purposes is the following. *Find* predicates of an entity that it has the experience of becoming aware of something, typically but not necessarily by moving so as to encounter the found object in some location. *Fekk*, by contrast, predicates of an entity that it moves so as to occupy the same location as some previously situated entity, but the Mover need not have a cognitive experience and may in fact be inanimate.

These facts about the lexical semantics of *fekk* explain why some Source examples, like 5a, seem natural with a word-for-word English translation, while others, like 5b, do not. *Fekk* sentences with an experiencer subject seem natural, those with an inanimate subject seem odd or poetic when translated into English. Moving Time uses of *fekk* such as example 6 below, which are unmarked in Wolof, seem marked from the English point of view for the same reasons that the Source sentences (like 5b) with inanimate subjects and human objects seem marked.

(6) *benn waxtu fekk* *na ko fa*
one hour <u>become.co-located.with</u> 3PERF 3OBJ there
"One o'clock became co-located with her there"; "One o'clock found her there." 'S/he was there at one o'clock.'

That is, example 6 foregrounds a time, construed metaphorically as an inanimate mover which arrives at the location of a person. The time is treated as the most prominent participant in the clause (by virtue of being the grammatical subject) and the person is treated as less prominent. By contrast, the corresponding unmarked English sentence, *She was there at one o'clock*, treats the person as more and the time as less prominent. (On the notion of *prominence*, cf. Langacker 1991.)

This discussion of *fekk* has shown how Wolof uses its lexical semantics to create an unfamiliar kind of expression using a familiar metaphor.[6]

3.2. *Jot* 'reach, obtain'. The final verb that I want to examine is *jot* 'reach, obtain'. *Jot* has many uses, only a few of which will be mentioned here. The use of *jot* that I take to be the most basic or central is exemplified in 7 below.

(7) *jot naa téere bi ci kaw armoor bi.*
reach 1PERF book the PREP top cabinet the
'I can reach the book on top of the cabinet.'

Some target uses are exemplified in 8 below. After we have examined the Target semantics, we will look at some more of the Source uses.

(8) a. *benn waxtu jot na.*
one hour reach 3PERF
"One o'clock has (just now) reached." 'It's (exactly) one o'clock.'

b. *fajar jot na*
dawn reach 3PERF
'It's dawn.'

Temporal uses of *jot* 'reach, obtain' such as those in 8 metaphorically instantiate a variation on the same schema in Diagram 1 that we looked at in connection with *ñów* 'come' and *agsi* 'arrive'. But there are some important differences in the temporal semantics of *jot* compared to *ñów* and *agsi*.

First, Wolof speakers report that in certain contexts *jot* denotes an instantaneous temporal occurrence; for example, in the appropriate context, 8a means that it is exactly one o'clock at speech time. *Ñów* 'come' and *agsi* 'arrive' do not have this property.

Second, constructions with *jot* 'reach, obtain' are an unmarked way of talking about the occurrence of a time. The *jot* expressions are ordinary in the same sense that their English translations in 8 with the "It's X" construction are ordinary. For example you can say *It's two o'clock* without any special context, but it would take a special context to say *Two o'clock has arrived*. The same observation pertains to the Wolof translations of these two sentences, with *jot* in the first sentence and *agsi* translating *arrive* in the second.

Textual evidence for the native speaker intuition just characterized was obtained by examining occurrences of *jot*, *agsi*, and *ñów* in four Wolof texts (totaling about 61,500 words).[7] In the texts, Moving Time *jot* is used in a wide variety of situations without any contextual buildup. By contrast, Moving Time *ñów* 'come' requires special contextual conditions. These conditions often involve an element of expectation, corresponding to that part of the Source scenario where the Figure is coming but has not arrived yet.

Also, Moving Time *jot* occurs more frequently than Moving Time *ñów* or *agsi*. In the texts, there are no instances of Moving Time *agsi*. There are 29 tokens of Moving Time *jot* out of a total of 55 tokens of that verb. By contrast, there are just 13 tokens of Moving Time *ñów* out of a total of 171 tokens of that word.

In order to account for the semantic differences between *jot* and *ñów/agsi*, I propose a metaphorical mapping for *jot* which is slightly different from that of *ñów/agsi*. The Moving Time mapping for *ñów/agsi* is presented below, where the arrow, "-->", stands for 'maps onto'. (Cf. the description of the Moving Time metaphor in the introduction.)

MOVING TIME MAPPING FOR *ñów/agsi* 'come/arrive'
A moving thing(F) --> A time.
Ego's location (G) --> Ego's present.
F's arrival at G --> The occurrence of a time.

The essential difference in the mapping for *jot* is that the occurrence of a time is construed in terms of something *making contact* with Ego rather than merely arriving at Ego's location. The mapping is stated as follows.

MOVING TIME MAPPING FOR *jot* 'reach, obtain'
A thing at a distance from Ego (F) --> A time that has not occurred.
Ego as a physical entity (G) --> Ego's present.
F's achievement of contact with G --> The occurrence of a time.

The proposed mapping motivates the semantics of "exactness" because the relation of physical contact is inherently more precisely specified than the relation of being in the same place. The unmarked character of *jot* expressions is motivated by the following two interdependent considerations. i) In the case of *jot*, attention is focused on only the final stage (i.e., the achievement of contact) of the scenario. The component of travel and expectation that is associated with *ñów* 'come' and *agsi* 'arrive' is not present. ii) The *ñów/agsi* scenario requires a previously established locative relation, the relation between Ego and her location. *Jot* does not require this extra layer of conceptual structure, and is thus appropriate to denoting the mere occurrence of a time, presented without invoking additional assumptions about the situation in which the time occurs.

Independent (morphosyntactic) evidence that the metaphorical Ground in *jot* expressions corresponds to Ego rather than Ego's location is found in sentences such as 9 below.

(9) *fu* *fajar* *jot-e* *mungi* *ci* *teen bi.*
where(ever) dawn reach-ADDARG 3PRSTV PREP well the
"Wherever dawn reaches, s/he's at the well." (I.e., 'Whenever...'). 'S/he's always at the well at dawn.' [The ADDARG suffix on *jot* shows that the locative pronoun *fu* is not a canonical argument of *jot*. It is safe to assume that *jot*'s canonical argument remains implicit in this example.]

In 9, Ego's metaphorical location is coded as an argument added onto *jot*'s ordinary valence, suggesting that *jot*'s default object here is a (metaphorical) person, not a place.

In addition to 7 above (the basic/central use), the Source uses of *jot* that seem to be most closely related to the Moving Time use are those in 10 below. The notion

of *contact* is salient in all of them. The other component of meaning crucial to the Moving Time uses is that in which contact occurs as a result of the motion of a discrete entity, as in 10a, 10b, and 10d.

(10) a. "Achievement of contact." The Reacher has come into contact with the Other. Neutral as to whether the Reacher or the Other moved:
jot *na* *(ci)* *armoor bi*
reach 3PERF (PREP) cabinet the
'S/he {got to/got} the cabinet.'

b. "Arrive-reach." A minimal variant of (a) in which the Reacher is specified as a Mover and the Other as a Location:
jot *na* *ndakaaru*
reach 3PERF Dakar
'S/he has reached Dakar.'

c. "Extent-reach." A minimal variant of the central use in which the Reacher is inanimate:
buum gi jot *na*
rope the reach 3PERF
'The rope reaches.'

d. "Catch." A variant of (a) in which the Other is moving away from the Reacher:
gaynde gi *jot* *na* *ko*
lion the reach 3PERF 3OBJ
'The lion caught her.'

3.2.1. *Jot* 'reach, obtain' and Wolof grammar. We have seen two cases in which Wolof has temporal expressions that are markedly different from English. In the case of *fekk* 'become co-located with', the difference can be attributed to the lexical semantics of *fekk*. But there does not seem to be anything about the lexical semantics of *jot* that would account for why its Moving Time uses are unfamiliar to English speakers. Why should Wolof but not English have constructions like the temporal *jot* constructions?

I would like to point to three kinds of phenomena in Wolof grammar that play a role in motivating the temporal *jot* constructions: i) Parameters of subject selection. ii) Membership in the category *verb*. iii) Tense/aspect marking and lexicalization patterns involving aspect. What I want to motivate in this section is the fact that it is unmarked in Wolof to talk about the occurrence of a time with a construction in which a point or period of time is the subject of a transitive verb (i.e., *jot*).

Let's look at subject selection first. In certain cases in which there is an interaction between a human participant and something inanimate in her experience, Wolof requires that the inanimate take the grammatical role of subject and the person take the role of object. The point is that Wolof has ordinary ways of talking about experience, in addition to the *jot* constructions, that parallel the way *jot* predications construe a time as contacting a person. We have already seen one example of this phenomenon in the *fekk* constructions (examples 5 and 6), where movement overrides animacy as a criterion for subject selection. Another example involves *xeeñ* in 11 below, where a smell is the subject and the person who smells it is the object.

(11) *cere* *ji* *xeeñ* *na* *ma*
couscous the smell 3PERF 1OBJ
'I can smell the couscous.' [Munro and Gaye 1991:143]

Since time is a component of setting, temporal subjects of *jot* fit into a broader pattern whereby setting subjects are unmarked in Wolof. Note example 12.

(12) *fii* *rafet* *na*
here be.beautiful 3PERF
"*Here is beautiful." 'It's beautiful here'; 'This is a beautiful place.'

Furthermore, it is common in Wolof to treat time-expressions (e.g. *ci kanam* 'later', *ëllëg* 'tomorrow') in subordinate clauses as verbs, thus construing the denotatum of the time-word as a process rather than a thing. This is exemplified in 13 below. What this has in common with the *jot* constructions is that they are both unmarked ways of construing a time as doing something when it occurs.

(13) *bu* "ten"-*ee*
REL ten-COND
"when it's ten o'clock"; "when it tens." 'at ten o'clock'

What we have just seen are arguments that apply to all uses of Moving Time *jot*. The following discussion pertains to *jot* as it is used in constructions like the one exemplified in 8a (*benn waxtu jot na* 'It's one o'clock'). This construction portrays the occurrence of a time both as a punctual event and as a transitory state that obtains at the moment of speech. What I want to discuss here is how Wolof grammar helps make it possible to use a punctual verb to talk about a present state. In addition to what we have already examined, two further phenomena of Wolof grammar are relevant here: the Wolof Perfect and a particular lexicalization pattern.

The first thing we should note about the Wolof Perfect is that it differs from the English Perfect in two interrelated ways: i) The Wolof Perfect does not encode tense; ii) The Wolof Perfect marks not only action verbs like *daanu* 'fall', but also state verbs like *baax* 'be good'. (See Anderson 1982 for a crosslinguistic characterization of the perfect.) What the Wolof Perfect marker marks is the verb root, not a participle. The Perfect-marked action verb typically denotes a past event. The Perfect-marked state verb typically denotes a present state. (Cf. Robert 1991.) This contrast is exemplified below.

(14) a. *baax* *na*
good 3PERF
'It's good.'

b. *daanu* *na*
make.fall:MIDDLE 3PERF
'It fell down.'

Next, we note a pattern whereby some Wolof verbs can encode either an event or the state that results from the event (cf. Jackendoff 1990, Talmy 1985:88-9). While investigation is still in progress, it appears that the class of Wolof verbs that have this alternation is much larger and covers a greater semantic range than the

corresponding English class. Notably, the Wolof class includes some verbs of translational motion, (e.g. *sore* 'get/be far from') as well as "quality" verbs like *forox* 'get/be sour.'[8] The Wolof alternation is compatible with but not limited to occurrence with the Perfect. It is exemplified below.

(15) a. *dee na ca ja ba*
die 3PERF PREP marketplace the
'S/he died in the marketplace.'

b. *dee na*
die 3PERF
'S/he's dead.'

Because the Perfect can mark either a past event or a current state, and the verb root itself can denote either an event or its resulting state, it is possible for a construction to be indeterminate between a past-event and a current-process reading. This is exemplified below. (Cf. Vendler 1967:97-121.)[9]

(16) *gis naa ko.*
see 1PERF 3OBJ
"I spotted it and I see it." 'I have caught sight of it now.'
'I saw it.' 'I have seen it.' 'I see it.'

Since Wolof does not encode tense in this construction, the temporal scope of predication can include both the punctual event and its immediately resulting state.[10] The English Perfect is different in the case of punctual events in that it explicitly refers to something in the past. Whereas 16 above can mean "I have spotted it and I see it," this meaning is not so readily denoted by a simple, unmarked English clause. Because of these facts of Wolof and English grammar, Wolof is better disposed than English to use a punctual verb to refer to the present moment, and is thus better disposed to have an expression like the one in 8a (*benn waxtu jot na* 'It's one o'clock.') According to this analysis, the semantics of Moving Time *jot* has a punctual and a stative component, corresponding to the "spotting" and the "seeing" components of *gis* 'see' in 16. Thus 8a is well paraphrased as "One o'clock has struck and it's one o'clock," where *jot* denotes the punctual occurrence and the resulting state.

4. Conclusions. We started this investigation by examining a simple schema of movement and some of the ways that schema is elaborated by the Wolof lexicon, and we examined the role that this schema and its lexical elaborations play in the domain of temporal experience. We saw that an investigation of the Moving Time metaphor in Wolof yields some interesting results regarding the precise nature of the Source-Target correspondences involved in the metaphor. These results show up because of the way words like *romb* 'go by', *weesu* 'go beyond', and *jàll* 'get past' interact with the metaphor. Finally, we looked at the interaction between the Moving Time metaphor and the lexemes *fekk* 'become co-located with' and *jot* 'reach, obtain'. I argued that the same fundamental metaphorical correspondences that yield familiar-sounding expressions in the former cases yield unfamiliar sounding expressions with *fekk* and (with slight modification) *jot*. In the case of *fekk*, this could be accounted for by a simple appeal to the lexical semantics of that

verb. In the case of *jot* it was necessary to take into account more general aspects of Wolof grammar. The findings reported here support the idea that a metaphor can function as a conceptual invariant relative to other conceptual and linguistic phenomena.

Notes

*Thanks to Paap Alassane Sow for data, guidance, and inspiration. And thanks to Stéphane Robert for inspiration and helpful suggestions. Also, this paper couldn't have been written without the help of the dissertation support group at U.C. Berkeley.

1. The Moving Time metaphor is equivalent to Special Case 1 of the metaphor Time Passing Is Motion in the terminology of Lakoff 1993.

2. The following abbreviations are used for the data: ADDARG: Added Argument; COND - Conditional; DUR - Durative; EMPHVERB - Emphatic of the Verb; NEG - Negative; OBJ - object; PRSTV - Presentative (similar to the English progressive in many uses); PRT - Particle; PERF - Perfect; PL - plural; PREP - Preposition (typically locative); REL - Relativizer. (On EMPHVERB, PRSTV, and PERF see Robert 1991.)

Examples are transcribed according to the official Senegalese transcription system (cf. Fal et al. 1990). Correspondences between the Senegalese system and the International Phonetic Alphabet are listed below, with the Senegalese symbol on the left and its IPA equivalent on the right. In all cases not mentioned, the Senegalese symbol has the IPA value. é = e; e = ɛ; ë = ə; a = (a low central vowel); à = a more open a̱ ; ó = o; o = ɔ; ñ = ɲ; j = ɟ (voiced palatal stop); y = j. Capital and lower case symbols have the same value. Geminates are indicated by doubling the symbol in question except for *q* which represents [q:]. In the case of long vowels, a single diacritic modifies both symbols. For example, óo represents [o:]. Word-final stops are devoiced.

