

Studies in African Linguistic Typology

edited by F. K. Erhard Voeltz

John Benjamins Publishing Company

Studies in African Linguistic Typology

Typological Studies in Language (TSL)

A companion series to the journal *Studies in Language*

General Editor Michael Noonan

Assistant Editors Spike Gildea, Suzanne Kemmer

Editorial Board

Wallace Chafe (Santa Barbara)	Charles Li (Santa Barbara)
Bernard Comrie (Leipzig)	Edith Moravcsik (Milwaukee)
R. M. W. Dixon (Melbourne)	Andrew Pawley (Canberra)
Matthew Dryer (Buffalo)	Doris Payne (Eugene, OR)
John Haiman (St Paul)	Frans Plank (Konstanz)
Bernd Heine (Köln)	Jerrold Sadock (Chicago)
Paul Hopper (Pittsburgh)	Dan Slobin (Berkeley)
Andrej Kibrik (Moscow)	Sandra Thompson (Santa Barbara)
Ronald Langacker (San Diego)	

Volumes in this series will be functionally and typologically oriented, covering specific topics in language by collecting together data from a wide variety of languages and language typologies. The orientation of the volumes will be substantive rather than formal, with the aim of investigating universals of human language via as broadly defined a data base as possible, leaning toward cross-linguistic, diachronic, developmental and live-discourse data.

Volume 64

Studies in African Linguistic Typology

Edited by F. K. Erhard Voeltz

Studies in African Linguistic Typology

Edited by

F. K. Erhard Voeltz

University of Cologne

John Benjamins Publishing Company

Amsterdam/Philadelphia



™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Studies in African Linguistic Typology / edited by F. K. Erhard Voeltz.
p. cm. (Typological Studies in Language, ISSN 0167-7373 ; v. 64)
Includes bibliographical references and indexes.
1. African languages--Congresses. 2. Typology (Linguistics)--
Congresses. I. Voeltz, Erhard Friedrich Karl; 1943- II. Series.

PL8002.S78 2005
496--dc22
ISBN 90 272 2975 9 (Hb; alk. paper)

2005053172

© 2005 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

This volume is dedicated to the memory of Joseph H. Greenberg.
No other linguist has so influenced African language studies.
We all still learn from him.

Table of contents

Introduction	IX
<i>F. K. Erhard Voeltz</i>	
Future tense and aspect markings in Southern Bantu	1
<i>H. M. Batibo</i>	
The marking of directional deixis in Somali: How typologically idiosyncratic is it?	13
<i>Philippe Bourdin</i>	
A typology of subject and object markers in African languages	43
<i>Denis Creissels</i>	
Head marking, dependent marking and constituent order in the Nilotic area	71
<i>Gerrit J. Dimmendaal</i>	
Agent phrases in Bantu passives	93
<i>Axel Fleisch</i>	
Grammaticalization of switch reference: Motivation and means	113
<i>Zygmunt Frajzyngier</i>	
Complex predicates based on generic auxiliaries as an areal feature in Northeast Africa	131
<i>Tom Güldemann</i>	
The OHO Constraint	155
<i>Richard J. Hayward</i>	
The word in Luganda	171
<i>Larry M. Hyman and Francis X. Katamba</i>	
Case in Africa: On categorial misbehavior	195
<i>Christa König</i>	
The typology of relative clause formation in African languages	209
<i>Tania Kuteva and Bernard Comrie</i>	
Deictic categories in particles and demonstratives in three Gur languages	229
<i>Kézié Koyenzi Lé bikaza</i>	

Preprefix or not – that is the question: The case of Kwangali, Kwanyama and Ndonga <i>Karsten Legère</i>	251
Nonverbal and verbal negations in Kabyle (Berber): A typological perspective <i>Amina Mettouchi</i>	263
Grammaticalization chains of the verb <i>Kare</i> ‘to give’ in Kabba <i>Rosmarie Moser</i>	277
Selectors in Cushitic <i>Maarten Mous</i>	303
How Bantu is Kiyansi? A re-examination of its verbal inflections <i>Salikoko S. Mufwene</i>	327
Diathesis alternation in some Gur languages <i>Brigitte Reineke and Gudrun Mieke</i>	337
Structure and function of incorporation processes in compounding <i>Claudia Maria Riehl and Christa Kilian-Hatz</i>	361
Toward a typological perspective for Emai’s BE constructions <i>Donald P. Schaefer and Francis O. Egbokhare</i>	377
Intrinsic focus and focus control in two varieties of Hausa <i>H. Ekkehard Wolff</i>	397
Language index	417
Name index	421
Subject index	425

Introduction

F. K. Erhard Voeltz

Universität zu Köln

The present volume is a selection of papers presented at the International Symposium: Typology of African Languages held at Sankt Augustin under the auspices of the Institut für Afrikanistik, Universität zu Köln. It was the aim of the symposium to give new impetus to African linguistic typological issues and to launch a large scale project roughly to be named '*Typology of African Languages*' the aim of which it was and is to build upon the foundation laid by Heine 1975.

The twenty-one papers brought together here reflect the broad perspective of African linguistic topology studies today. While similar volumes a generation ago would present language material from a very restricted area and perspective, the present contributions reflect the global interest and orientation of current African linguistic studies. The studies are nearly all implicational in nature. Based upon a detailed survey of a particular linguistic phenomenon in a given language or language area conclusions are drawn about the general nature of this phenomenon in the languages of Africa and beyond. They represent as such a first step that may ultimately lead to a more thorough understanding of African linguistic structures. This approach is well justified. Taking the other road, attempting to pick out linguistic details from often fairly superficially documented languages runs the risk that the data and its implications for the structure investigated might be misunderstood. Consequentially only very few studies of this nature giving the very broad perspective, the overview of a particular structure type covering the whole African continent are represented here.

African linguistic studies still form a cohesive body without too much theme-specific specialization. The papers are thus arranged here in simple alphabetic order as they were presented at the colloquium itself. We begin with Batibo's contribution, which deals with the source and development of certain TAM markers in two South African Bantu languages, Sotho and Zulu. While he is able to confirm observation made by Bybee et al. 1994 regarding some aspects of grammaticalization in the languages investigated, he nevertheless observes two points of broader issues not covered there: i. the use of a perfect marker '-ile' with 'future' meaning in Sotho, and ii. the very differential development of the TAM marker 'go' in Sotho and Zulu. In the former

language 'go' has developed into a 'near future' marker, while in the latter, the notions 'present' and 'far future' are now expressed.

Bourdin's detailed discussion very convincingly shows directional deixis and that the manner of encoding it is not an idiosyncratic property of Somali but based on extensive cross-linguistic evidence, a property of a fairly large number of languages. Creissells investigates subject and object marking in African languages, looking at pronominal/bound indexing on the verb. He comes to the following observations: A very large number of languages in Africa use bound morphemes to mark the subject. Fairly often the agreement markers fuse with the TAM or polarity markers. When languages have what he calls stage I subject markers, they can often not be distinguished from independent pronouns. It is fairly rare to find languages which allow only bound subject markers but not also bound object makers. 'Exotic' patterns of subject and object marking as found in 'ergative' languages are very rare in Africa. In ditransitive/three-argument verbs it is the overwhelming tendency to index the recipient/the patient/the 'indirect' object rather than the 'direct' object.

Dimmendaal deals primarily with the notion of areal types and areal diffusion. Making reference to Nilo-Saharan as a whole, he shows that in one sub-branch, Nilotic, head marking rather than dependent marking has become the rule. It is argued that this change can be related to the fact that the languages are predominantly verb-initial languages. As the next step in the development some Eastern Nilotic languages have developed (again) dependent marking while keeping head-marking at the clausal languages. In neighboring (related) Surmic languages variations of head- and dependent marking must be viewed as resulting from contact with other languages and a high degree of multilingualism.

Different strategies used for marking 'agent' in Bantu passives are the focus of Fleisch's contribution. While Bantu languages display a high degree of homogeneity in the formation of the passives, a fairly large number of strategies for marking agents can be observed. In most cases these agent phrases are morphologically transparent. The choice of which kind of agent marking is used appears to depend much more on areal than genetic features of a given language. The agent phrase is, moreover, very frequently also used in other grammatical constructions, showing that there is not an integral link between the passive verbal piece and the agent as such.

Frajzyngier demonstrates that grammaticalization can be motivated by the internal properties of a grammatical system. This kind of motivation does not play a role as a notion in most contemporary studies of grammaticalization. Secondly he presents a specific analysis of the grammaticalization of switch-reference of Mina, demonstrating how morphological switch-reference markers grammaticalized from two types of pronouns. It is shown that the coding of co-reference and switch reference is a consequence of the language having already grammaticalized the category 'subject'.

As already argued in Güldemann (2001) for the Bantu language, in his present paper Güldemann argues that elements introducing complex predicates do not derive from verbs with the basic meaning "to say" but rather from verbal lexemes which originally encoded semantically more generic notions (e.g., manner deixis and similarity

'like (this)', action 'do, make', inchoativity 'become', and equation 'be') and which have been subject to parallel grammaticalization in the domains of predicate formation and reported discourse.

Hayward convincingly argues for a constraint, loosely called the OHO constraint, for high-pitch placement in words and phrases in languages of the Ethio-Eritrean area. This constraint can be observed in a fairly large range of typologically different languages in the area and elsewhere in the world. What the languages have in common is that they are all head-final, i.e., all elements modifying precede rather than follow their head. Should the constraint turn out to be a quite general one, it will naturally lead to the important question of how to account for it in a theoretically satisfying way.

Based upon data from Luganda, Hyman & Katamba attempt to provide answers to the following three questions: i. Can the word be defined? ii. If not, why not? iii. If yes, is the word a universal? Although intuitively answers to these questions should be fairly straight forward, answers turn out to not to be simple at all. Two factors tend to influence any attempt to provide a universal answer to this question: a. Even within a given language there are conflicts between the different components of the word (morphology, syntax, phonology). b. And even within the same component conflicts arise with such properties as vowel length or tone.

König investigates instances of categorial misbehavior of "cases" in Ik. While the explanation for this misbehavior is preliminary, it is reasonable to hypothesize that all instances will have followed a similar path of grammaticalization as here outlined.

In their typological study of relativization, Kuteva & Comrie survey a broad range of languages across all of Africa, establishing that only three of the six known strategies for relativization can be found here. While only half of the known strategies are attested, there exists in some cases extreme parsimony of marking, e.g. Maale (Omotic) and great diversity, e.g. Ngemba (Bantoid), where up to five relativization markers can be found in a given phrase.

A detailed analysis of deixis for three Gur languages, Kabiye, Tem and Nawdem, is presented by Lebigaza. Particular attention is given the relationship of deictic particles to demonstratives, in an effort to explain their fairly differentiated use in syntax, discourse and given social settings.

Legère takes up the old Bantu issue of the role and status of preprefixes. Preprefixes are mostly vocalic elements, often shadow vowels, placed before the nominal class prefixes. They occur in a fairly large number of Bantu languages but their occurrence does not correspond to any boundaries recognized by linguistic subgrouping. Legère presents evidence to show that the pre-prefixing system in a number of South-West Bantu languages is being eroded in some syntactic environments while being retained in others.

Mettouchi shows that negation and aspect are intimately related on the semantic level. Investigating TAM and negation encoding in Kabyle as well as in a number of other African languages, she observes that in languages that use different TAM markers for the affirmative and negative subsystem and that also have different non-verbal negation the core values of aspectual forms mirror the core value of non-verbal

negation. She concludes with the hypothesis that languages that use strictly aspectual systems are more likely to make extensive use of non-verbal negation, whereas languages with tense based systems tend to replace them in time with copular forms with TAM distinctions.

Moser analyses in much detail the verb *kàre* with the basic meaning “give” in Kabba. Beyond its basic function as a recipient marker for intrinsically ditransitive verb, *kàre* also has a basic causative function (‘cause, make’), with the extended meaning ‘force’ and even ‘kill’, allowing any kind of sentential complements. Thus she provides evidence for the identical source of two quite different constructions, viz. verb serialization and verb complementation.

Many Cushitic languages have an additional inflectional element in the middle of the sentence that is separate from the verb. Mous calls these elements ‘selectors’. In his article he gives an overview of these elements, of their functions and of which categories are expressed on them. What most of these so-called selectors have in common is the marking of sentence type and/or focus, and of subject. The comparison results in three types of selectors: (i) those that define the left border of a syntactic unit such as the verbal piece in Somali; (ii) those that indicate focus as a pro-clitic to the verb, and (iii) those that indicate focus by their position in the sentence.

Mufwene deals with one of the elementary issues in comparative linguistics: The discrepancy between genetic and typological classification of languages. Looking at data from a fairly closely related set of languages he shows that in respect to the verb system one language in particular, Kiyansi, behaves very atypical and set apart from the normal pattern. Where there is no evidence of areal-typological influence, do such developments reflect earlier stages or language internal motivated changes? While no answer can yet be given, Mufwene’s discussion does draw attention to certain parallels in Germanic linguistics.

Reineke & Miehe investigate changes of valence of verbs in Gur and Mande languages. While such changes in other Niger-Congo languages are often indicated by head marking, Gur and Mande languages have developed a strategy whereby a second argument of a transitive verb is either present or suppressed. Verbs that are usually seen as being very rigid as far as their transitivity is concerned are becoming quite flexible. The role formerly played by, for example verb extension, is thus taken over by word order. Allowing the possibility of admitting or suppressing arguments of the verb largely increases the number of intransitives – normally a very restricted set in African languages – to a very large, perhaps unlimited set.

Riehl & Kilian-Hatz show that incorporation seems to be a very common strategy used not only in polysynthetic languages, but also in all types of languages to integrate lexical items into a larger lexical complex also known as nominal compounds. The function of this process is not only a semantic one, i.e. narrowing the scope of the predication, but mainly a syntactic one, i.e. back-grounding the given information. Whereas polysynthetic languages use noun incorporation as a strategy to background nouns in otherwise fore-grounded sentences, rather analytic languages use it to back-ground the whole predication by nominalization of the verb-noun complex.