3. The word *agsi* is actually morphologically complex, consisting of the root *àgg* 'arrive' and the suffix *-si* which indicates that the subject of the verb comes to the deictic center to perform the action denoted by the verb. This fact does not affect the current analysis because in the Moving Time metaphor, whenever a time arrives, it arrives at the deictic center (in both Wolof and English).

4. Actually, I searched for both *wees* and *weesu*, treating them as variants of the same lexeme.

5. Jackendoff 1990 and Talmy 1991 are examples of important research on predicators of motion that do not touch on this issue. However, Talmy (1985 and elsewhere) briefly mentions some similar cases. Herskovits 1986 and Vandeloise 1991 deal with closely related issues for prepositions.

6. In addition to the kind of use described here, *fekk* appears in a range of constructions that introduce clauses which indicate that a certain state of affairs already obtained at a given point in the narrative. More on *fekk* can be found in Moore *in press*.

7. I counted both *àgg* and *agsi* and their morphological variants. Cf. note 3.

The texts are Caam 1989 (about 10,650 words), Jeŋ 1992 (about 22,450 words), Kesteloot and Mbodj 1983 (about 19,500 words), and Ndaw 1993 (about 9000 words). Kesteloot and Mbodj 1983 is a collection of transcribed oral literature; the others are novelettes.

8. The investigation of this alternation has really just begun. *Forox* 'be sour' is the only quality verb I have investigated so far.

9. Comrie 1976:57, citing Welmers (1973:347-8) notes a similar phenomena in Kpelle, where "... even 'see' is expressed as a Perfect, i.e. *aâ `káa* can correspond to either 'he sees it' or 'he has caught sight of it'."

10. The notion *scope of predication* comes from Langacker, e.g. 1987. In the case in question here, the temporal scope of predication is the span of time that the denotation of an expression pertains to.

References

Anderson, Lloyd. 1982. The "perfect" as a universal and as a language-specific category. Tense-aspect: between semantics and pragmatics, ed. by Paul Hopper, 227-264. Amsterdam: John Benjamins.

Caam, Saxiir. 1989. Njaaxum: léeb. Dakar: IFAN Cheikh Anta Diop.

Cissé, Mame Thierno; Mammadu Gueye; and Mangoné Touré. 1982. Proverbes wolof. Dakar: Centre de Linguistique Appliquée de Dakar.

Comrie, Bernard. 1976. Aspect. Cambridge: Cambridge University Press.

Fal, Arame; Rosine Santos; and Jean Léonce Doneux. 1990. Dictionnaire wolof-français. Paris: Karthala

Herskovits, Annette. 1986. Language and spatial cognition: An interdisciplinary study of the prepositions in English. Cambridge: Cambridge University Press.

Jackendoff, Ray. 1990. Semantic structures. Cambridge: The MIT Press.

Jeŋ, Maam Yunus. 1992. Aawo bi. Dakar: IFAN Cheikh Anta Diop.

Kesteloot, Lilyan; and Chérif Mbodj. 1983. Contes et mythes wolof. Dakar: Les Nouvelles Editions Africaines.

Lakoff, George. 1993. The contemporary theory of metaphor. Metaphor and thought, ed. by Andrew Ortony, 202-251. Cambridge: Cambridge University Press [Second edition.]

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

Langacker, Ronald. 1987. Foundations of cognitive grammar 1. Stanford: Stanford University Press.

---------- 1991. Foundations of cognitive grammar 2. Stanford: Stanford University Press.

Moore, Kevin Ezra. In press. The conventional association of a lexeme with a metaphor: the case of the Wolof verb *fekk-e*. Lexical and syntactical constructions and the construction of meaning, ed. by Marjolijn Verspoor. Amsterdam: John Benjamins.

Munro, Pamela and Dieynaba Gaye. 1991. Ay baati Wolof: A Wolof dictionary. Los Angeles: Department of Linguistics, University of California. [UCLA occasional papers in linguistics, number 9]

Ndaw, Séex Aliyu. 1993. Buur Tilleen. Dakar: IFAN Cheikh Anta Diop.

Robert, Stéphane. 1991. Approche énonciative du système verbal: le cas du wolof. Paris: Éditions du Centre National de la Recherche Scientifique.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. Language typology and syntactic description 3, ed. by Timothy Shopen, 57-149. Cambridge: Cambridge University Press.

----------1991. Path to realization: A typology of event conflation. Berkeley Linguistics Society 17. 480-519.

Traugott, Elizabeth Closs. 1975. Spatial expressions of tense and temporal sequencing: A contribution to the study of semantic fields. Semiotica 15, 3: 207-230. Mouton Publishers.

----------1978. On the expression of spatio-temporal relations in language. Universals of human language 3, ed. by J.H. Greenberg, C.A. Ferguson, and E. Moravcsik, 369-400. Stanford: Stanford University Press:

Vandeloise, Claude. 1991. Spatial prepositions. Chicago: The University of Chicago Press.

Vendler, Zeno. 1967. Linguistics in philosophy. Cornell University Press.

Welmers, Wm. E. 1973. African language structures. Berkeley: University of California Press.

Argument structure and locus of affect in the Maasai External Possession construction*

Doris L. Payne
University of Oregon & Summer Institute of Linguistics

The Maasai External Possession (EP) construction can be fully accounted for only if the grammatical object of the verb is understood as simultaneously having two distinct semantic features: Possessor and Affected. Whether these features are both semantic roles bears on the adequacy of those theories of syntax which claim that a nominal in any given sentence must have at least one, but no more than one, semantic (thematic) role. This study in part explores the feature Affected and whether it is a core feature of a semantic role, or whether it is a distinct type of conceptualization intimately tied to the speaker's ability to shift his or her "take" on the Starting and Ending point of an event. The Maasai data also bear on the adequacy of lexically centered views of syntax, succinctly expressed as "syntax is projected from the lexicon," versus more complex verb-plus-construction views of the semantics-syntax interface. The Maasai EP construction adds to the body of data supporting the latter as a more adequate theoretical stance.

1. The general problem. The Theta Criterion states that an NP in a sentence must have at least one semantic role, but no more than one.[1] Otherwise, incomprehensibility follows because NPs are not "licensed," as in: **The elephant the cheese saw the idea the generator* (Napoli 1993:108). In some works the Theta Criterion is articulated with reference just to arguments (Napoli 1993:109), while in others it is stated with reference to all lexical NPs (van Riemsdijk and Williams 1986:131). Under the first view, an argument gets its semantic role from being governed by a verb (cf. Gruber 1976). Under the second view, a nominal may get its semantic role from being governed by a verb, an adposition, or a noun as in a possessed NP (ex. 1). Semantic roles are generally assumed to include at least Agent (AG), Patient (PAT) or Theme (TH), and Goal.[2] Many frameworks add Source, Instrument, Benefactive (BEN), Location, Direction. Baker (1988) also includes Possessor, as in (1c).

(1)	a. *The <u>tyrant's</u> destruction of the city*	AGENT
	b. *The <u>city's</u> destruction*	PATIENT
	c. *<u>John's</u> backpack*	POSSESSOR

Though the Theta Criterion has been widely accepted by linguists of many allegiances, it is challenged by data from a wide variety of languages. The following data illustrate five situations where, at first glance, we might be tempted to think that a phrase has two distinct semantic roles in that the phrase at least has "colorings" of two distinct semantic features.

Based on examples like those in (2), Jackendoff (1972:34-35) suggests that a single argument (underlined) can be simultaneously Agent and Theme, Agent and Source, or Agent and Goal. (2a) is especially of interest here in that *Max* both undergoes the action and is somehow responsible for the action.

(2)	a. *By his own volition, Max rolled down the hill.*	AG/THEME
	b. *Reuben sold Fred some hashish on purpose.*	AG/SOURCE
	c. *Fred bought some hashish from Reuben on purpose.*	AG/GOAL

As a second instance, consider certain "Dative"-shift phenomena. The core argument frame of verbs like *bake* contain only an Agent and a Patient (3a). That they do not include a Benefactive (or a Goal) is shown by the contrast between (3b-c), where *me* can be interpreted as a Benefactive only if the oblique (OBL) marker *for* occurs.

(3)	a.	*She*	*baked*	*a cake.*	
		AG		PAT	
	b.	*She*	*baked*	*a cake*	*for me.*
		AG		PAT	OBL-BEN
	c.	*?She*	*baked*	*me.*	
		AG		PAT	

Example (3d) contains the "shifted" variant of (3b). The question arises as to the semantic role of *me* in the shifted form. The *but*-clauses in (3b', d', e) negate an intended Goal. Because some native speakers find (3b') less odd than (3d'), it could be argued that *me* simply counts as a Goal in (3d), and not as a Benefactive (cf. Goldberg 1995:141). If the role is simply Goal, there is no challenge to the Theta Criterion. However, that some sense of the Benefactive role is retained in the shifted form is suggested by the contrast between (3d') and (3e). In (3e) it is almost impossible to add the Goal-negating phrase *but didn't intend to give it to me*, whereas it presents much less of a problem for (3d').

(3)	d.	*She*	*baked*	*me*	*a cake.*
		AG		BEN/GOAL?	PAT

b'. *She baked a cake for me, but didn't mean to give it to me.*
d'. ?*She baked me a cake but didn't mean to give it to me.*
e. **She wanted to give me a cake but didn't mean to give it to me.*

Whatever one might decide about role assignment for the data in (3), the English shifted sentences are the rough equivalent of what many other languages accomplish with morphological applicatives. Applicative objects are a third case where colorings of two distinct semantic features arise. An applicative overtly

signals that a syntactic object has some semantic role other than Patient, even though the basic frame of the verb in question might normally require a Patient object (cf., Nomatsiguenga [Arawakan], Wise 1971; Kinyarwanda [Bantu], Kimenyi 1978). There is often a greater sense that the applied object is the final locus of effect (Croft 1992), but also very importantly something other than Patient (such as Goal, Benefactive, Instrumental, Locative).

As a fourth instance, consider Comitatives. It seems intuitively clear that the Comitative in (4a) is colored with agentive-ness in that James also rode to the store, while the Comitative in (4b) is colored with patient-ness in that the mashed potatoes are affected by the action of the verb.

(4) a. *Mary rode her bike to the store with James.*
AG PAT COMITATIVE=AG
b. *Mary ate her peas with mashed potatoes.*
AG PAT COMITATIVE=PAT

As a fifth instance, Klaiman (1988) and Kemmer (1993) characterize reflexives and middles as constructions in which the subject simultaneously has features of both Source (or Agent) and Affected. Klaiman elaborates that "Another middle voice function, referred to by Barber as the 'plain middle' [Barber 1975], is that of showing the subject's beneficiary status vis-à-vis the action; i.e., for conveying the subject's dual status, as source (performer) of the action and as affected entity, or locus of the action's effects" (1988:31). The following Greek examples illustrate:

(5) a. hair-o: moiran
take-ACTIVE share
'I take a share.'

b. hair-oumai moiran
take-MIDDLE share
'I choose (take for my own benefit) a share.'

(6) a. politeu-o:
be.citizen-ACTIVE
'I am a citizen/have civic rights.'

b. politeu-omai
be.citizen-MIDDLE
'I act as a citizen/carry out my civic rights for myself.'

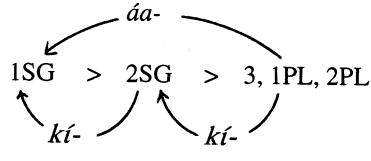
In sum, we find a variety of phrases expressing a combination of semantic features which are commonly viewed as pertaining to two distinct semantic roles.

Whether or not such data falsify the Theta Criterion of course depends on how a given theory determines what semantic roles are: whether their existence depends on the verb or on the construction; whether roles are discrete and distinct versus fuzzy in categoriality; how broad or narrow the inventory of roles is; and whether each of the relevant features in the preceding examples does in fact pertain *to* a semantic role.

In what follows, I will try to demonstrate two things: First, in the Maasai EP construction grammatical objects necessarily have semantic features of both Possessor and Affected. Some linguists identify Affected as a core feature of Patient/Theme. For others, it is indicative of a distinct type of conceptual status. Second, it is the EP *construction* which assigns both of these features to the grammatical object – not the verb.

2. A primer of Maasai morphosyntax, and the EP construction. Maasai has a relatively clear division between basic one-argument and two-argument verbs. The former – verbs like *kʊɛt* 'run' or *ishu* 'be alive' – cannot take inverse person prefixes (cf. 8); the latter can – verbs like *duŋ* 'cut' or *dɔl* 'see' (cf. 9). The inverse prefixes are a clear sign of transitive predicates since they signal that the subject is lower on a person-number hierarchy than the object. The Maasai hierarchy for singular persons is: 1 > 2 > 3 (Payne, Hamaya and Jacobs 1994). If second singular is subject with first singular as object, or if third (or any plural) is subject with second singular as object, then the inverse prefix is *(e)kí-* (INV). If third (or any plural) is subject with first singular as object, then the inverse prefix is *áa-* (3>1). The arrows go from person-of-subject to person-of-object:[3]

(7) Maasai person-number hierarchy:



(8) a. *áa-ishú enk-aí
3>1-be.alive FSG-God.NOM
('God will enliven me / cause me to be alive.')

b. *áa-kúɛ́t
3>1-run
('He will run me / He will cause me to run.')

c. *kí-pír
INV-be.fat
('He will fatten you / will be fat with reference to you.')
('You will fatten me / will be fat with reference to me.')

(9) a. áa-dɔ́l en-kínè
3>1-see FSG-goat.NOM
'The goat will see me.'

b. kí-duŋ
INV-cut
'He will cut you / You will cut me.'

Tone distinguishes two morphological cases in Maasai. By tradition these are referred to as "nominative" and "accusative." The "nominative" marks grammatical subjects when they follow the verb, and phrases governed by the oblique preposition *tɛ*. All other phrases occur in the "accusative" (Tucker and Mpaayei 1955).

In possessed NPs, the head noun precedes the genitive Possessor. If the Possessor is expressed by a lexical noun, a genitive particle occurs between the two nouns indicating gender and number of both Ns. If the Possessor is expressed by a possessive pronoun, no genitive particle occurs (Tucker and Mpaayei 1955). In the following, the possessed NPs are bracketed, reflecting the fact that the Possessor is internal to the possessed NP.

(10) a. ké-yyetú-tò [ɛn-kɛráí [àì]
3-scream-PROG FSG-child.NOM my
'My child is screaming.'

b. ké-yyetú-tò [ɛn-kɛráí o ɔl-payyán]
3-scream-PROG FSG-child.NOM MSG.POSR MSG-man.ACC
'The man's child is screaming.'

c. ké-yyetú-tò [ol-ayyoní l-o
3-scream-PROG MSG-boy.NOM MPOSD-M.SG.POSR

ɔl-payyán]
MSG-man.ACC
'The man's boy is screaming.'

Maasai has rich morphology for modifying argument structure, including impersonal passive, middle, antipassive, causative, dative and instrumental applicative suffixes. For some verbs, directional affixes also affect argument structure. Ex. (11) illustrates the dative applicative, which adds a Benefactive or a Goal-reached (not just Goal toward which action is directed) to the argument frame. In (11a, c, e), the direct prefix *e-* codes the subject as third and the object

as either third or some plural person. In (11b, d), the inverse *áa-* prefix codes the subject as third person and the object as first singular.