As many other African languages, Emai exhibits a number of forms of the verb 'to be' expressing both predication and identification. The distribution of the various forms is not entirely symmetric. Schaefer & Egbokhare attempt to account for these asymmetries by proposing a hierarchy of stability that might prove useful in the typological characterization of BE constructions in other languages, especially those in Africa and the Edoid family.

In the second paper in this volume dealing with 'focus' Wolff provides a detailed analysis of the various focus parameters operating in Hausa. He observes that both *assertive* and *contrastive focus* operate as distinct semantic-syntactic categories in Hausa; that both PERFECT and PROGRESSIVE have *intrinsic focus* properties in terms of *assertive predication focus*; and that *assertive predication focus* is a valid and dynamic semantic-syntactic operation with some history and variance in modern Hausa varieties including Standard Hausa. To govern these three parameters, there are two different control mechanisms at work in the language, i.e. speaker's choice (i.e. *pragmatic control*) and system-internal *grammatical control*. The significant distinction between pragmatic and grammatical control in Hausa, as in many other Chadic languages, is not restricted to issues of information structure, i.e. various dimensions of focus, but also operates in the domain of verbal plurality, which, in Chadic, tends to cross the borderline between inflectional and derivational morphology in both directions.

It is our pleasure to thank the director of the Arnold-Janssen-Haus of the Steyler Missionare, Sankt Augustin, Dr. Hermann Kochanek and Ms. Marlies Dahmen for their hospitality and the congenial and relaxed atmosphere during our symposium.

The symposium was supported with a generous grant from the Deutsche Forschungsgemeinschaft to whom we wish to express our sincere thanks and the Institut für Afrikanistik, Universität zu Köln and its then director, Bernd Heine, who has never failed to support and encourage our endeavors, however unconventional they may have been.

References

- Bybee, J. L., Perkins, R. D., & Pagliuca, W. (1994). *The Evolution of Grammar, Tense, Aspect, and Modality in the Languages of the World*. Chicago, IL: University of Chicago Press.
- Güldemann, T. (2001). Quotative Constructions in African Languages: A synchronic and diachronic survey. 'Habilitation' thesis. Institute für Afrikanistik, Universität Leipzig.
- Heine, B. (1976). *A Typology of African Languages Based on the Order of Meaningful Elements*. Berlin: Dietrich Reimer [Kölner Beiträge zur Afrikanistik 04].

Future tense and aspect markings in Southern Bantu

H. M. Batibo

University of Botswana

Like most other languages of the world, Southern Bantu languages are characterized by verb forms, which refer to the time, state or circumstances associated with an action, event or process. These verb forms are said to mark tense, aspect and mood (TAM). Most of the Southern Bantu languages have a wide range of TAM markers, which provide a rich array of temporal and aspectual notions.

This chapter examines the most common future TAM markers in Southern Bantu languages. The main argument of the study is that many forms have evolved in Southern Bantu, mainly through the process of grammaticalization of certain types of verbs, which refer to a variety of future notions. Although one notices many variations of form in the future TAM markers across the Southern Bantu languages, there are several common features, which raise issues of theoretical interest.

1. Tense and aspect in Bantu

1.1 The notions of tense and aspect

As it is also the case in many other languages, the Bantu language verbal forms are characterized by tense and aspect in their time event marking notions.¹ While tenses represent time in relation to the event within the universe time, aspects represent the process or state of the event within a given time (Dubois et al. 1973). The two notions are usually complementary, although one form may project one of the notions more than the other. The usual tense markers consist of the present, the past and the future and may sub-divide into specified stretches of time within these time categories. Tense marking is remarkably varied in Bantu languages. The commonest pre-stem marker across Bantu is *-a-*. This form comes in several shapes (short vowel, long vowel, high-toned, low-toned) and has a variety of signification, the commonest being present, near past and far past.

The usual aspect categories include the perfect, the habitual and the progressive. The most common forms in Bantu languages are the perfect suffix *-ile*, believed to

be one of the earliest temporal notions in Bantu (Nurse 2003:96), and the imperfective suffix *-a(n)g-*, which often appears as a habitual, an iterative or a continuous. Since many of the temporal notions combine tense with aspect and mood, they are commonly known as Tense, Aspect and Mood (TAM) markers.

1.2 Modal notions in Bantu

Bantu languages have also other temporal notions. These include modal forms, which represent the speaker's communicative intentions or attitude. The most common modal forms in Bantu languages include the indicative, usually represented by the suffix *-a* and the subjunctive, represented by the suffix *-e*. Other modal forms have been identified such as the conditional (or contingent), the potential and the participial (Doke 1967:71). Moreover, there are other forms which are neither modal nor aspectual, which are common in Bantu languages. These notions include the narrative, the persistive and the consecutive (Hyman & Watters 1984). It could also be argued (Nurse pc) that the narrative and the consecutive are relative tenses, while the persistive is an aspect.

1.3 Future tense and aspect marking in Bantu

As rightly stated by Nurse (2003:93), future is problematic as it represents an event which has not yet happened. Usually if it is only a short temporal distance away, it may be represented by progressive or habitual notions, as many Bantu languages do. This is a natural semantic extension of the present. All futures are characterized by notions of intent, wish, expectation and uncertainty. Hence, they have given rise to many forms expressing different levels and types of futurity. According to Bybee et al. (1987, 1994; Dahl 1985), the agent intention is the crucial bridge to the formation of future notions. Hence, the most common sources for such forms are those which yield the intention inferences most easily, such as desire, obligation and movement towards a goal. Although there are no widespread future forms in Bantu, many of the Bantu languages appear to have grammaticalized the motion verbs for "come" and "go" into future notions. One other phenomenon about Bantu languages is that usually the future forms are less in number than the past tense/aspect forms. This is presumably because of the uncertainty of future events. Some languages, like Kiswahili, have only one probable future form.² The Kiswahili future form is *-ta-* as shown below:

- (1) *tu - ta - nunu.a*
we.FUT.buy.FV
'We shall buy'

Other Bantu languages may have as many as four forms. This is the case of Logooli, as exemplified below (Nurse forthcoming).

- (2) a. *ku - ra - gor.a*³
 we - FUT₁ - buy.FV
 ‘We will buy’ (Near Future)
- b. *na - ku - gor.i*
 FUT₂ - we - buy.FV
 ‘We will buy’ (Middle Future)
- c. *ku - ri. ka - gor.a*
 we - FUT₃ - buy.FV
 ‘We will buy’ (Far Future)
- d. *ku - ri - gor.a*
 we - FUT₄ - buy.FV
 ‘We will buy’ (Uncertain Future)

The most widespread forms for futurity, particularly in North – Eastern Bantu are *-laa-* and *-ka-*. In many Bantu languages, however, futurity is not only interpreted in the tense and aspect notions but also in the modal and conditional dimensions.