(11) a. é-dúŋ ɛn-kɛráí en-kíné
3-cut FSG-child.NOM FSG-goat.ACC
'The child will cut the goat.'

b. áa-duŋ-okí ɛn-kɛráí en-kíné
3>1-cut-DAT FSG-child.NOM FSG-goat.ACC
'The child will cut the goat for me.'

c. é-yá ɛn-kɛráí en-kíné tɛ nanú
3-carry FSG-child.NOM FSG-goat.ACC OBL 1SG.NOM
'The child will carry the goat toward/from/with reference to me.'

d. áa-ya-kí ɛn-kɛráí en-kíné
3>1-carry-DAT FSG-child.NOM FSG-goat.ACC
'The child will carry the goat to me / The child will carry the goat for me.'

e. é-yá-kɪ ɛn-kɛráí ɔl-payyán en-kíné
3-carry-DAT FSG-child.NOM MSG-man.ACC FSG-goat.ACC
'The child will carry the goat (all the way) to the man.'

There are a few ditransitive roots, such as *ɪshɔ* 'give'. These roots simply take two accusative case phrases for the Goal and Patient, without addition of an applicative.

We now turn to the Maasai EP construction. In this construction, the Possessor need not be expressed internally to the NP which contains the Possessed N. Rather, for first and second person, the Possessor is marked as the object on the verb, as shown by the inverse prefixes (compare 12a with 12b-c). For third person possessors, the Possessor precedes the Possessed N with no genitive particle occurring at all (12d).

(12) a. é-yyétú-tò [ɛn-kɛráí àì]
3-scream-PROG FSG-child.NOM my
'My child is screaming. '

b. k-áa-yyetú-tò ɛn-kɛráí
DSCN-3>1-scream-PROG FSG-child.NOM
'My child is screaming.'

c. kí-yyétú-tò ɛn-kɛráí
INV-scream-PROG FSG-child.NOM
'Your child is screaming.'

d. é-ya ɔl-tóŋání [en-kitók] [ol-coní]
3-take MSG-person.NOM FSG-woman.ACC MSG-skin.ACC
'The person/man will take the woman's skin.'

There are important dialect differences regarding what kinds of items an external possessor (EP), marked as object on the verb, can be construed as owning (Table 1). Most data in this paper draws are from IlKeekonyokie Maasai.

Table 1. EP Dialect variation (Payne 1997)

	External possessor can be construed with:
Restricted, Arusha	body part objects
Intermediate, IlUasinkishu	body part objects, intrans unaccusative subjects
Liberal, IlKeekonyokie	liberal range of nouns: objects, trans & intrans subjects

Despite rich morphology for changing argument structure, the EP construction has no verbal marking of any argument-changing operation. Rather, the only index of the construction is the presence of one extra argument in the clause than the verb in question normally allows. In this one situation, basic intransitives behave as transitives, taking inverse prefixes. For example, we have seen in (8a) that *ishu* is intransitive, further supported by the data in (13a-b). But if the object marked on such an otherwise intransitive verb is construable as Possessor, the sentence is acceptable (13c-d). This is possible only if a lexical phrase is available to be interpreted as a possessed item.

(13) a. é-íshú en-kínè
3-be.alive FSG-goat.NOM
'The goat is alive. '

b. * é-íshú enk-aí en-kíné
3>1-be.alive FSG-God.NOM FSG-goat.ACC
('God will enliven the goat/be alive with reference to the goat. ')

c. áa-ishú en-kínè
3>1-be.alive FSG-goat.NOM
'My goat is/will be alive (and I am benefited thereby).'
(*'The goat enlivens me/is alive with reference to me.')

d. kí-pir en-kínè
INV-be.fat FSG-goat.NOM
'Your goat is/will be fat (and you are benefited thereby).'

Similarly, basic transitives can behave as ditransitives with no argument-changing morphology when the verb-marked object is interpreted as Possessor. Example (14a) shows that *dɔl* 'see' is transitive and not ditransitive, in that it can only take one Agent-Subject and one Patient-Object. The sentence is ungrammatical in the IlUasinkishu dialect which only allows EPs to be construed with body part accusatives. The speaker explained what was wrong with (14a) by saying that it first sounds as if the goat sees 'me', but then 'the woman' too. If, however, a body part like 'hand' is substituted in the otherwise identical sentence, it is completely acceptable with the understanding that it was 'my hand' that the goat saw (14b).

(14) a. *áa-dól en-kínè en-titó (IlUasinkishu dialect)
3>1-see FSG-goat.NOM FSG-woman.ACC
'The goat will see me (Patient)... the woman (Patient).'

b. áa-dól en-kínè enk-áiná
3>1-see FSG-goat.NOM FSG-hand.ACC
'The goat will see my hand (to my detriment or benefit).'

To summarize, one-argument verbs can take two arguments only if the second argument is construable as Possessor (marked in the inverse prefixes in 13c-d), but not if construed as Patient (8a-c). The same is true, *mutatis mutandis*, for basic transitive verbs (14a). Thus, Possessors are unique in that a Possessor can be mapped onto an object when the Patient role cannot. Recognition of the Possessor semantics is essential, as simply trying to reconceptualize the EP object as a Patient gives no way of accounting for the ungrammaticality of (8a-c) as opposed to the grammaticality of (13c-d). The Possessor is also unique compared to Benefactive, Goal and Instrumental roles, in that these can only be expressed as non-oblique arguments of otherwise (in)transitive verbs if there is an overt applicative on the verb. The Maasai EP construction, in contrast, takes no applicative morpheme.[4] What may allow this uniqueness is that a semantic Possessor can be conceptually integrated with another participant (whether Agent or Patient) before the complex participant as a whole is conceptualized as a core participant of the verb.[5]

3. Affectedness and the EP construction. In terms of what has been presented so far, one might assume that the role of the EP-object is everywhere just Possessor. But a second semantic issue arises when we observe that the EP construction is used only when an owner is conceptualized as being somehow

affected by the entire state of affairs or action expressed by the predication. That is, the speaker wishes to convey that the ultimate Locus of Affect (Klaiman 1988, Croft 1992) is the owner – and not the owned item or any other participant that might be present.

Thus, in some dialects the EP construction is allowable only with transitive roots and with body part accusatives or items metaphorically construable as extensions of the body (e.g., Arusha; Table 1). This restriction makes pragmatic sense because if part of my body is affected, then I am also necessarily affected. In other dialects, there can be syntactically ambiguous EP sentences with competition between two nouns for which will be interpreted as possessed. In such sentences, what is often grammaticized as an inalienable noun in other languages will be necessarily interpreted as the possessed item (i.e., relationship expressions and particularly body parts; cf. IlKeekonyokie Maasai). This preference makes pragmatic sense because if someone or something "close" to me is affected, I may be indirectly affected -- perhaps in an emotional sense or in terms of general well-being. Finally, if a speaker cannot conceive of any way in which a Possessor would be even indirectly affected by the scene depicted in a sentence, then the EP version is rejected. Thus, some "remote" items such as kraal gates can occur as the head N in possessed NPs (ex. 15b), but not as possessed by an EP. Furthermore, there may be some pragmatic preference for physical contact verbs in the EP construction, as opposed to perception or cognition verbs like 'see', further underscoring the Affectedness feature.[6]

The following judgments, confirmed by two IlKeekonyokie speakers, illustrate. Example (15a) has an item very close to the body as the accusative NP, perhaps as a metaphorical body part.[7] In contrast, (15b) does not have anything even metaphorically construable as a body part, and the sentence is less acceptable. Presumably this is because a kraal gate is not viewed as closely associated with any particular individual.

(15) a. áa-ból ɔl-páyyàn ɪm-beniá
3>1-open MSG-man.NOM FPL-pockets.ACC
'The man will open my pockets. (*'My husband will open the pockets.')

b. ?áa-ból ɔl-páyyàn ɛn-kishómì
3>1-open MSG-man.NOM FSG-gate.ACC
?'My husband will open the gate.' (*'The man will open my gate.')

Relative to (15a), the 'my husband' reading is possible over 'my pockets' if the verb is put into the antipassive form (16). In the antipassive, the EP can no longer be construed as owning the pockets, even when an extra accusative NP does occur. (An accusative NP is not generally permissible in a simple antipassive clause.)

(16) áa-bol-ishó ɔl-páyyàn ɪm-beniá
3>1-open-ANTIPAS MSG-man.NOM FPL.pockets.ACC
'My man/husband has the habit of opening people's pockets.'
(*'The man has the habit of opening my pockets.')

In IlKeekonyokie Maasai, and for intransitive subjects in IlUasinkishu Maasai, kin terms and pragmatically alienable items can be construed with the EP when not in competition with a body part.

(17) a. áa-bʊak-ɪtá ɛn-kerái
3>1-bark-IMPF FSG-child.NOM
'My child is shouting.'

b. áa-bʊak-ɪtá ol-díà
3>1-bark-IMPF MSG-dog.NOM
'My dog is barking.'

Again, a contrast is seen between kin terms and other potentially closely-associated items in that the former will preferably be construed with the EP. A particularly insightful expression of this is quoted below (18).

(18) áa-pík ɛn-kɛráí ɪn-kɪlání ɛn-káré
3>1-put FSG-child.NOM FPL-clothes.ACC FSG-water.ACC
'My child will put clothes in the water.'

"Here the ambiguity of whether the child or clothes or both are mine hardly arises. It is obvious that it is the child that is mine. The clothes may be mine but that ownership compared with that of the child is quite distant." (Philip Koitelel)

The extent to which the EP construction is associated with an affect on the Possessor is surely conveyed by the translation that was given to (19c).

(19) a. ké-purróò ɛn-kɛráí in-tokitín
3-steal FSG-child.NOM FPL-things.ACC
'The child will steal things' (with no sense of shame on any parents).

b. ékí-púrróò ɛn-kɛráí in-tokitín
INV-steal FSG-child.NOM FPL-things.ACC
'Your child will steal some goods' (and you as the parent are going to be shamed with a bad name)

c. áa-purr-isho εn-kεráí
3>1-steal-ANTIPAS FSG-child.NOM
'The problem I have is that my child is a thief, so I have a lot of shame.'

To summarize, the more intimate an item is to the EP marked in the verb, the more likely the EP will be construed as owning *that* item over other possible entities in the sentence. This follows from the basic Affectedness constraint governing felicitous use of this construction: if a "closely possessed" item is involved in an event or situation, whatever happens to that item is more likely to affect the Possessor – as the final Locus of Affect – than if a pragmatically alienable or very "remote" item is involved in the situation. In sum, Affectedness appears to be part of the *meaning* of the Maasai EP construction -- not just something that might be inferred from it.

4. Affected: A semantic role? Or something else? We have seen that two semantic features are necessarily involved in the Maasai EP construction: the object marked on the verb is necessarily interpreted as Possessor (not as Patient); and the participant referenced by the verb-marked object is also necessarily interpreted as Affected. If both are semantic roles, then these data falsify the Theta Criterion. One ready way to avoid such a violation would be to claim that in a sentence like (13c) there are really two propositions or clauses: 'The goat is alive, and it [=the goat being alive] affects me.' Each argument then carries only one semantic role in its respective proposition. Subsequently, some sort of clause integration occurs such that there is only one surface clause. However, this solution is necessarily abstract (positing a zero predicate), and seems primarily motivated by trying to save the one-role-per-argument principle.

For some linguists, the EP might simply be said to have the role of Patient/Theme, either on the premise that Affected is the prototypical feature of Patient (cf. discussion in 4.1 below), or on the premise that Possessor is not a bona-fide semantic role because no basic verb ever assigns such a core role. Overall, I believe this view effectively ignores the issue of the Possessor feature.

For strong adherents of the Government Binding tradition, the EP must have the role of Possessor in order to satisfy both the Theta Criterion and the Projection Principle (e.g., Baker 1988); this effectively ignores the Affectedness feature (cf. Shibatani 1994 for some discussion). In this tradition there is little attempt to probe the semantic duality of EPs or of the ontological bases of semantic roles like Patient and Possessor.

Yet other linguists have advocated conceptual-semantic categories of Starting and End Point/Locus of Affect that mediate between the familiar semantic roles and grammatical relations. This type of analysis does justice not only to the EP construction, but also to other types of morphosyntax that allow the speaker to express varying conceptualizations of event boundaries, without ignoring basic lexicalization patterns in verb roots.

From here on, I will simply assume that Possessor (or more accurately, Genitive understood in a semantic sense) is as much of a semantic relation as are other typically-oblique roles like Benefactive, Source, Instrument, and even Goal – which many languages can bring into the core argument frame of derived verb stems via applicative morphology. I will now turn to two proposals bearing on the feature Affected.

4.1. Prototype views of the Patient semantic role. Givón (1984:139) assumes thematic roles of Agent, Dative (covering Experiencier, Recipient, and Benefactive), Patient, Locative, Instrument, Associate, and Manner. He takes a prototype approach (cf. Rosch and Mervis 1975), in which particular exemplars can be good or less-good instances of their category. In his view, "total affectedness" is characteristic of the prototype Patient, in that total affectedness is involved in a prototypically transitive action necessarily including a Patient (1984:88, 164). Given that the Maasai EP is necessarily interpreted as affected by the situation described by the sentence, under Givón's view we might suggest that the EP thus approximates to the Patient role.

Also following a Roschean framework, Dowty (1991) argues for two proto(typical)-roles of Agent and Patient. The Patient role is characterized by the features outlined in (20). The more of these features a given argument has, the better instance of a Patient it is; but failure to have one or more features does not, in itself, remove the argument from the Patient category.

(20) a. undergoes change of state
b. incremental theme (i.e., the theme can be incrementally affected during the realization of a telic predication)
c. causally ***affected*** by another participant
d. stationary relative to movement of another participant
(e. does not exist independently of the event, or not at all)

Givon's and Dowty's systems provide partial motivation for why EPs should be coded as objects in Maasai, in that affected participants are Patients, and Patients are preferably coded as objects rather than as subjects. However, they do not provide any motivation for languages like Chickasaw, Choctaw, and Korean which, following the Relational Succession Law (Perlmutter and Postal 1983), can code EPs as subjects in certain clauses.

More to the point, we cannot ignore the fact that *both* Affected and Possessor are crucial semantic features for the EP construction, and prototype analyses do not help us understand why otherwise intransitive verbs can take grammatical objects when the object codes a Possessor-Patient (if that is what we want to call it) -- but not a simple Patient.

The explicitness of Dowty's Patient prototype allows us, in fact, to see that the EP is not a very good Patient at all. To take specific examples like either (14b) or

(16), the EP does not undergo a change of state, it will not be incrementally affected in any clear fashion, it is not causally affected by another participant but rather is affected by the entire situation involving the child, and no Agentive participant is moving with reference to it. Thus, aside from the affectedness parameter, the EP is not a very good Patient at all. If anything, it is perhaps more akin to an "Experiencier" or "dative" (whether metaphorical Goal, Benefactive, or Malefactive) -- and this doubtless explains the tendency in Indo-European languages, at least, to use grammatical dative or indirect object forms to code EPs. But in Maasai there is an explicit dative applicative suffix (11), and this is not employed in the EP construction. Thus, there is fairly straightforward evidence that the semantics involved in Maasai are not that of Benefactive or Goal which are elsewhere signaled by the dative applicative.

4.2. Locus of Affect as distinct from a semantic role. Like Gruber and Jackendoff (and presumably Dowty and Givón), Fillmore's classic and influential (1977) article, "The Case for Case Reopened," adopted a view in which verb properties determine what semantic roles the arguments bear. The motivation for Fillmore's system came from the conviction that there was a level of semantic roles different from "deep" grammatical relations; coupled with the desire to work towards a constrained view which avoided ad-hoc proliferation of roles. For Fillmore, the three core thematic roles are Patient (the thing which gets manipulated), Goal (the item on which the manipulated thing acts), and Agent (the manipulator).