2. Future tense and aspect marking in Southern Bantu

2.1 Sources of future tense and aspect marking in Southern Bantu

In this study we make two assumptions. The first assumption is that there is a certain historical unity between the Southern Bantu languages, that is most of the languages found in Guthrie’s Zone S (Guthrie 1948, 1967–71), which is demonstrated in the various linguistic forms. Such unity has been observed elsewhere by some of the other studies (Doke 1954; Fagan 1984; Louw & Finlayson 1990; Janson 1991/92; Batibo et al. 1996; Batibo 1998). The languages which belong to Southern Bantu include all the languages and the respective varieties, belonging to Nguni, Sotho/Tswana, Venda, Tsonga and Inhambane groups. Unlike Doke (1954), Shona is excluded in our grouping. The second assumption is that following this historical unity; the notions and forms for future tense/aspect marking are identical.

According to a survey which was carried out on the languages of Southern Bantu,⁴ most future tense and aspect markers appear to originate from the motion verb ‘come’ which may be used as Tense, Aspect and Mood (TAM) marker within a verb form or as an auxiliary form followed by the main verb. The main verb may be in the infinitive form or may be inflected by a verbal prefix. This scenario makes it possible to establish two stages through which the Southern Bantu languages appear to be going, or to have gone through, in their process of grammaticalization of the verb ‘come’ into a future TAM marker.⁵

a) *Stage I: VP – Aux Infinitive or VP – Aux VP – Verb*

This is the first stage in the process of grammaticalization in which the verb ‘come’ has reached in a number of Southern Bantu languages. The verb has assumed auxiliary functions, although in many of the languages it still maintains its full verbal characteristics, including the range of its semantic uses elsewhere. However, it has assumed grammatical dimensions through its present use as an auxiliary to mark futurity. The verb that follows may have first served as a complement in its infinitive form, but its role has evolved to that of the main verb. Cases with infinitival forms are very common in Southern Bantu, as they are found in most of the Nguni languages as well as Sotho/Tswana and Venda.

(3) a. Isizulu:

ngi - za uku - thanda > ngi - zoo - kuthanda

I - come to - love

‘I shall love’

b. Northern Sotho (Sepedi):

o - tla (g)o - reka > o - tl - o - reka

he - come to - buy

‘He will buy’

c. Southern Sotho (Sesotho):

Ke - tl - ile (g)o - reka > ke - tlil - o - reka

I - come - PERF to - buy

‘I shall have bought’

d. Venda:

ndi - da (k)u - vhona > ndi - doo - vhona

I - come to - see

‘I shall see’

There are, however, a few cases in which the main verb has been inflected with the Verb Prefix. Hence the Verb prefix is found on both the auxiliary and the main verb. This is the case of the present progressive form in Isizulu, where the main verb prefix is inflected in the present progressive form.

(4) Isizulu:

ngi - ya ngi - hamba

I - go I - travel

‘I am traveling’

b) *Stage II: VP - TAM - Verb*

This is the second stage in the process of grammaticalization in which the languages of Southern Bantu have gone through. The verb ‘come’, in its auxiliary function has progressively developed into a future tense, aspect and mood (TAM) marker, thus structurally transforming itself from a syntactic constituent to an affixal element. In most cases, the forms have continued being used both as lexical and grammatical units. In

some cases morphophonological contractions have taken place in reducing the auxiliary forms to prefixes. Also, in languages like Setswana, a new convention has been established to distinguish between the form of the future TAM and that of the verb ‘come’. Cases in which auxiliary forms have developed into future TAM’s are found in most of the Sotho/Tswana languages as well as some of the Tsonga languages, as exemplified below.

- | | |
|--|---|
| <p>(5) a. Setswana:</p> <p><i>o - tlaa - bona</i>⁶</p> <p>he - FUT - see</p> <p>‘He will see’</p> | <p>d. Southern Sotho (Sesotho):</p> <p><i>ke - tla - reka</i></p> <p>I - FUT - see</p> <p>‘I shall buy’</p> |
| <p>b. Sekgalagadi:</p> <p><i>o - da - bona</i></p> <p>He - FUT - see</p> <p>‘He will see’</p> | <p>e. Gitonga (Tsonga):</p> <p><i>ndi - ta vona</i></p> <p>I - FUT - see</p> <p>‘I shall see’</p> |
| <p>c. Sebirwa:</p> <p><i>o - ta - bona</i></p> <p>he - FUT - see</p> <p>‘He will see’</p> | |

However, it is important to state here that, although the usual grammaticalization direction is from a periphrastic to affixal formation, it should not be expected that all the future markers in Southern Bantu will eventually evolve into TAM prefixes. In fact as Bybee et al. (1994:267) observe, most future markers, cross-linguistically, tend to appear as auxiliaries rather than affixal elements, thus many of these forms may remain auxiliaries. Another important remark is that the formal changes which have occurred have usually proceeded in parallel with the semantic changes in the grammaticalization process, where the original lexical usage has been progressively shifted from ‘come’ through the allative meaning of ‘come to’ and then metaphorically transformed from spatial to temporal connotation (Sweetser 1988; Emmation 1992).

2.2 Other future TAM forms in Southern Bantu

Although the majority of the Southern Bantu languages, particularly those of Nguni, Sotho/Tswana and Venda origins, have adopted the motion verb ‘come’ (< Proto-Bantu *-ja), some of the languages have other forms which may be used to capture other future notions.

A number of languages have adopted the verb ‘have/ be’ (i.e. -na), which, in fact, originated from the connective *na* ‘and, with’ (Guthrie 1967:243). For example, in Gitonga (Inhambane), the form -na- is used as the general future TAM.

- (6) *s.anana si - na - emba*
 children they - FUT - sing
 ‘The children will sing’

In Isixhosa, the form *na* has been adopted to refer to a general future.

- (7) *u - na u(k)u - bona > u - noo - bona*
 he - is to - see
 ‘He will see’

In fact, the grammaticalization of the verb ‘have, to be’ to futurity is one of the common developments cross-linguistically (Bybee et al. 1991; Heine et al. 1991). This is because the respective semantic content is often associated with intentions and directionality. Southern Bantu has extensively exploited this source.

On the other hand, both Nguni and Southern Sotho have added a new notion of futurity by adopting the verb “to go” (Proto-Bantu **gia*). This has allowed these languages to assign several time spans to futurity.

- (8) a. Isixhosa:
ndi - ya ku - hamba
 I - go to - travel
 ‘I shall travel’ (Far Future)
- b. Isizulu:
ngi - ya uku - hamba > ngi - yoo - kuhamba
 I - go to - travel
 ‘I shall travel’ (Far Future).
- c. Southern Sotho (Sesotho):
ke - ea - reka
 I - FUT - buy
 ‘I shall buy’ (Near Future).

It is interesting to note that the values of futurity accorded to the adopted form do not always correspond. While in the Nguni languages, it has acquired the notion of the “Far Future”, it is “Near Future” in Southern Sotho. On the other hand, the auxiliary form *-ya* is used in the Nguni languages, without the infinitive marker, to refer to the present. This is the case of Isizulu where the form has reached Stage II in its present TAM marking, but still at stage I in its future form. Thus, in the example below, *-ya* has reached Stage II in (10a), while it is still at Stage I in (10b).

- (9) a. *ngi - ya - hamba*
 I - FUT - travel
 ‘I am traveling’ (Present).
- b. *ngi - ya uku - hamba > ngi - yoo - kuhamba*
 I - go to - travel
 ‘Shall travel’ (Far Future).

2.3 The use of the Potential/Conditional form *-nga-*

In many of the Southern Bantu languages, the form *-nga-* (Conditional, Potential) is used to express futurity. This form (**nga* in Proto-Bantu, with the meaning of 'if') tends to focus mainly on modal notions as it combines potentiality with conditionality to express future probability or eventuality with elements of uncertainty.

- (10) a. Venda: *ndi - nga - vhona*
I - POTEN - see
'I might see'
- c. Setswana: *o -ka - lema*
he - POTEN - cultivate
'He/she can cultivate'
- b. Ronga (Tsonga) *ndi - nga - bona*
I - POTEN - see
'I might see'

In such languages, there are usually at least two forms, the normal future expressing probable happenings and the potential future, expressing events which might happen. In the Nguni languages, however, this form is combined with the past form of the auxiliary *-ya* to refer to events which might have taken places.

- (11) *Nga - ngi - ya u(k)u - hamba* > *nga - ngi - yoo - hamba*
POTEN - I - go to - travel
'I could have traveled, but...'

3. Negative forms

Although most of the Southern Bantu languages have regular negative markers emanating from the Proto-Bantu forms **ka* and **ta*, the negative future marking tends to occur in various and often complex forms. As remarked by Doke, sometimes a probable future marker may be replaced by a potential form in the negative, as in the following Venda example (Doke 1967:70).

- (12) a. *ndi - da u(k)u - divh.a* > *ndi - do - divha*
I - come to - know
'I shall know' (Affirmative/Probable)
- b. *ndi - thi - nga - divh.i⁷*
I - NEG - POTEN - know
'I shall not know' (Negative/Potential).

In the negative sentence, the potential form has to be used. At the same time the suffix of the main verb is changed to *-i* to comply with the usual negative construction.

On the other hand, in some of the future negative formations, the infinitive form is introduced even where in the affirmative it was not. Evidently, this could be seen as a retention of an earlier construction. This is the case of Setswana.

- (13) a. *b.ana ba - tlaa - rata dijo*
 children they - FUT - like food
 ‘The children will like the food’ (Affirmative)
- b. *b.ana ga ba - na go - rata dijo*⁸
 children NEG they are to - like food
 ‘The children will not like the food’ (Negative)

4. Compound forms involving futurity

In all Southern Bantu languages, it is possible to place futurity within various time, aspect and modal contexts. This may result in compound or poly – clausal forms, involving the verb “to be” as an auxiliary verb. Usually, the tense will be associated with the auxiliary, while the aspect and mood will go with the main verb. This is demonstrated below:

- (14) a. Isizulu:
ngi - za uku - be ngi - thanda > ngi - zoo - kube ngi - thanda
 I - come to - be I - love
 ‘I shall be loving’ (combining futurity with continuity).
- b. Southern Sotho:
ke - tla - be ke - sa - rat.e
 I - FUT - be I - still - love
 ‘I shall not be loving’ (combining futurity with progressiveness).
- c. Southern Sotho:
ke - tla - be ke - rat - ile
 I - FUT - be I - love - PERF
 ‘I shall have loved’ (combining futurity with the perfect notion).
- d. Northern Sotho:
ke - tla (g)o - ba n - ka - be ke - rat.a
 I - come to - be I - POTEN - be I - love
 ‘I shall be able to be loving’
 (combining futurity with potentiality and progressiveness).

5. Conclusion

There are certainly many common elements in the Southern Bantu languages with regard to future tense and aspect marking. It is, however, difficult to state whether this

similarity is due to a common history or a mere typological coincidence. The common features which have been identified include:

First, in most languages, the motion verb ‘come’ has been adopted as the future tense and aspect marker. If we go by the usual grammaticalization process (Heine et al. 1991), one would have expected both motion and volition verbs. No volition verbs were found in our data. The predominance of the verb ‘come’ is also felt in some of the other zones, namely in south Central Bantu (e.g. Chikuhane) and South Western Bantu (e.g. Otjiherero).

Second, compared with the other Bantu languages, particularly those in North-Eastern Bantu, such as Logooli and Mwera with at least four future forms, Southern Bantu languages can be said to have very limited future markings. Many languages have no more than two forms, one as the probable future and the other as the improbable or potential future. The restricted number of fully grammaticalized forms may imply that until very recently the temporal distinction in Southern Bantu was only between past and non-past markers.

Third, there appears to be an on – going process to expand the number of futurity notions through the process of grammaticalization and re-assignment of temporal roles. Some languages, like Sesotho (Southern Sotho), have managed to build up their futurity notions through the adoption of new forms and the re-assignment of new temporal roles to already existing forms.

- (15) a. *ke - ea - rek.a*
 I - FUT - buy.FV
 ‘I shall buy’ (Near Future, with verb ‘to go’)
- b. *ke - tla - reka*
 I - FUT - buy
 ‘I shall buy’ (General Future, with verb ‘to come’)
- c. *ke - ile (g)o - reka⁹*
 I - GO.PERF to - buy
 ‘I shall buy’ (Immediate Future, with the perfect form of the verb ‘to go’)
- d. *ke - tl.ile (g)o - rek⁹*
 I - come.PERF to - buy
 ‘I shall buy’ (Probable Future with the perfect form of the verb ‘to come’).

One important question is what motivates languages to expand and re-define their tense/aspect notions, given that some languages, like Kiswahili, as seen above, are comfortable with only one probable future form. Kiswahili, however, has created further past or perfect forms, such as the recent creation with the verb *-isha* ‘finish’ which has been grammaticalized in the colloquial usage to *-sha-* ‘perfect’. The source of motivation, in many of the Bantu languages, could involve internal reorganization but also external forces such as contact with other languages, as observed among the East African Bantu languages which have come into contact with Nilotic languages (Emanation 1992).

Fourth, the several variations in Southern Bantu, sometimes even between closely related languages, imply that there is a lot of dynamism in the evolution of the tense and aspect marking systems, particularly with regard to the futurity notions. Most of the languages appear to have adopted innovative strategies in meeting the optimal expression of futurity in their languages.