In Fillmore's system, Affected is irrelevant to the role of Patient. Indeed, Affectedness and Patient are importantly distinct. Patient is a semantic role which cannot be lost when a verb is used in different syntactic constructions. But shifting conceptions of Affectedness can be conveyed by shifting to a construction in which the Affected item is brought into "perspective" and is expressed as the direct object:

> When an AGENT moves a PATIENT against a GOAL, and as a result the GOAL participant *moves* or *changes*, the element in the GOAL [semantic role] has acquired the saliency sufficient for it to be included in the perspective. (1977:76)

For Gruber and many others who followed him, the *item which moves* and *item which changes* are taken to be semantic features identifying Patient/Theme participants. But Fillmore does not see the additional semantic "colorings" of *moving* and *changing* as being either the addition of a second semantic role onto the Goal participant, or as replacing the original Goal. Rather, he links the *movement* and *change* feature directly to the grammatical relation of direct object, as seen in his discussion of the examples in (21).

(21) a. *I* *cut* *my foot* *on* *a rock.*
AG PAT GOAL

b. *I* *cut* *my foot* *with* *a rock.*
AG GOAL PAT

> In the sentence with *with* the foot has the goal relation to the action, and the rock is treated as the thing which acted against the foot; in the sentence with *on* the foot has the patient relation to the action, and the rock is seen as the thing against which the foot moved. *The thing which underwent the change of state -- in each case, the foot -- is expressed uniformly as the direct object, independently of its case role in the underlying action scene.* [Italics mine - DP] (Fillmore 1977:78)

And further:

> ... the relationship between a change-of-state verb and the entity which undergoes this change in state is reflected, not in the underlying [semantic role] case structure..., but in the grammatical relation DIRECT OBJECT.

Features that Fillmore identifies as affecting which items are likely to be "brought into perspective" (thus coded as objects and subjects) include humanness, movement, definiteness, and totality of affectedness Fillmore's list appears to entirely involve semantic features, with no mention of such things as pragmatic "discourse topicality" or cognitive "focus of attention," though it would not seem too far afield to infer that by "bringing something into perspective" Fillmore did have in mind a certain cognitive conceptualization of the situation.

Like Fillmore, Klaiman (1988) clearly distinguishes the Affected participant from the semantic role of Patient. Klaiman refers to the Locus of Affect as a "conceptual status" which may correspond to either Agent or Patient, depending on the particular basic voice of a clause. (Thus, she also differentiates Locus of Affect from Foley and Van Valin's (1984) Undergoer and Kemmer's (1993) Endpoint.[8]) Klaiman's motivation comes from trying to grasp the semantic typology of middle constructions across Sanskrit, Greek, Korean, and English. In middle constructions generally (cf. 5-6), "the 'action' or 'state' *affects* the subject of the verb or his interests" (Lyons 1968:363, quoted in Klaiman 1988). As seen in the discussion surrounding examples (5-6), Klaiman observes that in middles, a given semantic role like Agent/Source or Theme/Patient is necessarily also understood to have a distinct semantic feature of Affected (or sometimes of Benefit; see 5b):

> ... it is clear that diathesis [basic voice] in the IE system marks verbs according to whether their subjects have affected entity conceptual status, irrespective of their *possible concomitant status* as actors, sources, controllers, or catalysts of

> action... in classical IE, diathesis signals not the subject's thematic relation to the action, but its conceptual status as affected or nonaffected. [emphasis mine - DP]. (1988:37).

Relative to Maasai, if we are simply concerned to see how the Theta Criterion might still stand, Fillmore and Klaiman's approaches provide one way out, in that one could first say the only bona fide semantic role in the EP construction is that of Possessor. Second, the construction is akin to an applicative (though without the overt morphological marking) in bringing in an otherwise oblique-like role into the (derived) argument frame of the verb just when the Possessor is conceptualized as the primary Locus of Affect.

Croft (1992) suggests there is an idealized cognitive event model which contains a Controller Starting Point, and a Locus of Effect Endpoint. He discusses how causative, passive, antipassive, middle, and applicative morphosyntax express new construals of the primary Starting Point and primary Locus of Effect (or Endpoint), differently from those that are lexicalized as the default Starting and Ending points for given verbs. Whether one wishes to view such morphosyntax as manipulating "voice" versus "event construal" is not my primary concern here. I do suggest, however, that the Maasai EP construction belongs to the domain of morphosyntactic constructions that manipulate basic lexicalizations such that the speaker can express construal of the Possessor as the primary Locus of Effect in the situation described by the sentence -- rather than the Patient, Agent, or some other core argument of the lexical root.

To summarize, Fillmore most closely ties Affected with the grammatical relation of Object. Klaiman specifically avoids such a statement (as do Foley and Van Valin, and Kemmer), clearly keeping the affected participant in the "conceptual" (Klaiman, Kemmer) or "semantic" (Foley and Van Valin) domain. They all, however, separate Affected from the semantic roles of Patient, Recipient, Benefactive, Goal, and Agent. In contrast, Givón and Dowty tie "totally affected" to the prototype of Patient, as a semantic role.

5. Conclusions and further questions. The Maasai EP construction is felicitous only when the speaker can conceive of the event or situation as somehow being to the Possessor's Benefit or Detriment. Typically, the Possessor ends up being happy or sad, or somehow being emotionally affected, by the event or situation. That is, the possessor is seen as being the ultimate Locus of Effect of the situation. Similar to the examples in (3d) and (21), when the "oblique" Possessor is brought "into perspective" and coded as the direct object on the verb, it does not lose its Possessor semantic role and take on a Patient role. It retains the same role that it had as a non-core argument of the verb, governed by a noun; but as it is construed as the conceptual Endpoint, it is conceptualized as the primary Locus of Effect.

If one tries to explain syntax starting from a verbal argument-frame centered view, there seems to be no easy way to explain all the semantic features of the EP construction, regardless of whether one takes a prototype view of roles, or a more discrete view. This is first of all because the verbs in the construction do not have Possessor as one of their basic core arguments. A Bakerian approach might suggest incorporation of a zero possessor-applicative into the verb, such that the derived verb could then govern the Possessor and assign the Possessor semantic role. This approach, however, is silent about the Affected feature that is part-and-parcel of the EP construction.

If one starts from a construction centered view, the facts fall out more cleanly (cf. Goldberg 1995). It is the EP *construction* that licenses the Possessor role, a role which is distinct from Patient and from Goal/Benefactive; thus, a distinct Patient may still occur in the clause, as in (14b). As we have seen, the distinctness of the Possessor from Patient is shown by those lexically transitive roots that bring with them a basic Patient and which do not allow a second Patient to be added (cf. 14a). Its distinctness from Goal/Benefactive is shown by the failure of this construction to employ the dative applicative, and also by the distinctive Possessor semantics. Additionally, this construction is chosen precisely to convey an alternative construal of the event or situation lexically coded by the verb root, a construal in which the Possessor is conceived of as the final Locus of Affect.

Finally, these data raise questions for how *constructions* in the sense used here might arise – though I cannot provide any resolution here. If most constructions arise via generalization from the argument frame of paradigm verbs (e.g., the ditransitive CAUSE-RECEIVE <Agent Recipient Patient> arises via generalization of the *give* frame as in *Mary gave John a cake*; Adele Goldberg, p.c.), it is difficult to see how such a source could account for the development of the EP construction, as it is not clear that there is any basic verb with Possessor as part of its core argument frame. The closest verb type that does suggest itself as a possible candidate for the origin of the EP construction is a simple ditransitive verb, like *ɪshɔ 'give,'* which allows a third participant without a dative applicative; as we have seen, affected Possessors are not prototypical Patients by Dowty's criteria, and many languages do code EPs as the third argument of a ditransitive clause. In this regard, it is probably not irrelevant that all Maasai dialects allow the EP construction with transitive verbs (Table 1), but not all with intransitives.

NOTES

* This work could not have been pursued without the patient and insightful help of Leonard Kotikash, Philip Koitelel, Jonathan Ololoso, Renoi and Morompi Ole-Ronkei, and Nancy McGill. In addition, I am grateful to Christine Bartels, Scott DeLancey, Adele Goldberg, and Tom Payne for comments pertaining to this paper; and to Immanuel Barshi, Mitsuyo Hamaya, and Leonard Kotikash for the field experiment referred to in note 6.

1. The term *Theta Criterion* has been popularized by Government Binding Theory, but the idea is not exclusive to that theory.
2. In this paper I use Theme and Patient interchangeably.
3. Abbreviations are: ACC accusative, AG agent, ANTIPAS antipassive, DAT dative, DSCN discontinuous discourse thread, F feminine, IMPF imperfective, INV inverse, M masculine, NOM nominative, OBL oblique, PAT patient, PL plural, POSD possessed, POSR possessor, PROG progressive, SG singular, TH theme, 3>1 third (or any plural) subject with first singular object. The Maasai orthography used here generally follows the conventions of Tucker and Mpaayei (1955) with the exception that the 'strong' glide phonemes are written *yy* and *ww*, rather than *yi* and *wu*.
4. In this respect Maasai differs from languages which do employ applicatives in EP constructions; cf., Aissen (1987) on Tzotzil.
5. Shibatani (1994) offers an alternative explanation in terms of "relevance." He argues that Affectedness -- either along the lines outlined shortly for inalienable possession, or adversative/benefactive affectedness -- is a feature which makes something highly "relevant" to the scene. If this line of reasoning were to be taken to its logical conclusion, the more affected a participant is, the more easily integratable it should be. This, however, does not happen with the Maasai EP construction. As we will see in Section 4.1, the Maasai EP is not a prototypical Patient in Dowty's sense precisely because it is not totally affected.
6. In one preliminary field experiment speakers were asked to add nouns to sentences beginning with a variety of inflected verbs. The speakers could continue the sentence as either EP or non-EP sentences. The verbs presented included cognition/perception and physical contact meanings. In all cases where the sentence was continued as an EP construction, the verb was one of physical contact and the item added was a body part.
7. Even when on the body, object clothing items were not generally construable with the EP in the IlUasinkishu dialect. "Pockets" was a definite exception.
8. Foley and van Valin (1984) posit macro-(semantic) roles of Actor and Undergoer. Each of these roles may encompass a range of more specific roles – e.g., Undergoer may code Patient or Goal. Kemmer (1993:39, 50) develops Initiator and Endpoint "participant-structure" concepts, which sound rather like a mix of Givón's prototype notion of semantic roles, and Foley and van Valin's macro-semantic roles (both of which Kemmer cites as antecedents).

REFERENCES

Aissen, Judith. 1987. *Tzotzil Clause Structure.* Dordrecht. D. Reidel.

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

Barber, E. 1975. Voice -- beyond the passive. *Berkeley Linguistics Society* 1:16-24.

Croft, William. 1992. Voice: beyond control and affectedness. *Voice: Form and Function*, ed. by Barbara Fox and Paul Hopper, 89-117. Amsterdam: John Benjamins.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67.547-619.

Fillmore, Charles. 1977. The case for case reopened. *Syntax and Semantics 8, Grammatical Relations*, ed. by Peter Cole and Jerrold Sadock, 59-81. New York: Amsterdam.

Foley, William and Robert van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Givón, T. 1984. *Syntax: A Functional-typological Introduction, vol. 1*. Amsterdam: John Benjamins.

Goldberg, Adele. 1995. *A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Gruber, Jeffrey. 1976. *Lexical Structures in Syntax and Semantics*. Amsterdam: North Holland.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Kemmer, Suzanne. 1993. *The Middle Voice*. Amsterdam: John Benjamins.

Kimenyi, Alexandre. 1978. *A Relational Grammar of Kinyarwanda*. Berkeley: University of California Press.

Klaiman, Miriam. 1988. Affectedness and control: a typology of voice systems. *Passive and Voice*, ed. by Masayoshi Shibatani, 25-84. Amsterdam: John Benjamins.

Lyons, John. 1968. *An Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

Napoli, Donna Jo. 1993. *Syntax, Theory and Problems*. Oxford: Oxford University Press.

Payne, Doris. 1997. The Maasai external possessor construction. *Essays on Language Function and Language Type*, ed. by John Haiman, et al., 395-422 Amsterdam: John Benjamins.

Payne, Doris, Mitsuyo Hamaya and Peter Jacobs. 1994. Active, Passive and Inverse in Maasai. *Voice and Inversion*, ed. by T. Givón, 283-315. Amsterdam: John Benjamins.

Perlmutter, David and Paul Postal. 1983. The Relational Succession Law. *Studies in Relational Grammar 1*, ed. by David Perlmutter, 30-80. Chicago: The University of Chicago Press.

van Riemsdijk, Henk and Edwin Williams. 1986. *Introduction to the Theory of Grammar*. Cambridge, MA: MIT Press.

Rosch, Eleanor and Carolyn Mervis. 1975. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* 8.382-439.

Tucker, A. N. and J. Mpaayei. 1955. *A Grammar of Maasai with Vocabulary*. New York: Longmans, Green and Co.

Shibatani, Masayoshi. 1994. An integrational approach to possessor raising, ethical datives, and adversative passives. *Berkeley Linguistic Society*, 461-486.

Wise, Mary Ruth. 1971. *Identification of Participants in Discourse: A Study of Aspects of from and Meaning in Nomatsiguenga*. (Summer Institute of Linguistics Publications in Linguistics and Related Fields 28.) Norman, OK: Summer Institute of Linguistics and the University of Oklahoma.

From body to argumentation: grammaticalization as a fractal property of language
(the case of Wolof *ginnaaw*)

Stéphane ROBERT
CNRS-LLACAN (Paris) robert@cnrs-bellevue.fr

This paper uses the example of *ginnaaw*, a polysemous morpheme of Wolof, to demonstrate that grammaticalization can be understood as one of what may be called the FRACTAL PROPERTIES of language. In fact, *ginnaaw*'s synchronic uses across three syntactic categories (noun, preposition and subordinating conjunction) can be described as a common semantic structure applying at different levels inside the utterance, thanks to the syntactic flexibility of the term. This fractal model can thus account for the phenomenon of grammaticalization and, more generally, the transcategorial functioning of linguistic morphemes, by relating semantic variation (and argumentation) to syntax: the variation of the syntactic scope of the morpheme produces its polysemy. The term *ginnaaw* also reveals connections between body, space, causality and argumentation. These different domains have common topological properties which allow the same term to refer to all of them. *Ginnaaw* expresses a spatial framing in which discourse is shaped in a topological way as a landscape with orientations. The orientation defined by *ginnaaw*'s semantics explains the argumentative values of this morpheme[1].

Introduction. Synchronically *ginnaaw* occurs in three different grammatical categories. As a noun, it names a body part, 'back'; as a preposition it means 'behind', and in some extended uses 'after' or 'except'. *Ginnaaw* also has a more striking use as the subordinating conjunction 'since' in its causal meaning, much like French 'puisque' with its argumentative properties. Examples (1), (2) and (3) exemplify the different uses[2].

ginnaaw's senses:	
noun	- back (body part)
preposition	- behind (extended uses: after, except)
subordinating conjunction	- since (causal not temporal)

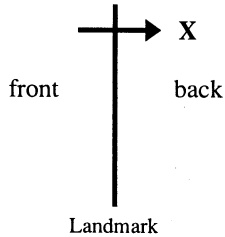
(1) *Jigéénu Senegaal dañuy boot seen doom ci* GINNAAW
woman+conn. Senegal EmphVb.3pl+uncompleted carry their children prep. *ginnaaw*
Senegalese women carry their children on their BACKS

(2) *Mi-ngi dëkk ci* GINNAAW *jàkka ji*
3sg-Presentative live prep. *ginnaaw* mosque the
He lives BEHIND the mosque

(3) GINNAAW *faral nga ko, maa-ngi dem.*
ginnaaw to.side.with Perfect.2sg him, 1sg-Presentative go
SINCE you have taken his side, I am leaving

Thus, we are dealing with the well-known phenomenon of grammaticalization. However, since we have here a morpheme which has three functions in the synchronic system, rather than a historical process, I prefer to use the term *transcategorial morpheme*. By 'transcategorial', I am focusing on the fact that *ginnaaw* functions in different syntactic categories. *Ginnaaw*'s polysemy then involves syntax, semantics and also argumentation. This synchronic functioning requires a semantic analysis which can both provide a unitary analysis of the meaning, and also account for the various senses of the morpheme.

The analysis presented here assumes that *ginnaaw* defines an asymmetrical space with a front / back orientation proceeding from a LANDMARK (or LOCATOR) and REFERS TO THE SPACE BEHIND IT (excluding the landmark). 'Landmark' can be understood in Langacker's sense (1987:10) with the additional qualification[3] made by Culioli in his use of *repère* (1978a, 1978b), often translated as 'locator'. A locator refers to any kind of entity (a notional or a temporal reference point, a physical landmark, a noun, a proposition, a speaker...), used in an utterance as a reference point to locate (and thus specify) another entity. For the purpose of this paper, 'landmark' and 'locator' are essentially equivalent.



ginnaaw
Diagram 1

With this definition, the different uses of *ginnaaw* can then be explained according to which element serves as the landmark. This element plays the role of the variable producing the observed polysemy. The syntactic scope of each different usage reveals the different levels in the sentence at which *ginnaaw* applies. In the various uses of *ginnaaw*, the same semantic structure applies at different syntactic levels. This property is what makes the functioning of a morpheme, a 'fractal' functioning. Thus *ginnaaw* refers to different 'domains' which are presented as structured spaces with their specific landmarks: body, spatial relations and discourse. I'll try to show, as well, that this semantic analysis can also account for the argumentative values of *ginnaaw* and its pragmatic effects.

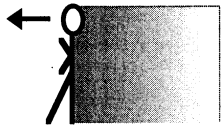
1. *Ginnaaw* as a noun and as a preposition. In its first use, *ginnaaw* functions as a noun. It can be used after a preposition, as shown in example (4), or take a possessive determiner, as shown in examples (5) and (6).

(4) *Jigéénu Senegaal dañuy boot seen doom ci GINNAAW*
woman+conn. Senegal EmphVb.3pl+uncompl. carry their children prep. *ginnaaw*
Senegalese women carry their children on their BACKS.

(5) *Xoolal GINNAAWam*
look+imper.2sg back+his
Look at his BACK / look behind him (ambiguous)

(6) *Xoolal ci sa GINNAAW*
look+imper.2sg prep. your back
Look in/to your BACK = the space behind you

Notice that syntactically, there is no noun after *ginnaaw*: *ginnaaw* has a nominal status (it receives the nominal determiners), thus a referential scope and a denotational value. Since there is no other element in the clause to play the part of the landmark, *ginnaaw* refers to the space behind the primary landmark, namely the human body. Examples (5) and (6) show that the space referred to is not only the body part 'back' but can be extended to the space associated with the body part, the space behind the person. It is worth noting that the human body has an intrinsic orientation and this orientation is relevant to *ginnaaw*; as depicted below.



By way of contrast, in example (7), *ginnaaw* governs a noun, *jàkka ji* - 'the mosque', in which case it behaves like a preposition with its syntactic scope and specific semantics. It introduces an argument and does not refer to the body part any more. The noun plays the role of the landmark and *ginnaaw* refers to the space behind this landmark.

(7) *Moodu, mi-ngi dëkk ci GINNAAW jàkka ji*
Moodu 3sg-Presentative live prep. *ginnaaw* mosque the
Moodu lives BEHIND the mosque

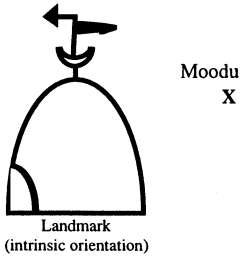


Diagram 2

Ginnaaw is used to locate an argument of the clause (namely 'Moodu', the subject) behind the landmark specified by the complement ('the mosque'). The use of body

parts as spatial prepositions is very common crosslinguistically[4]. Wolof has an entire system of body parts used in compound prepositions, as shown below. The compound prepositions are prepositional phrases made up of the only real preposition of Wolof (the locative *ci*), the body part noun and the connective suffix [u][5].

kanam	face	*ci kanamu X*	in front of X
biir	belly	*ci biiru X*	inside X
wet	side	*ci wetu X*	beside X
ginnaaw	back	*ci ginnaaw(u) X*	behind X
ndigg	waist	*ci diggante X ak Y*	between X and Y

As mentioned previously, *ginnaaw* only refers to a space behind a landmark. Thus, as a spatial preposition, it requires a locator (namely the Wolof locator *ci*) to relate the two arguments (the one located, i.e. the subject, and the one specifying the landmark of the space, i.e. the circumstantial complement) in the predicative relation. This constitutes a syntactic constraint related to the semantics of *ginnaaw*. *Ginnaaw* is then used in a complex prepositional phrase. By way of contrast, in its temporal uses[6], as in *ginnaaw ëllëg* - '(the day) after tomorrow', and its argumentative uses (see section 3), *ginnaaw* is related directly to the main verb (without *ci*) because it is not used to localize one argument in the space defined by another argument. Rather, it defines the 'space' in which the PREDICATE is validated, specified by the landmark.

In the case we have just examined, the landmark -- namely, the mosque -- had an intrinsic orientation. What happens in such a case when a landmark, such as a hill, has no intrinsic orientation?

(9) *Moodu, mi-ngi dëkk ci ginnaaw tund bi*
Moodu 3sg-Presentative live prep. *ginnaaw* hill the
Moodu lives behind the hill.

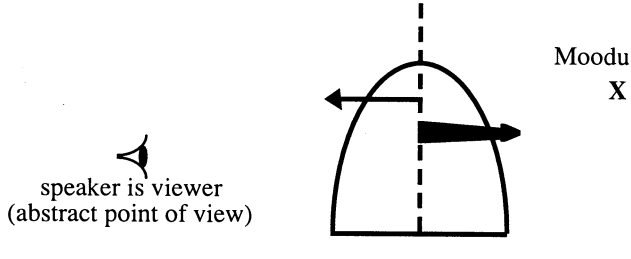


Diagram 3

The orientation is then given by the point of view of the speaker. The speaker is a viewer[7] and creates the orientation of the landmark as FACING him.

2. GINNAAW as a subordinating conjunction. Now, let us turn to the third use of *ginnaaw* as the subordinating conjunction 'since' in its causal sense.

(10) *GINNAAW faral nga ko, maa-ngi dem.*
ginnaaw to.side.with Perfect.2sg him, 1sg-Presentative go
SINCE you have taken his side, I am leaving.

(11) *GINNAAW añ nañu ba noppi, mën nañu naan àttaaya*
ginnaaw have.lunch Perfect.1pl until cease, can Perfect.1pl drink tea
SINCE we have finished lunch, (now) we can drink tea.

Here *ginnaaw* is used in a complex sentence and governs the subordinate clause, designated P in Diagram 4. The subordinate clause always precedes the main clause, designated Q. According to the general analysis proposed in Diagram (1), *ginnaaw* refers to the space behind a landmark. Here, however, the landmark is the clause P, in which case *ginnaaw* expresses a locational relationship between two clauses. 'Behind (i.e. given) the fact that you have taken his side (P), there is the fact that I'm leaving (Q)'.

Landmark Landmark

(*ginnaaw* refers to the shaded space)

Diagram 4

How does *ginnaaw* come to mean 'since' in its causal sense? The answer relies on understanding what a 'landmark' is in discourse. In this third use, the syntactic scope of *ginnaaw* is a clausal complement, not a noun. We are dealing with a complex sentence at the discourse level, i.e. a complex assertion. According to *ginnaaw*'s semantics, the clause P ('you have taken his side') is the landmark behind which the clause Q is located and *ginnaaw* REFERS to the space behind this landmark. Thus, the main clause ('I am leaving') is the scope of assertion, the focus, and the *ginnaaw*-clause is presented as the starting point of the utterance, a topic.

This point is confirmed by the syntactic constraints on the order of the clauses. As shown in examples (12a) and (12b), *ginnaaw*-clauses always appear first. Wolof word order does not parallel English word order in this type of sentence. The *ginnaaw*-clause can occur after the main clause only when the sentence is marked by a special cohesive anaphoric intonation which confirms its topical status. In contrast with another causal morpheme *ndax*, *ginnaaw* always appears in first position (compare examples 12 and 13).

(12) a. *GINNAAW mënuloo ànd ak man, maa-ngi la fiy bàyyi.*
ginnaaw can+Neg.2sg accompany with me, 1sg-Presentative you here+uncompleted leave
SINCE you can't come with me, I am leaving you here

(12) b. **maa-ngi la fiy bàyyi, GINNAAW mënuloo ànd ak man*
I am leaving you here, SINCE you can't come with me

(13) *Maa-ngi la fiy bàyyi, NDAX mënuloo ànd ak man.*
1sg-Presentat. you here+uncompl. leave, because can+Neg.2sg accompany with me
I am leaving you behind BECAUSE you can't come

The topical status of the *ginnaaw*-clause is also confirmed by the impossibility of an answer with *ginnaaw* to the question 'why are you leaving?'. As shown in (14), the normal answer to a why-question, is with *ndax*. An answer that begins with *ginnaaw* is understood to be an unfinished sentence (cf. 14c and d).

(14) *Lu tax ngay dem ?*
What to.cause Aorist.2sg+uncompl. go
Why are you leaving ?
(14) a. *NDAX mënuloo ànd ak man.*
Because you can't come with me.
(14) b. **GINNAAW mënuloo ànd ak man*
Since you can't come with me
(14) c. *GINNAAW mënuloo ànd ak man, KAAY...*
Since you can't come with me, THEN...
(14) d. *GINNAAW mënuloo ànd ak man, MAA NGI LA FIY BÀYYI.*
Since you can't come with me, I'LL GO BY MYSELF

Thus, the two markers express causality but with an opposite 'figure / ground' organization (Talmy 1978) at the discourse level. The *ginnaaw*-clause is not new information (figure); it is only the topic (ground) i.e. the starting point or reference point of the utterance.

Thus, in its interclausal use, *ginnaaw* does not express a temporal sequencing of the events P and Q in spatial terms (*behind = after P, there is Q). Rather, it expresses a relationship between two propositions in the *assertive space*, namely a localizing relation between a topic and a focus. This is confirmed by the fact that *ginnaaw* in its subordinating uses apparently never has the temporal meaning of 'after'[8]. Moreover, as shown in example (15), where the subordinate clause refers to a future event, *ginnaaw* does not imply temporal antecedence of the first proposition vis-à-vis the second.

(15) *Ginnaaw mu-ngi fay dem, jarul may bind*
ginnaaw 3sg-Presentativ. there+uncompl. go, worth+Neg.Perf.3sg Aor.1sg+uncomp. write
Since he is going there, there is no need for me to write

Here, the causality is presented as a spatial orientation between two propositions, the first one being the topic after which the second one can be asserted. The preceding statement ('he is going there') has created a spatial situation orienting toward a conclusion ('there is no need for me to write'), following 'behind' those premises. The consequences 'following' *ginnaaw* are not temporal, but argumentative. In example (15), the event referred to by the *ginnaaw*-clause is still

to come at the time of utterance but the speaker infers from this first statement consequences for the present situation.

Thus, at the utterance level, *ginnaaw* validates the main clause as a following consequence of the topic. The spatial relationship between the two clauses expresses both a sequencing in CAUSALITY and a sequencing IN THE SPEECH ACT. Two crucial points are involved here:

(a) Causality is conceived as a localization in a space comparable to the model of space built up from the orientation of the body, where the landmark-clause is the causal source ('causal landmark') behind which there is a following event.

(b) Argumentative inference is conceived of as an orientation in the 'assertive space'. The topic, or given information, is the starting point of the utterance[9]. It is the 'discursive landmark' from which the speaker's stance follows. The focus follows the topic as a consequence in the 'assertive space'.

The specificity of this causality appears in the contrast between *ndax* and *ginnaaw*. *Ndax* expresses an explanation, the causal clause is the focus. By contrast, the *ginnaaw* clause expresses the discursive landmark, from which the speaker's stance follows ('since you can't come with me, then I'll leave you here'). For this phenomenon, I use the term ARGUMENTATIVE CAUSALITY.

ndax	because	- causal clause is the FOCUS ('figure'), speaker's assertion. - explanation
ginnaaw	since	- causal clause is the TOPIC ('ground'), hearer's assertion[10]. - causal clause is the LANDMARK OF ASSERTION - argumentative causality ('I'm not responsible'): comment 'follows' from the topic

3. Discourse as a landscape: topology of argumentation. With *ginnaaw*, discourse is presented as a landscape where some propositions are landmarks defining spatial ordering, orientations and paths between propositions. Argumentative inference is conceived as a path leading from one statement to another. Actually, in order to describe the clause Q as located in a space behind P, you need a viewpoint. Since we are at the utterance level, we can assume that the speaker is the viewer of this landscape -- a conceptualizer with a point of view, vantage point and orientation in the abstract space of discourse. A schematic representation of the speaker's point of view is given in Diagram 5.

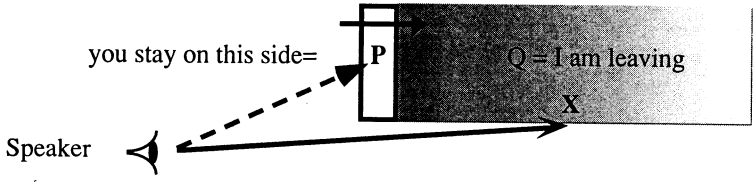


Diagram 5

The use of *ginnaaw*[11] in the sense of 'besides', 'except', confirms this analysis of assertion in terms of abstract spaces and abstract point of view, as shown in example (18).

(18) *GINNAAW Moodu, ñépp ñëw nañu*
ginnaaw Moodu, all come Perf.3pl
BESIDES (except) Moodu, they all came.

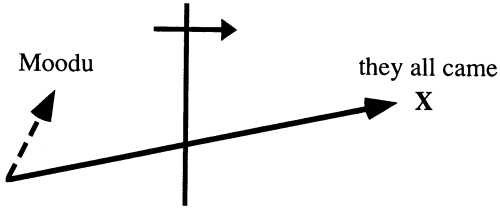


Diagram 6

According to the general analysis, *ginnaaw* creates an asymmetrical space and REFERS to (thus validates) the 'space' behind the landmark, excluding the landmark. In other words, the proposition 'they all came' IS TRUE only behind the landmark 'Moodu'. Hence the sense 'except Moodu, they all came'.

In this spatial framing of the discourse, when *ginnaaw* relates two clauses, argumentation is laid out in such a way that from the point of view of the conceptualizer, the comment is considered as proceeding from the topic. In order to be a topic, i.e. a DISCURSIVE LANDMARK, the *ginnaaw*-clause has to be a stable reference point, therefore a clause presented as EPISTEMICALLY GROUNDED so that the speaker can use it as a starting point for AN ARGUMENTATIVE SEQUENCE. Thus, with *ginnaaw*, the causal clause is presented by the speaker as part of the common ground of the discourse, a previous statement, independently established (whether discursively true or not) and independent of the current speech act. The current assertion Q follows from the orientation created by the previous one P, therefore the speaker is not responsible for the consequences following this first statement. The various argumentative effects proceed from this epistemic status of the *ginnaaw*-clause.