Lastly, most of the findings in this study have confirmed the generalizations made in Bybee et al. (1994) who studied the evolution of the tense, aspect and mood forms cross-linguistically. However, this study has revealed at least two developments of theoretical interest. The first one concerns the use of a non primary source, like the perfect form *ile*, in a language such as Sesotho, to represent the notion of 'immediate future', when according to Bybee (op. cit. 244) usually immediate future notions are represented by forms from primary sources like verbs of motion, volition and obligation. The second point of theoretical interest is that, while the verbs with the meaning of 'come' and 'go' have been grammaticalized to represent notions of relative nearness and distance respectively with regard to futurity, coinciding with the cross-linguistic findings of Bybee et al. (1994: 280), our study has shown also that each language has its own grammaticalization path towards a specific time stretch within the futurity span. For example, while the form for 'go' has developed to represent the notion of 'near future' in Sesotho, an identical form has evolved into the notions of 'present' and 'far future' in Isizulu.

Notes

1. I wish to thank very sincerely my colleagues, particularly Derek Nurse, Bernd Heine, Erhard Voeltz, Gerrit Dimmendaal and Christa Koenig for their comments and suggestions to the earlier form of this paper.
2. Of course, Kiswahili uses other forms to express certain aspects of futurity, such the extension of the present and the form *-ki-* for conditional or uncertain future events.
3. For practical reasons, the examples will be written in the orthographic conventions of the respective languages or as transcribed in the original sources. Since most of the conventions have not marked the tones, tones will not be marked, except where special distinctions are desired.
4. This survey was mainly documentary. However, some interviews were also conducted with the speakers to verify some of the data.
5. One example to provide evidence to the above stages is the case of Kiswahili, where the verb *ku - taka* 'to want' was first used as an auxiliary with the main verb in the infinitive form (e.g. *a - taka ku - nunua* 'he wants to buy'). Later, during the second stage of the grammaticalization process, not only the infinitival form of the main verb dropped out but also the form *- taka* contracted to become a monosyllabic prefix (e.g. *a - ta - nunua* 'he will buy'). However, the infinitival form resurfaces when the stem of the main verb is monosyllabic (e.g. *a - ta - kuja* 'he will come'). Also the original auxiliary form *-taka* reappears in certain constructions, such as *a - taka - po - kuja* 'when he comes'.

6. In order to distinguish between *tla* 'come' and *tla* 'future TAM' an orthographic convention has been established for Setswana in which the vowel of the future TAM is doubled, i.e. *tlaa*.
7. It could also be argued that the formative *-nga-*, in this context, has a different origin and that it has no potential functions.
8. In Setswana, the formative *na* (originally meaning 'and, with') is only found in negative and compound forms, with the meaning of 'be'.
9. The fact that *-ea* and *tla* can be conjugated into perfect forms would imply that, although they are now considered as FUT.TAM markers, they still maintain some auxiliary characteristics. However, in our study we consider the forms with the perfect as newly grammaticalized FUT.TAM markers, formally and functionally different from those without the perfect.

References

- Batibo, H. M. (1998). A lexicostatistical survey of the Bantu languages of Botswana. *The South African Journal of African Languages*, 18 (1), 22–28.
- — —, Moilwa, J. M., & Mosaka, N. M. (1996). The historical implications of the linguistic relationship between Makua and Sotho. In K. Darkwah (Ed.), *Festschrift for Professor Leonard D. Ngcongco* (pp. 23–29). (A Special Issue of *Pula*, 11, 1).
- Bybee, J. (1987). The evolution of future meaning. In A. G. T. Ramat et al. (Eds.), *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam: John Benjamins.
- Bybee, J. L., Parkins, R. D., & Pagliuca, W. (1991). Back to the future. In E. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*, Vol. II (pp. 17–58). Amsterdam: John Benjamins.
- Bybee, J. L., Parkins, R. D., & Pagliuca, W. (1994). *The Evolution of Grammar, Tense, Aspect, and Modality in the Languages of the World*. Chicago, IL: University of Chicago Press.
- Dahl, O. (1985). *Tense and Aspect Systems*. Oxford: Blackwell.
- Doke, C. M. (1967). *The Southern Bantu Languages*. London: Dawson of Pall Mall for the International African Institute.
- Doke, C. M. (1954). *The Southern Bantu Languages*. London: Oxford University Press. (reprinted by Dawsons of Pall Mall, London for the International African Institute, 1967 [Handbook of African Languages]).
- Dubois, J. et al. (1973). *Dictionnaire de Linguistique*. Paris: Librairie Larousse.
- Emanation, M. (1992). Chagga 'come' and 'go' metaphor and the development of tense–aspect. *Studies in Language*, 16, 1–33.
- Fagan, B. M. (1984). The Zambezi and Limpopo Basins: 1100–1500. In D. T. Niane (Ed.), *Africa from the Twelfth to the Sixteenth Century*, (pp. 525–550). Paris: UNESCO and Heinemann.
- Guthrie, M. (1948). *The Classification of the Bantu Languages*. London: International African Institute.
- — —. (1967–71). *Comparative Bantu*. 4 Volumes. London: Gregg.
- Heine, B., Claudi, U., & Hunnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago, IL: The University of Chicago Press.
- Hyman, L., & Watters, J. R. (1984). Auxiliary focus. *Studies in African Linguistics*, 15, 233–73.
- Janson, T. (1991/92). Southern Bantu and Makua. *SUGIA*, 12/13, 63–106.
- Louw, F., & Finlayson, R. (1990). Southern Bantu origins as represented by Xhosa and Tswana. *South African Journal of African Languages*, 10 (4), 401–410.

- Nurse, D. (2003). Aspect and tense in Bantu languages. In D. Nurse & G. Philippson (Eds.), *The Bantu Languages*. London: Curson.
- Sweetser, E. (1988). Grammaticalization and semantic bleaching. *Berkeley Linguistic Society*, 14, 389–405.

The marking of directional deixis in Somali*

How typologically idiosyncratic is it?

Philippe Bourdin

Université Paris X and York University (Toronto)

Standard accounts of directional deixis fail to explain adequately the attested uses of ventive *soo* and itive *sii* in Somali. This is because their semantics involves a dual architecture: exophoric anchoring (is the goal of motion the deictic center?) and endophoric anchoring (is the moving entity the primary participant in the associated non-motional process?). *Soo* signals identification in both domains, while *sii* encodes otherness in one of the two. Beyond explaining several facts specific to Somali, the proposed account is typologically relevant on three counts. With respect to endophoric anchoring, in particular, the behavior of *soo* and *sii* exhibits both convergence with, and divergence from, that of ‘come’ and ‘go’ in languages as diverse as Korean, Tongan and Mohave.

1. Introduction

Most reference grammars, including those dealing with well-known languages, give relatively short shrift to the semantics of *directional deictics* – by which term are standardly, though not uncontroversially, designated ‘come’-type and ‘go’-type directionals, i.e. *ventive* and *itive* markers respectively. If they go into any detail at all, they seldom do much more than specify whether the ventive verb (or affix, particle. . .) may refer to motion towards the speaker or addressee, as in English, or strictly to motion towards the speaker, as in Spanish:

- (1) [Situational context: someone’s knocking at the door.]

I’m coming!

**I’m going!*

- (2) [Same situational context]

¡Ya voy!