Speaker and addressee may have different positions with respect to the focus of the *ginnaaw*-utterance. If the addressee has the same position as the speaker with respect to Q, as in example (11) where Q is good or neutral, the argumentative

effect is a confirmation. If the addressee has a different position with respect to Q, as in example (10) where Q is bad, the argumentative effect is what I have called the 'return to sender' effect (Robert 1990). That effect says if you are not happy with this, go back to the person that created the first situation ('you have sided with him'), from which my assertion is only a consequence ('I am leaving').

Here we have two metaphors, in Lakoff's sense (1993:207). Causality is conceptualized in spatial terms, one event behind another. In addition, with *ginnaaw* another metaphor is also at work which we can call the metaphor of discourse as landscape -- i.e. the structural properties of space (source domain) are mapped onto discourse (target domain). Argumentative sequences are constructed in terms of spatial relations (localization, orientation and paths proceeding from landmarks) between propositions. One statement is located behind another thanks to the orientation created by the discursive landmarks. In other words, discourse is framed as what I call 'an assertive space'. In that space, topics and comments follow each other, creating argumentative inferences. This spatialization of discourse also occurs in argumentative morphemes of other languages, as shown below, with English *way* and French *ailleurs*.

Engl.	*anyway*		
Fr.	*par ailleurs*	(lit. 'through elsewhere')	= Engl. *besides, moreover*
Fr.	*d'ailleurs*	(lit. 'from elsewhere').	= Engl. *on the other hand*

4. Grammaticalization as a fractal property of language. Through the various uses of *ginnaaw*, we see the same image-schematic structure functioning at different levels inside the utterance, as given in Diagram 7. The context specifies the level at which this semantic structure functions by defining the syntactic scope of the item and the nature of the landmark. When *ginnaaw* is in nominal function, no other term in the utterance plays the role of the landmark; the morpheme has an extra-linguistic reference; the landmark is the primary landmark -- i.e. the human body. In prepositional use, the landmark is the noun governed by *ginnaaw*. In subordinating use, the landmark is the clause introduced by *ginnaaw*.

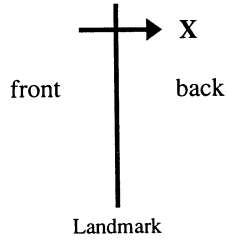


Diagram 7

Landmark = Ø = the body	*ginnaaw* = noun	sense = 'the back'
Landmark = a noun	*ginnaaw* = preposition	sense = 'behind'
Landmark = a clause	*ginnaaw* = sub. conj.	sense = 'since'

Grammaticalization and, more generally, the transcategorial functioning of morphemes such as *ginnaaw*, reveal a FRACTAL property of language. Indeed, objects are said to be 'fractals' (Mandelbrot 1975) when they have the property of SCALE INVARIANCE and SELF-SIMILARITY (Sapoval, 1997:73, 136): a similar structure appears at different scales. Those objects are invariant when undergoing a dilatation. A coast for instance is a fractal object (see Gleick, 1991:128).

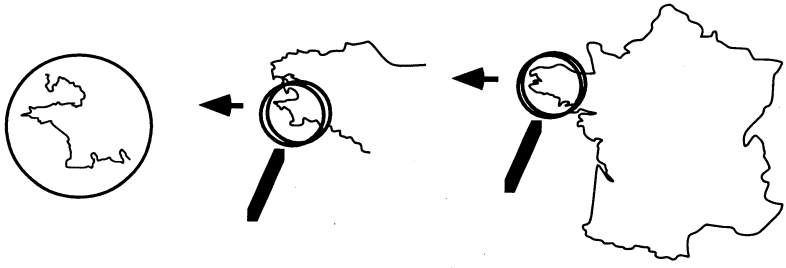


Diagram 8

Thus, fractals have SCALING LAWS (Sapoval 97) by which a common structure appears at different scales. But each scale also has specific scale properties so that there is no strict identity between the same structure appearing at different levels. Rather, we have an 'analogic' structure.

In the same way, we can say that with *ginnaaw* a similar semantic structure applies at different 'scales' inside the utterance. The linguistic 'scale' is the syntactic 'level' at which the unit functions. In language, the different syntactic levels are embedded in each other -- nominal level, prepositional phrase level, clausal linkage and discourse level. Through its transcategorial functioning, the semantics of *ginnaaw* undergoes dilatation of its syntactic scope with scale invariance and scale properties.

Attributing SCALE PROPERTIES to *ginnaaw* means that each level of the utterance (defined by the syntactic scope of *ginnaaw*) has specific properties despite a common semantic structure. Those properties are defined by the syntax and create the polysemy. At the nominal level, *ginnaaw* has referential scope and a denotational value. At the prepositional phrase's level, *ginnaaw* builds a relation between two arguments. At the utterance's level, we have the point of view of the speaker who endorses the utterance; a topic and a focus; a framing of the clauses with topological relationships. At this level, we also have argumentative effects produced by the topical status of the landmark organizing the framing of one proposition by another. The different paradigmatic oppositions of the various senses also appear as scale properties that specify the semantics of the term in each of its various uses. These are summarized in the chart below.

Scale properties of *ginnaaw* (with specific paradigmatic oppositions at each scale)	
• nominal scale:	- referential scope and denotational value
• prepositional scale:	- relation between two arguments
• utterance scale:	- speaker's point of view and assertion
	- topic and focus
	- framing of the clauses (topological relationships)
	- argumentative effects

Thus, the analysis of grammaticalization in terms of a topologically structured image schema, abstracted and preserved from one domain to another (Sweetser 1988), allows an account of the SEMANTIC INVARIANCE of a transcategorial morpheme and motivates the grammaticalization. Moreover, the fractal model proposed here specifies the nature of the various DOMAINS involved in transcategorial marker uses and accounts for the gain and loss of the meaning in the different uses by relating the SEMANTIC VARIATION to the change of syntactic scope in a functional manner.

Conclusion. The morpheme *ginnaaw* reveals connections between body, space, causality and argumentation. These different domains have common topological properties which allow the same term to refer to all of them. With *ginnaaw*, causality (argumentative causality) is conceptualized as a localization in a time-space comparable to the model of space built up from the orientation of the body. The foregoing space corresponds to the causal antecedence and the previous statement corresponds to the source of discursive inference. Argumentation thus appears to be also describable as an orientation in the assertive space. The analysis of grammaticalization in terms of fractal functioning relates syntax, semantics and argumentation in the dynamic process by which the meaning of a term is constructed inside an utterance.

Notes

1. I am grateful to Kevin Moore and Miriam Petruck for their helpful comments on this paper.

2. The official orthography of Senegal is used here.

3. In Culioli's conception, the referential value of a word and the meaning of an utterance are not given but yielded by a series of 'locating' operations at work inside the utterance: by relating a located term to an anchoring point (the 'locator'), the locating operation produces a new specification for this located term. This basic operation applies at different levels : the notional level (lexicon); the predicative level; the higher level of the speech act. At the predicative level, a predicative relationship is constructed. At the level of speech act, the predicative relationship is associated with a speaker and a time-place, and also with a previous verbal context defining the topic.

4. Locative noun phrases with an animate in genitive function (*ci ginnaaw Faatu* lit. "on Faatu('s) back") are ambiguous, referring either to the body part ("on Faatu's back") or to the space behind the landmark ("behind Faatu"). They represent an intermediate stage leading from nominal status (body part) to the use in a prepositional phrase with inanimate complement (*ci ginnaaw*

jàkka ji "behind the mosque"), in which case the syntactic properties of *ginnaaw* are different and the reference to the body part is absent.

5. The suffix [u] appears throughout the whole paradigm of these compound prepositions. The apparent exception in the case of *ci ginnaaw X* might be explained by phonetic reasons -- i.e. assimilation between the suffix [u] and the final bilabial glide [w].

6. The temporal sense of *ginnaaw* is possible in its prepositional use but non central. In that case, the temporal domain is shaped as a space. This temporal use however seems to be impossible when *ginnaaw* is used as a subordinating conjunction (see section 2.).

7. In the sense of an abstract point of view, since the speaker, who still functions as a reference point with its spatial orientation, might actually not see the hill.

8. "After" is expressed with another morpheme (*bi/ba*): *Bi mu lekkee la dem* (when Aor.3sg eat+anterior. EmphComp+3sg go) 'After he had eaten, he left'.

9. The 'constitutive locator' in Culioli's terms (1990 : 138-9).

10. Or a previous speaker's assertion. See section 3.

11. Thanks to Kevin Moore for calling my attention to this use.

References

Culioli, Antoine. 1978a. Valeurs modales et opérations énonciatives. Reprinted in Culioli. 1990. 35-155.

Culioli, Antoine. 1978b. Valeurs aspectuelles et opérations énonciatives: la notion d'aoristique. Reprinted in 1990. *Enonciation: aspect et détermination*, ed. by S. Fisher *et al.*, 99-114. Paris: EHESS.

Culioli, Antoine. 1990. *Pour une linguistique de l'énonciation. Opérations et représentations*, vol. 1. Paris: Ophrys.

Culioli, Antoine. 1995. *Cognition and Representation in Linguistic Theory*, texts selected, edited and introduced by Michael Liddle. Amsterdam: John Benjamins.

Fillmore, Charles. 1982. Frame Semantics. *Linguistics in the Morning Calm*. Seoul: Hanshin. 222-254.

Gleick, James. 1987. French translation, 1991. *La théorie du chaos*. Paris: Flammarion.

Lakoff, George. 1993. The Contemporary Theory of Metaphor. *Metaphor and Thought*, ed. by Andrew Ortony, 119-154. Cambridge: Cambridge University Press. [second edition]

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, vol. 1. Stanford: Stanford University Press.

Mandelbrot, Benoît. 1975. *Les Objets fractals, forme, hasard et dimension*. Paris: Flammarion.

Robert, Stéphane. 1990. *Puisque* et le dos en wolof, tempête dans le signifié. *Le Gré des langues* 1. 82-92.

Sapoval, Bernard. 1997. *Universalité et fractales*. Paris: Flammarion.

Sweetser, Eve. 1988. Grammaticalization and semantic bleaching. *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. 389-409.

Talmy, Leonard. 1978. Figure and Ground in Complex Sentences. *Universals of Human Language*, vol. 4. ed. by J. Greenberg *et al.* Stanford: Stanford University Press. 119-154.

Grammar and pragmatics in the Swahili auxiliary focus system

Benji Wald
University of California, Los Angeles

0. Linguistic problems of tense-aspect-mood are well-known for challenging lines drawn between grammar and pragmatics, or either and semantics; the speaker seems to have much freedom of choice in using the grammar to organize events into coherent discourse. At first thought, it would seem that at least time relationships among events would have a constraining effect on speakers' freedom of choice. However, it turns out that very often in languages the importance of the time relationships between events is overshadowed by the purposes for which speakers choose to mention such events. From this evolves the problem of whether in a given language a time relationship should be recognized as an element of the grammar, and its relevance to the larger discourse unit (story or whatever) considered a matter of pragmatic inference on the part of a listener, or whether, conversely, some other relationship between events is encoded in the grammar, from which a time relationship can be pragmatically inferred.

My point will be that a particular distinction within the Swahili tense-aspect system, which has previously been taken to be the problem of a grammatical or semantic distinction within tense-aspect, is something else. It is a declining distinction within an auxiliary focus system which happens to be embedded in the Swahili tense-aspect system, as is the case in many Bantu languages. I will further suggest that if grammaticalization is the process by which a set of strategies which start off being pragmatic may evolve into grammatical patterns, meaning that they somehow become obligatory in describable contexts, then there seems to be a reverse phenomenon. That is, a distinction which was once grammatically obligatory at an earlier stage of Swahili or its ancestor now seems to have reverted to a set of pragmatic strategies, much to the confusion of scholars who, over the past century, have examined Swahili, particularly its standard and Southern varieties. This process might be called "de-grammaticalization".

1. Example (1) below illustrates the problem. It is taken from a Swahili folktale recorded by Steere (1870). The variety of Swahili is a pre-standard urban Zanzibari variety, of the type which later came to be the foundation for the standard language reflected in the Swahili literature usually examined for linguistic analysis.

(1) hata watu **wa-<u>me</u>-zi-jua** nyimbo zile kidogo kidogo, lakini yeye, na mamaye, na mtumwa wake **wa-<u>na</u>-zi-jua** sana. Na maana yake zile nyimbo mamaye **<u>a</u>-zi-jua**, na wale watu mjini hawajui maana yake
"so then the people <u>were getting to know</u> (**me**) those songs little by little, but he, and his mother, and his servant <u>knew</u> (**na**) them a lot. And his mother <u>knew</u> (**a**) the meanings of the songs, (while) the people of the city didn't know their meaning." (Steere 1870: 442)

In (1) the same verb occurs three times in succession, each time with a different tense marker (henceforth TM). In (2) below, the TMs of interest are isolated and

described.

(2) Structure: Subject.Marker -**TM** (Tense.Marker) - Object.Marker - Verb Root...

a. tu-**me**-zi-jua (1p-**TM=me**-cl.10-know) 'we are / were getting to know them / we have / had come to know them'

b. tu-**na**-zi-jua (1p-**TM=na**-cl.10-know) 'we know / knew /are / were gonna know them'

c. tw-**a**-zi-jua (1p-**TM=a**-cl.10-know) 'we know / knew /are / were gonna know them'

Example (2) shows the syntactic position of the TM in Swahili, and in Bantu in general. With regard to meaning, the translation possibilities out of further context show that there is no exact correspondence with any single English device. The TMs all take their time orientation from some time reference point which depends on a larger context. That's why they may be used in either past or non-past contexts. **me** in (2)a is generally translatable distinctly from the other two, and will not be considered further here.[1] The problem is in distinguishing **na**, as in (2)b, from **a**, as in (2)c. Both fit into the general concept of "imperfective" aspect markers, to use traditional terminology (cf. Comrie 1976). The only thing I will say here about that is that this includes both progressive and non-progressive uses. Among non-progressive uses are Vendlerian state uses such as habitual, dispositional, gnomic and such things (Vendler 1967). It is such a state use that is exemplified in the passage in (1), for both **a** and **na**. Beyond that, the difference between them has never been adequately analyzed or even described in the literature. It is impossible to figure out from example (1), or any other set of examples.

One final thing needs to be said about the difference between **na** and **a**, which is at once obvious and turns out to be important in an unrecognized way; **na** is phonologically more substantial than **a**. It allows the TM an entire syllable on its own. **a** shares a syllable with the preceding obligatory Subject Marker. Many Swahili speakers actually call **a** the "short" form of **na**. This suggests that there is a pragmatic rather than semantic basis for the choice between them. And, in fact, **a** is more common in colloquial speech than in more formal social contexts of speech, in those dialects that use it at all.

Meanwhile, analysts of standard Swahili have insisted that **a** and **na** differ in meaning. Part of the problem is the concept of "standard Swahili" itself. All I can do here is make the observation that the written varieties of Swahili which are considered standard vary greatly in how much they use the TM **a**. All use the TM **na** much more often. Some do not use the TM **a** at all. In this they agree with certain rural dialects of Zanzibar, for which spoken texts show no use of **a**, amidst copious use of **na** (cf. Whiteley 1959). But there is no direct connection between these rural dialects, largely unknown outside of their small communities of speakers, and the standard writers, who are often second language speakers of Swahili from areas of Tanzania distant from Zanzibar. This suggests independent trends which indicate that something principled from the earlier state is continuing to evolve along similar paths in various dialects.