**¡Ya vengo!*

They may also, occasionally, provide some clarification as to the type of *associated motion* (Koch 1984; Tunbridge 1988) denoted by the directional – i.e. whether the

motional event is preparatory to the main event described by the verb, as in Wolof, concomitant with it, as in Datog, or subsequent to it, as in Pero:¹

- (3) [Wolof (Niger-Congo, Atlantic; Senegal and Mauritania); Samb (1983:43)]
muucu-si-'l yax bi
 suck-VEN-IMPER.2 bone DEF
 'Come (here) to suck on the bone!'
- (4) [Datog (Nilo-Saharan, Nilotic; Tanzania); Rottland (1982:187)]
bwa:f-a:n-an
 dig-VEN-INSTRUMENTAL
 'Dig (with an object) while coming here.'
- (5) [Pero (Afro-Asiatic, Chadic; Nigeria); Frajzyngier (1989:94)]
cúg-ínà tù púccù
 fall-COMPLETIVE:VEN PREPOSITION over.there
 'He fell over there and then he came here.'

Such relative paucity of cross-linguistic information sets stringent limitations on the kind of typological investigation that may currently be carried out in the area of directional deixis, and it accounts for the largely pioneering nature of such important work as Ricca (1993) or Wilkins and Hill (1995). Against this background, my goal will be to begin taking apart the particular system of Somali directional deictics so that it then becomes possible to compare other systems to it, rather than using as standard benchmarks such "usual suspects" as French, German, Spanish or English.

Superficially, and in morphosyntactic terms, the Somali markers of directional deixis look very much like their putative German counterparts. On the one hand, the verbs *imaw* and *tag* are standardly glossed 'kommen' and 'gehen', respectively:

- (6) a. *wuxu ka yimid Cadan*
 FOCUS:3MASC.SG from come:PAST:3MASC.SG Aden
 'He came from Aden.'
- b. *wuxu tag-ay Cadan*
 FOCUS:3MASC.SG go-PAST:3MASC.SG Aden
 'He went to Aden.'

On the other hand, the preverbal clitics *soo* and *sii* are reminiscent of 'her' and 'hin':²

- (7) a. *wuu soo soc-daa*
 FOCUS:3MASC.SG VEN walk-PRES:3MASC.SG
 'He's coming towards me.'
 Or: 'He's coming towards you.'
- b. *wuu sii soc-daa*
 FOCUS:3MASC.SG ITIVE walk-PRES:3MASC.SG
 'He's going (towards a location distinct from where I am and from where you are).'

To be sure, *imaw* can sometimes alternate with motion verb + *soo*:

- (8) a. *halkan buu yimid shalay*
 here FOCUS:3MASC.SG come:PAST:3MASC.SG yesterday
 'He came over here yesterday.'
- b. *halkan buu soo mar-ay shalay*
 here FOCUS:3MASC.SG VEN drop.in-PAST:3MASC.SG yesterday
 'He came over here yesterday.'

However, there is only limited denotational overlap between *imaw* and *soo*, and even less between *tag* and *sii*. There are at least two reasons for this.

The verb *imaw* is only sporadically deictic. Thus, somewhat like German *kommen*, it routinely focuses on the endpoint of motion without any sort of implication as to the speaker's or addressee's location:

- (9) *afar-tii-buu yimid Cadan*
 four-DEF-FOCUS:3MASC.SG come/arrive:PAST:3MASC.SG Aden
 'He arrived in Aden at 4 o'clock.'

If anything, the verb *tag* is even more weakly deictic. Besides, it may either refer to departure or to arrival, so that (10), for instance, may be an answer to 'Has he left?' or to 'Do you think he has arrived in Aden by now?':

- (10) *afar-tii-buu tag-ay Cadan*
 four-DEF-FOCUS:3MASC.SG go-PAST:3MASC.SG Aden
 (i) 'He left for Aden at 4 o'clock.'
 (ii) 'He arrived in Aden at 4 o'clock.'³

In terms of *Aktionsart*, there seems to be a punctual or telic quality to *imaw*, which places some constraints on its ability to describe motion in progress. The instantiation of the so-called "Moving-Time" metaphor provides telling evidence of this:

- (11) a. *waaya-dii teg-ey*
 time-DEF:REMOTE go-PAST:3
 '[in] times gone by'
- b. *waayo-oyin-ka soo socda*
 time-PL-DEF:NON.REMOTE VEN walk:PRES:3
 '[in] times that are coming'
- c. **waayo-oyin-ka imanayaa*
 time-PL-DEF come:PRES.PROG:3

While it is unproblematic to refer to past time by means of the verb *tag* (or *teg*), as in (11a), it is not possible to refer to future time by means of the verb *imaw*; instead, the ventive clitic *soo* must be used, as in (11b), in conjunction with the verb *soco*, which refers to locomotion in general. This is because while past time is conceptualized as time which has finished moving, future time is thought of as an entity which is currently moving towards the speaker as she speaks; and it so happens that *imaw* has

an *Aktionsart* profile which seems to preclude, or at the very least constrain, reference to motion currently under way.⁴

The semantics of the verbs *imaw* and *tag* are only peripheral to the thrust of this paper, as is the limited degree to which they alternate with *soo* and *sii* plus verbs of locomotion. Although *imaw* and *tag* certainly contribute to the encoding of directional deixis, they just do not play the central role that Somali assigns to the clitics *soo* and *sii* in that respect – a state of affairs which is not unlike what may be observed in German, where the preverbs *her* and *hin* are more consistently and forcefully deictic than the verb *kommen*, let alone *gehen* (Vuillaume 1983; Di Meola 1994).

Soo and *sii* will be the main focus of this paper. First, the seemingly idiosyncratic uses will need to be delineated from uses that appear, at least superficially, to be far more typical cross-linguistically. This will lead to the formulation of a hypothesis, or rather the elements of one, regarding the logic governing the semantics of both markers. Finally, I will look at how this logic fits in with some selected features of the behavior of ventive and itive markers across a number of languages; in so doing, I will be taking a fresh look at directional deixis in general and will be working my way towards a redefinition of what it means, cross-linguistically, for a marker to be *ventive* or to be *itive*.

2. The seeming atypicality of *soo* and *sii* as directional deictics

Ventive markers are standardly defined as referring, prototypically, to motion towards the deictic center, while itive markers refer to motion towards some other location. As shown by Fillmore (1975) in his seminal analysis of *come* and *go* in English, the selection of the deictic center is sensitive to a number of factors, which actually vary from language to language. For the sake of simplicity, and unless otherwise specified, the deictic center will be assumed – in this section of the paper – to be the speaker's spatial location at speech time.

Intuitively, the behavior of *soo* and *sii* would seem to fall under two broad types: the expected and the unexpected. Examples like (7a, b) above fall under the expected type, as do (12a, b):

- (12) a. *subaxkasta gurig-ii Cabdi buu soo*
 every.morning house-DEF Abdi FOCUS:3MASC.SG VEN
mara
 drop.in:PRES:3MASC.SG
 'Every morning, he drops in on Abdi (on his way to work where I am).'
- b. *subaxkasta gurig-ii Cabdi buu sii*
 every.morning house-DEF Abdi FOCUS:3MASC.SG ITIVE
mara
 drop.in:PRES:3MASC.SG

‘Every morning, he drops in on Abdi (on his way to work where I am not).’