2. One of the most interesting and novel analyses of the standard Swahili tense-aspect system is Contini-Morava (1989). Using a variety of standard Swahili

written texts, some of which did provide a number of examples of **a** as well as **na**, she suggested the following distinction:

> "The difference between **na** and **a** is that **na** singles out a point in time to which an event is related, and indicates a type of [BW: temporal] ordering relationship, whereas **a** does not explicitly relate an event to any point in time (although it can be inferred to have a time-reference, given the appropriate context)." (Contini-Morava 1989: 67-8)

Elsewhere (p.93), she states that the semantic features of **a** are "not-negated" (i.e., cannot co-occur with negative markers) and "time-relevant" (i.e., there is some relevant time orientation), and that **na** has the same semantic features plus the additional feature "INCLUDES time orientation". Thus, "INCLUDES time orientation" is the specific "type of ordering relationship" referred to in the above quote.

It will not be necessary to give a full account of the motivations underlying Contini-Morava's analysis of the intended semantic distinction between **a** and **na**. Basically, it is simply that **na** contains more information than **a**, and more specifically that **na** refers to the fact that some phase of the event it marks INCLUDES a pragmatically determined time reference point. Such inclusion accounts for why **na** may be used in either progressive or Vendlerian state contexts. That is, she proposes that progressive contexts include a single point in time as the time orientation, while Vendlerian state contexts include any or all points in a relevant stretch of time (durative) or a series of recurrent points (iterative) as the time orientation. Thus, for example in (1), the time reference point can be pragmatically inferred to be any or all points in the past reached in the development of the story by the time of the passage in (1), perhaps including, for example, the last previously mentioned perfectively marked verb (i.e., the last preceding verb marked with the TM **ka**). Since **a** has an even more general meaning, according to Contini-Morava, it can also be used in the same contexts. Thus, there is no pragmatic difference in the interpretation of **a** and **na** in example (1).

Whatever our own theoretical inclinations, I think that we can agree, with regard to this or any similar analysis, that, according to classic linguistic analysis, if **a** has a more general meaning than **na**, then there should be some contexts in which **a** can occur but **na** cannot. However, Contini-Morava (1989: 68ff, esp. 78) admits that she cannot identify them. Instead she tries to show by statistical methods that **na** is more favored than **a** in contexts where the time relationship is worthy of greater precision. In this she comes close to an older and less accurate analysis by showing that **na** is particularly favored in progressive contexts. It is less favored, though still more frequent than **a**, in Vendlerian "state" contexts. In (3) I have reproduced as Table 1 her table which shows this.

(3)	State	Activity	Total	(n) = number of occurrences
a	58% (76)	42% (55)	100% (131)	
na	36.8% (156)	63.2% (268)	100% (424)	

Data: HLB, HBS, AB, PP $p < .001$

Table 1. Distribution of **a** and **na** with respects to State [BW: Vendlerian] vs. Activity [BW: generally translatable as Progressive]
(from Contini-Morava 1989: 71)

The logic may be that in progressive contexts there is acknowledgment that the event marked as progressive is "temporary" and will end sooner or later. Habitual and other state contexts do not make an issue out of coming to an end. In terms of acknowledging that there will be a final phase to the event, such states might be said to contain less information than progressives. Or it might simply be said, to paraphrase Contini-Morava, that temporary activity generally has a perceptual salience that is worthy of the more precise specification.

Note that Table 1 shows that **a** exhibits a marked preference for states (or at least certain states) and **na** for progressive contexts (or at least certain progressive contexts). However, also notice that in either progressive or non-progressive contexts **na** outnumbers **a** in raw frequency of occurrence. This would suggest that even though progressive contexts are more often worthy of greater specification than states, states themselves are, more often than not, also worthy of the greater specification.

One attractive feature of Contini-Morava's analysis is that it allows **a** and **na** to both be used in progressive and state contexts, since that is indeed the case for standard and many Southern varieties of Swahili. And yet it helps explain the impression of many earlier observers for such varieties that **na** is more closely associated with progressive contexts than **a**.

3. When I was able to interpret Contini-Morava's analysis in the preceding way, I recognized that the difference between **a** and **na** is a matter of the amount of focus on the time relationship, NOT one of different degrees of specificity of the time relationship. The favorability of **na** to progressive contexts remains for the same reasons that Contini-Morava suggested. I also recognized that the focus difference was a consequence of Swahili's evolution from an East Bantu type of auxiliary focus system similar to some of those described by Hyman & Watters (1984). Such auxiliary focus systems pair certain TMs (but not all), in order to indicate the location or scope of the constituent with maximal focus in the clause. I reasoned that in the course of time, the difference between **a** and **na** had narrowed in its domain from differences in constituent focus to different degrees of focus on the time relationship that **a** and **na** were originally paired to have in common, and still maintain. In fact, the narrowing is not toward specializing **na** as a marker of the progressive, but simply a stage in the eventual narrowing of the difference to zero, with the imminent loss of the **a** TM.

I will now consider the earlier focus system from which **a** and **na** have evolved. One of the simplest and well-known examples of such a system, not mentioned specifically by Hyman & Watters, is the Zulu system, typical of Southeast Bantu. The system involves devices for assigning the maximal focus of a clause to one or

another constituent of that clause, and/or varying the scope of maximal focus to include the verb as well as postverbal material. The actual uses of such focus are various, such as highlighting new information or casting certain constituents into contrastive or assertive focus in the larger discourse context. Which specific use of maximal focus is intended in a particular context is a matter of pragmatic inference. In contrast, the focus system itself is grammatical, as some of the constraints on the Zulu system show.

The point of departure is a grammar in which post-verbal position has been grammaticalized as having higher focus than pre-verbal position or the verb itself. Most comparable to the reconstructable distinction between Swahili **a** and **na** are the two Zulu TMs **0** and **ya**. The examples in (4) illustrate:

(4) ZULU: Southeast Coast Bantu / Nguni group. (Doke 1968: 334-41)
Post-V Focus.
a. ngi-**0**-bona abantu (I-**TM=0**-see people) 'I see (the) people'
V included in Maximal Focus.
b. ngi-**ya** /***0** (-ba)-bona (I -**TM=ya**-(OM=them)-see) 'I (DO) see (them)'
c. ngi-**ya** / **0** -ba-bona abantu 'I DO see *or* am seeing / habitually see (the) people' (cf. Doke 1968: 339)

In (4)a, there is actually no TM at all. The information status of the post-verbal constituent is flexible, but it is considered to be more in focus than the verb. Such focus is suitable to post-verbal position as the site for introducing either new or contrastive information. Most telling is the syntactic prohibition against the **0** TM when the verb is final, as in (4)b. Since focus grammatically increases toward the end of the clause, if the verb is at the end of the clause it has maximal focus, and therefore it must be marked by **ya** since it must be included in the scope of maximal focus.

With regard to (4)c, Doke (1968: 339) observed that **ya** is more conducive to a progressive interpretation than **0** in some contexts. About this he wrote:

> "In some cases when the same adjunct is used with either tense [TM], the former [**0**] has the idea of habitual action, the latter [**ya**] of continuous action..." (Doke 1968: 167)

However, he did not identify those contexts other than to provide examples which need not have such progressive interpretations. The examples he actually offers are:

(5)a. ba-**0**-yi-dumisa inyoka (they-**TM=0**-cl.9-worship snake)
'they conduct snake-worship'
b. ba-**ya**-yi-dumisa inyoka (they-**TM=ya**-cl.9-...)
'they are worshipping the snake' Doke (1968: 339)

It is evident that, for the same or similar reasons to those discussed above in connection with Table 1 in (3) for Swahili, the distinction given for (5)a-b is a pragmatic effect salient to speakers when they are asked to distinguish out of further context this minimal pair differing only in the TM used. Elsewhere Doke (p.167) notes "The two tenses ... are not really distinct in meaning or significance". In appropriate contexts, either (5)a or (5)b can be either progressive or habitual (i.e.,

Vendlerian state). Thus, (5)a can also be used for progressive "they are worshipping the SNAKE", with maximal focus restricted to the postverbal constituent, and (5)b can also be used for habitual "they DO (so) worship the snake", with the verb included in maximal focus. In sum, the effect in (5)a-b involves the higher focus of **ya** than **0**, and the greater pragmatic favorability of progressive than state contexts to such higher focus.

Note also that the greater substance of the TM **ya** than **0** also contributes iconically to retracting the scope of focus from strictly postverbal material. Recall the similar iconic relation between **na** and **a** in Swahili.

3.1. The system of associating certain TMs to vary maximal constituent focus within the clause is a feature of many East Bantu languages, but there are interesting differences in detail among them. Closest to Swahili among the better studied auxiliary focus systems is the Shambaa system, described as such by Odden (1982), and less explicitly in various articles by Ruth Besha (e.g., 1989). The system is more complex than Zulu's. It has three degrees of focus for some tense-aspect contexts. Example (6) illustrates:

(6) SHAMBAA: Northeast "Coast" Bantu / Seuta group.
Post-V Focus.
a. ni-**0**-dika manga (I-**TM=0**-cook cassava)
'I'm cooking CASSAVA (NOT something else)'
Neutral Focus.
b. n-**àà**-dika (manga) (I-**TM=àà**-cook ...) 'I'm cooking (cassava)'
V included in Maximal Focus.
c. ni-**ta**-dika (manga) (I-**TM=ta**-cook...)
'I'm COOKing (NOT EATing) (cassava) / I am SO cooking (cassava) / etc.'

Typologically, we can see the Shambaa system as further elaboration of the auxiliary focus system on a Zulu-like base. As in Zulu, the **0** TM requires a postverbal constituent to take the maximal focus. But unlike Zulu, there is a further contrast between a neutral focus marker and a marker which includes the verb in the maximal focus. The neutral focus marker does not require a postverbal constituent, but if there is one it does not have the heightened focus of its counterpart following the **0** TM. Finally, the marker in (6)c explicitly includes the verb in the maximal focus of the clause. The alternative translations of (6)c show that pragmatics is involved in deciding just what the scope of focus is. It can be just the verb, or it can be the entire predicate. Also notice that iconically, the phonological substance of the three markers reflect the degree of retraction of the scope of focus from postverbal position: 0 < vowel < syllable.

Swahili and Shambaa are closely enough related to see the cognate relationship between Shambaa **àà** and Swahili **a**. It can be further suggested that the Shambaa focus relationship between **àà** and **ta** is similar to the one which once obtained between Swahili **a** and **na** respectively. The difference is that the Swahili focus relationship no longer refers to the scope of maximal focus in the clause, but more narrowly refers to the amount of focus intended for a particular time relationship between the event marked by the TM and a time reference point. Beyond that, pragmatics must be used to decide whether the focus refers to a progressive time relationship or some other kind.

4. Next, it is important to note that Swahili also has a **0** cognate for Shambaa **0**. To my knowledge, the Swahili **0** TM has never before been viewed in this way. Just as with **na** and **a**, the Swahili **0** TM is more restricted in its focus domain than its Shambaa counterpart. It is restricted to a certain syntactic type of relative clause, as exemplified in (7)a below.

(7)a. Post-V Focus. (**0**-Rel) SWAHILI
u-**0**-taka-**cho** (2s-**TM=0**-want-**Rel.M**) 'what you want / need'

Significantly, in this type of relative clause there is an obligatory postverbal element, the relative marker (Rel.M), which refers to the head of the relative clause, whether expressed or understood, as in the example. Thus, we can see that the head of the clause has more focus than the verb, and is, in fact, the maximal focus of the clause, necessarily excluding the verb. The post-verbal relative marker represents the superior focus of the head.

Another, and perhaps less language-specific way to look at it, is that relative clauses are inherently lower in focus than main clauses, where relative focus takes on a significance to clause-size constituents. From this perspective, the domain of the Swahili **0** TM has narrowed to the point that it is too low in focus to dominate a main clause.

It is worth mentioning that the particular syntactic formation for relative clauses exemplified in (7)a is restricted to a relatively small area of Northeast Bantu which includes all the closest relatives of Swahili and extends to Shambaa, e.g.,

(7)a'. Post-V Focus. (**0**-Rel) SHAMBAA (Roehl 1911: 158)
ndima ni-**0**-kunda-**yo** (work 1s-**TM=0**-want-**Rel.M**)
'the work I want'

Thus, Shambaa also has this context for the **0** TM, though it has a greater range of contexts, as seen in (6)a. The generalization remains the same in both languages: the verb is not included in the maximal focus of the clause.

Further removed is the Southeast Bantu relative construction with the **0** TM and additional (tonal) marking of the subordinate status of the relative clause, e.g.,

(7)a". Post-V Focus. (**0**-Rel) ZULU (Doke 1968: 322)
umuntu e-ngi-**0**-m-bona(-**yo**) manje (person Rel.M-1s-**TM=0**-3s-
'the person I('m) see(ing) now' see(-**Rel.M**) now)

The most relevant difference between the Southeast construction of (7)a" and the Swahili-Shambaa area construction of the Northeast Coast is that in Zulu the postverbal relative marker is optional if there is another postverbal constituent in the relative clause. In the Swahili-Shambaa area the postverbal relative marker with the **0** TM is obligatory in all contexts, e.g.,

(7)a'". Post-V Focus. (**0**-Rel) SWAHILI
mtu ni-**0**-mw-ona-**ye** sasa (person 1s-**TM=0**-3s-see-**Rel.M** now)
'the person I('m) see(ing) now'

At the same time, all these languages agree that there must be a postverbal relative marker to take the maximal focus if there is no other postverbal constituent in the relative clause.

It should be noted that though a relative clause may be inherently lower in focus than a main clause, there are other Swahili relative clause strategies which allow its verb to be marked by the same TMs as in main clauses. (7)b and (c) illustrate them.

(7)b. Neutral Focus: Restricted TM. (**na**-Rel) SWAHILI
u-**na**-**cho**-taka (2s-**TM=na**-**Rel.M**-want)
'what / which you want / need'
Neutral Focus: Unrestricted TM. (amba-Rel)
c. amba-**cho** w-**a**-taka (COMP-**Rel.M** 2s-**TM=a**-want)
'what / which you want / need'

By either device the relative marker precedes rather than follows the verb, and the verb may be final in the clause, as in the examples. The construction exemplified in (7)b is restricted to certain TMs, which vary slightly according to dialect. The construction in (7)c is common to all dialects and allows the full range of main clause TMs. The lack of restrictions on the *amba*-type relative clause of (7)c is most suitable to non-restrictive relative clauses. The non-restrictive relative contains new information, so that the focus difference between it and the matrix clause is minimal, while restrictives contain given or head-defining information. Nevertheless, it is obvious that the *amba*-type of (7)c may also be used for a restrictive relative clause, since there are some TMs which can only occur in this type of relative clause. In such cases, it might be said that the focus value of those particular TMs has been grammaticalized as so high that they have become incompatible with the other forms of relativization. Again, whether a relative clause is restrictive or not is a matter of pragmatics in Swahili.

4.1. Most revealingly, the problem of the difference between the **na** TM and the **a** TM in shared grammatical contexts is paralleled by the problem of the difference between the **na** TM as in (7)b and the **0** TM as in (7)a in relative clause contexts. For example, Ashton (1944: 111) implies that **na**-Rel and **0**-Rel are distinct in the same way that **na** and **a** are in other contexts. To my knowledge, later scholars have not made more accurate observations. Contini-Morava is not unusual in completely ignoring the **0** TM (cf. p.15, where she lists all other TMs, including them with verb-final vowels as "verb markers").

Of great importance to the coherence of the auxiliary focus system is that just as the **a** TM has been declining in standard and Southern varieties of Swahili, so has the **0** TM, though with a slight time lag, leaving behind some frozen constructions of high frequency connected with particular verbs in particular uses. Some statistical figures demonstrating the parallelism of the decline are given in Table 2, presented in (8) below.

(8)

	% **na** (n = **a** + **na**)	% **na**-Rel (n = **0** + **na**-Rel)
Steere (1870)	60 (n=98)	42 (n=50)
AB (1935)	69 (n=188)	56 (n=50)
MWK (1960)	96 (n=150)	92 (n=52)

Table 2. Comparison of the percentage of **na** out of all occurrences (n) of **a** + **na** with percentage of **na**-Rel out of all occurrences (n) of **0** + **na**-Rel (except auxiliary taka) for one mid nineteenth century Zanzibar sample and two twentieth century standard samples.

All these samples represent coastal first language Southern varieties of Swahili. AB is a sample from one of the texts included in Contini-Morava's Table 1 in (3). MWK is another text which figures in her study, but it is clear from Table 2 that it was too poor in examples of **a** to be useful for inclusion in her count. Table 2 shows that it is also too poor in **0**-Rel in comparison to **na**-Rel. It is evident that the original system of focus, already very much narrowed in its domain of operation in comparison with such languages as Shambaa, has been continuing to collapse in favor of the highest focus marker of the set, **na**.[2] Sooner or later the problem I have been discussing will disappear from standard and various Southern varieties of Swahili.

5. So far discussion has only considered Southern and standard varieties of Swahili. In order to provide further evidence for the proposed declining auxiliary focus system of Swahili, I must turn to the status of this system in spoken Mombasa Swahili, where the **a** and **0** TMs are still very much alive, and much more favored than **na** in colloquial speech. It was Mombasa Swahili that first provoked my interest in the problems I have discussed above. One thing I discovered in Wald (1973) was that in Mombasa Swahili speech, there were different sets of verbs which have different favorabilities to **a** and **na** for discourse-pragmatic reasons. I have since found that these differential favorabilities cut across varieties, so that they are also properties of southern and even standard Swahili, to the extent that **a** remains at all in such varieties.

The set of verbs in all samples that are most favorable to **a** I call A verbs. A verbs include those most commonly used for representing states, of which the Swahili equivalents of "know", "be able", and "want" are most frequent in discourse. One thing I immediately noticed about the A verbs was that most of them were capable of taking infinitival complements, and thus were auxiliary-like. There were, however, three A verbs, also frequent in discourse, which showed the same favorabilities but could not take an infinitival complement. They are the Swahili equivalents of "call", as in "he's called Ali", and the two verbs equivalent to "tell" and "say". It took me until I arrived at the focus analysis of the Swahili tense-aspect system to realize that what these verbs have in common with the others is that they all pragmatically favor focus on postverbal complements. The complements of "tell" and "say" are the reported speech which most frequently follows these verbs as used in discourse. In this way it is evident that as the domain of **0** TM became restricted to relative clauses, the **a** TM extended to include **0**'s former function in main clauses of restricting the maximal focus of the clause to post-verbal position.

Currently, even in such non-Southern dialects as Mombasa, the focus system oriented toward postverbal constituents is no longer evident without such further conditions, which now appear to be disjunct from each other. Thus, both infinitival complements and reported speech remain as conditions favoring the use of the TM **a**. What these contexts may have in common, as opposed to such extinct conditions as postverbal objects or adverbs, is that they are potentially more complex syntactically and informationally heavier than the latter, because they imply an entire predicate which itself may include a postverbal object or adverb. Thus, they implicate more information than a postverbal object or adverb, and are worthy of higher focus than the latter. Still there is no grammatical obligatoriness to the choice of **a** or **na** in such contexts. While historically we can see such large constituents as residual holdouts to the complete loss of the constituent focus

system, synchronically we can suspect that **a** is preferred to **na** in such contexts because it is phonologically less substantial and thus is iconically more convenient for indicating the inherently superior focus of such postverbal material.

In (9) below, Table 3 compares a Mombasa speech sample with the historical samples of Table 2 ((8) above) for three verb classes.

(9)	A	MIXED	N
Mombasa (1973)	68 (n=294)	53 (n=152)	23 (n=31)
Steere (1870)	74 (n=19)	16 (n=32)	00 (n=2)
AB (1935)	38 (n=52)	21 (n=52)	00 (n=8)
MWK (1960)	13 (n=53)	03 (n=39)	00 (n=7)

Table 3. Percentage of **a** of total (n) occurrences of **na + a** for the three verb classes in a MOMBASA Swahili speech sample of 18 speakers (from Wald, 1973) and selections from the three Zanzibar / standard texts featured in Table 2 in (8) above.

The three verb classes are discourse-pragmatic sets, constant across all samples. A verbs include *taka* 'want, need', *weza* 'can, be able', *jua* 'know, find out, learn...', and *ita* 'call' *sema* 'say' *ambia* 'tell', among others. MIXED verbs are frequently occurring verbs in the Mombasa (1973) sample which exhibited no criterial preference for **a** and **na**. They include such common verbs as *ona* 'see, feel, think', *enda* 'go', *ja* 'come', among others. N verbs are a distinct set of verbs that are discourse-pragmatically most favorable to "resultant state" (i.e., achievement) contexts. They include *oa* 'get / be married', *simama* 'stand up /be erect', *kaa* 'stay, live, sit', *wa* 'be / become', among others. Table 3 shows the cross-dialectal validity of the three verb classes, as well as the precipitous decline in the use of the **a** TM in the twentieth century Southern / standard dialects.

The sample for Mombasa in Table 3 includes colloquial and more formal styles. Otherwise the percentage of **a** for all verb classes would be even higher. As it is, it can be seen that the Zanzibar pre-standard of Steere (1870), over a hundred years older than the Mombasa sample, is already polarized toward favoring **a** only with the A verbs. The twentieth century samples show a strong tendency toward across-the-board decline in the use of **a**, even though the A verbs remain relatively favorable. Note, in particular, that MWK (1960) exhibits only minimal retention of **a** across contexts. Nevertheless, it still shows greater favorability with A verbs than with other verb classes.

6. To conclude the theme of distinguishing semantics and pragmatics, the Swahili tense-aspect system seems to provide a case not well attested in the literature. Usually studies have focused on phenomena which start out pragmatic, grammaticalizing into obligatory syntactic features. In the collapse of the standard and Southern Swahili auxiliary focus system, certain focus relations which were earlier semantic, with syntactic consequences, present the appearance of being used pragmatically rather than with the obligatoriness associated with grammaticalization. Thus, there is no obligatoriness left to the choice between **a** and **na** in main clauses or **0** and **na** in relative clauses. There is simply a lingering tendency, much more strongly operative in Mombasa than in the South, to iconically prefer the phonologically less substantial marker when the verb is less in focus than

postverbal material, either in the same clause or in a syntactically connected following clause.

In ending, it should be mentioned that there is much more to be said about a larger range of interconnected shifts in the tense-aspect systems of Swahili and other East Bantu as the auxiliary focus system continues to decline. The narrowing domain of the **a** and **0** TMs was not the beginning of this decline, nor is the decline restricted to Swahili among the Northeast Bantu languages. In most varieties of Swahili and many adjacent coastal languages the former "perfect" markers in the auxiliary focus system went into decline earlier, and that decline had other effects on the use of the Swahili TM **na** and the auxiliaries **mala* and *isha* (both meaning "finish / end"), particularly in Mombasa and the rural coastal dialects, which have not been discussed here.[3] Thus, the preceding paper should be seen as preliminary to investigation and discussion of the larger range of shifts in the East Bantu tense-aspect systems which started earlier and whose effects continue to evolve throughout the area.

NOTES

[1] Although the TM **me** is not further considered in the text, the final paragraph of the text alludes to its further importance to an understanding of the larger series of shifts in the tense-aspect system that also affect the TMs which are the subject of this paper. In such a larger discussion the semantic and pragmatic relations of **me** to the other two markers would have to be taken into account.

[2] Extension of a higher focus marker at the expense of the lower focus counterpart is paralleled elsewhere in East Bantu. Thus, it can be deduced for the Northeast Interior, where the higher focus marker is **ne-*, prefixed to verbs, and strictly functioning to retract maximal focus in the clause to include the verb. Gikuyu reflects the earlier system to the extent that **ne-* is incompatible with a post-verbal question word, e.g., (***nĩ-**) ũ-k ũ-gwata kĩ ĩ(***Foc.M-**) 2s-TM-take what) 'what are you taking hold of?' (cf. Barlow 1960: 44). Grammaticalization of a postverbal question word as obligatorily higher in focus than the verb is typical of functioning East Bantu focus systems. However, parallel to the case of the Swahili high focus marker **na** extending its domain at the expense of the lower focus markers **0** and **a**, some of the Lake Victoria Bantu languages show innovative compatibility of **ne-* with a post-verbal question word, e.g., Nkore-Kiga: **n**-oo-kora ki (***ne**-2s-do what) 'what are you doing?' (Taylor 1985: 189). Revealingly, describers of such languages do not recognize its function as the (former) focus marker **ne*, but characterise it as a "progressive" marker.

[3] Briefly, the earlier decline involved the loss of the postverbal "perfect" marker **-ile*, its replacement in the south by the TM **me** (< **mal-ile* 'finish-Pf'), originally an auxiliary, and the rise of a new "high focus" perfect using the auxiliary *-isha* 'finish'. In the Mombasa colloquial and the rural dialects the TM **me** did not replace **-ile*. Instead, the TM **na** adjusted to fulfill this function. In Mombasa the loss of the high focus previously associated with **na** in this function involved the reduction of **na** to a mono-segmental homorganic nasal **n**. Needless to say, discussion of the evidence for and precise details of the shifts involving the perfect markers in the former Swahili focus auxiliary system must be reserved for a separate occasion.

REFERENCES

AB. Anonymous. 1935. Hekaya za Abunuwas na hadithi nyingine. London: Macmillan and Company Ltd.

Ashton, E.O. 1944. Swahili Grammar (Including Intonation). London: Longman.

Barlow, A. Ruffell. 1960. Studies in Kikuyu Grammar and Idiom. Edinburgh: William Blackwood & Sons Ltd.

Besha, Ruth M. 1989. A Study of Tense and Aspect in Shambala. Berlin: Dietrich Reimer Verlag.

Comrie, Bernard. 1976. Aspect. Cambridge: Cambridge University Press.

Contini-Morava , Ellen. 1989. Discourse Pragmatics and Semantic Categorization: The Case of Negation and Tense-Aspect with Special Reference to Swahili. Berlin & New York: Mouton de Gruyter.

Doke, Clement M. 1968. Textbook of Zulu Grammar. Johannesburg: Longmans Southern Africa Ltd.

Hyman, Larry M. & John R. Watters 1984. Auxiliary focus. Studies in African Linguistics. 15 (3).233-73.

MWK. Abdulla, Muhammed S. 1960. Mzimu wa watu wa kale. Nairobi: East African Literature Bureau.

Odden, David. 1982. Tonal phenomena in KiShambaa. Studies in African Linguistics.13 (2).177-208.

Roehl, Karl. 1911. Versuch einer systematischen Grammatik der Schambala-sprache. Hamburg: L. Friederichsen & Co.

Steere, Edward. 1870. Swahili Tales as told by Natives of Zanzibar. London: The Society for Promoting Christian Knowledge.

Taylor, Charles. 1985. Nkore-Kiga. London: Croom Helm.

Vendler, Zeno. 1967. Linguistics in Philosophy. Ithaca: Cornell University Press.

Wald, Benji. 1973. Variation in the system of tense markers of Mombasa Swahili. Columbia University Ph.D. dissertation. University Microfilms International.

Whiteley, W.H. 1959. An introduction to the rural dialects of Zanzibar. Parts 1,2. Swahili. 30.41-69, 31.200-205.

BERKELEY LINGUISTICS SOCIETY

1203 Dwinelle Hall, University of California, Berkeley, CA 94720 USA
Telephone: (510) 642-5808, (510) 642-2757; e-mail: bls@socrates.berkeley.edu
http://www.linguistics.berkeley.edu/BLS

The Berkeley Linguistics Society is a graduate student organization at the University of California, Berkeley. The Society holds its annual meeting during Presidents' Day weekend in the middle of February. The meeting consists of three parts: a General Session on topics of varied linguistic interest, a Parasession on a particular linguistic topic, and, beginning in 1990, a Special Session on general linguistic topics of a certain geographic area.

Book Prices*

		Parasession & Special Session (S) Topics	Individuals	Institutions
BLS-1	(1975) through BLS-7 (1981)		$10.00	$15.00
BLS-8	(1982)		$14.00	$20.00
BLS-9	(1983)		$10.00	$15.00
BLS-10	(1984)	Subordination	$14.00	$20.00
BLS-11	(1985)	Poetics, Metrics, and Prosody	$14.00	$20.00
BLS-12	(1986)	Semantic Typology	$14.00	$20.00
BLS-13	(1987)	Grammar and Cognition	$14.00	$20.00
BLS-14	(1988)	Grammaticalization	$14.00	$20.00
BLS-15	(1989)	Theoretical Issues in Language Reconstruction	$14.00	$20.00
BLS-16	(1990)	Legacy of Grice	$16.00	$20.00
BLS-16S	(1990)	General Topics in American Indian Linguistics	$8.00	$12.00
BLS-17	(1991)	The Grammar of Event Structure	$16.00	$20.00
BLS-17S	(1991)	African Language Structures	$8.00	$12.00
BLS-18	(1992)	Place of Morphology in a Grammar	$16.00	$20.00
BLS-18S	(1992)	Typology of Tone Languages	$8.00	$12.00
BLS-19	(1993)	Semantic Typology and Universals	$16.00	$20.00
BLS-19S	(1993)	Syntactic Issues in Native American Languages	$8.00	$12.00
BLS-20	(1994)	Anniversary Session: Dedicated to the Contributions of Charles J. Fillmore	$16.00	$20.00
BLS-20S	(1994)	Historical Issues in African Linguistics	$8.00	$12.00
BLS-21	(1995)	Historical Issues in Sociolinguistics /Social Issues in Historical Linguistics	$16.00	$20.00
BLS-21S	(1995)	Discourse Issues in Southeast Asian Languages	$8.00	$12.00
BLS-22	(1996)	The Role of Learnability in Grammatical Theory	$16.00	$20.00
BLS-22S	(1996)	Historical Issues in Native American Languages	$8.00	$12.00
Author Index to BLS 1-16S			$2.00 (includes shipping)	

California Sales Tax

8.25% for residents of Alameda, Contra Costa, Los Angeles, San Benito, San Mateo, and Santa Clara counties.
8.00% for residents of Santa Cruz county.
7.75% for residents of Del Norte, Fresno, Imperial, Inyo, Madera, Orange, Riverside, Sacramento, San Bernardino, San Joaquin, and Santa Barbara counties.
7.50% for residents of Sonoma county.
7.00% for residents of San Diego county.
6.50% for residents of Monterey county.
7.25% for all other California residents.

Shipping and Handling	Within USA	Outside USA
BLS 1,2,6,7,9,11-15,16S-22S	$2.00	$3.00
BLS 3,4,5,8,10,16-22	$2.50	$4.00

ORDERS MUST BE ACCOMPANIED BY PAYMENT
Make checks and money orders payable to the Berkeley Linguistics Society
*The prices may go up any time. For current prices see the most recent BLS volume.