However, *soo* and *sii* also behave in ways that seem completely unexpected. A case in point is sentences (13a) and (13b), which lend themselves to the readings indicated below, amongst others:

- (13) a. *soo seexo*
 VEN sleep:IMPER.2SG
 ‘Go have a rest (and then come back here).’
 b. *sii seexo*
 ITIVE sleep:IMPER.2SG
 ‘Have a rest (until I come back).’

The most striking feature of (13a) is that by using ventive *soo* the speaker invites her addressee, not to move towards her, but actually to move away from her! Clearly, the deictic logic underlying the use of *soo* in (12a) is no longer operative here. On the face of it, the semantics of itive *sii* in (13b) is also typologically idiosyncratic. Thus, although the sentence is in the imperative, the moving entity – i.e. the *figure* in Talmy (1985)’s terminology – is the speaker rather than the addressee: in other words, it is not coreferential with the grammatical subject of the verb.

The semantics of *soo* and *sii* in the following sentences is just as dissonant from the “standard” behavior of ventive and itive markers across languages:

- (14) a. *aad baan u soo cun-i*
 much FOCUS:1SG to/for[GOVERNS *aad*] VEN eat-INFIN
 ‘I’m going to go stuff my face (while you’re waiting for me).’
 b. *aad baan u sii cun-i*
 much FOCUS:1SG to/for[GOVERNS *aad*] ITIVE eat-INFIN
 ‘I’m going to stuff my face (while you’re out).’

Just as they did in (13a) and (13b), *soo* and *sii* here entail a roundtrip motion, with an outgoing segment which is thither-oriented and a return segment which is hither-oriented: (14a) implies that the speaker is on his way out, but is going to come back; and (14b) that the addressee is on his way out, but is going to come back. In other words, the roundtrip, in all four sentences above, has the deictic center as its eventual destination. Crucially, this is true irrespective of whether the clitic is *soo* or *sii*, which implies that it is *not* the deictic orientation of the trajectory which accounts, here, for the semantic contrast between the two markers.

How, then, is this contrast to be explained?

It was observed above, *à propos* (13b), that the identity of the moving entity was not in accord with typological expectations. Neither is it in (14b). Indeed, whether it is even consonant with what (13b) might lead us to extrapolate depends entirely on the conceptual basis chosen for the extrapolation. For instance, a logic based on the

Speaker/Addressee pattern turns out to be completely inoperative: *sii* in (13b) refers to motion by the speaker, and in (14b) to motion by the addressee.

Far more promising is a logic based on the Same-Subject/Different-Subject pattern. While in both (13a) and (14a), with *soo*, the figure is the individual denoted by the grammatical subject, in (13b) and (14b), with *sii*, the figure is another individual – namely, by default, the protagonist in the speech situation who is not coreferential with the subject. It is readily apparent that this “disconnect” between grammatical subject and figure is key to an understanding of *sii*; and that invoking sameness between subject and figure is just as crucial to accounting for the semantics of *soo*.

3. The underlying logic: Towards a working hypothesis

Fundamentally, *soo* and *sii* refer to a motional event (M) which takes place in physical space. Syntactically, they modify a verb which itself refers to a process (P). As will be seen below, the relationship between P and M may be one of identity, coincidence or contiguity. There is identity if the process designated by the verb is M itself. There is coincidence if P and M are concomitant, and there is contiguity if M occurs prior to P or after it, or if it frames P. For the sake of expository clarity, *soo* and *sii*, at this stage, are best tackled separately. The evidence gathered so far suggests that the core, non-metaphorical uses of *soo* are governed by two distinct but conjoined constraints.

One is a *Space-Deictic constraint*: *soo* refers to a motional event which has the deictic center as its ultimate goal. The other is a *Same-Subject constraint*: the figure involved in M is also the primary participant in P, i.e. the process denoted by the verb which *soo* modifies. This terminological choice is quite deliberate: the “Same-Subject” label is being chosen *because* it is the label that happens to be used by typologists in connection with the mechanism of switch reference (see §4, below).

The Space-Deictic constraint and the Same-Subject constraint constitute the boundaries within which the precise semantics of *soo* is going to play out in a given context. How exactly it is going to play out depends on a number of variables, which include – but are not limited to – the semantics of the verb, the person of the subject, the tense and mood specifications of the verb, as well as the medium of communication, e.g. face-to-face communication vs. phone conversation.

Example (13a), repeated here as (15), gives a flavor of this range of interpretative variation:

- (15) [Speaker = Anab / Addressee = Abdi]

soo *seexo*

VEN sleep:IMPER.2SG

(i) ‘Come sleep over here!’ [face-to-face communication]

(ii) ‘Go get some rest (and then we’ll get together)!’ [face-to-face communication]

(iii) ‘Get some rest (and then we’ll get together)!’ [phone conversation]

- (iv) ‘Get some rest (while you’re traveling towards here)!’ [phone conversation]

Interpretation (i) presupposes a scenario whereby Anab, the (female) speaker, and Abdi, the (male) addressee, are both in a fairly large room and Anab invites Abdi to move over to the sofa near which she is sitting: the motional event to which *soo* refers here is preparatory to the process of sleeping, and it is of course of the hither-oriented variety.

Interpretation (ii) involves the complex trajectory that was described above in connection with (13a) and (14a). It is made up of an “outgoing”, and thither-oriented, segment which is preparatory to P and of a hither-oriented segment which will be subsequent to P. What motivates the use of *soo* is the deictic orientation of the *subsequent* segment, i.e. the fact that speaker and addressee will end up in the same spot, which may be Anab’s present location or some other place where she will then happen to be. The outgoing segment is merely an automatic consequence of the mechanics of face-to-face communication: in order for Abdi to reunite with Anab subsequently to P, it must be the case that P will take place away from their current location; this, in turn, entails an outgoing trip which will take Abdi away from their present location.

That the outgoing segment is a consequence of real-world constraints and *not* a part of *soo*’s intrinsic meaning is not a hypothesis which is being put forward because it conveniently jibes with the Space-Deictic Constraint. There is, in fact, empirical evidence for it, which is none other than Interpretation (iii): if Anab and Abdi are talking on the phone, the real-world constraints no longer require an outgoing trip and indeed the only segment of the trajectory that remains is the one that will follow P and will end up at the deictic center, defined as Anab’s location. Clearly, it is this hither-oriented segment that *soo* properly encodes.

A phone conversation scenario, however, allows for a fourth interpretation, whereby Abdi is on his way to Anab’s location and Anab is suggesting that he should get some rest in the course of the trip. One interesting feature of interpretation (iv) is that the motional event, this time, is concomitant with P.

To recap, all four interpretations follow from the Space-Deictic constraint in tandem with the Same-Subject constraint. When it comes to the motional event encoded by *soo*, it is preparatory to P (Reading (i)), subsequent to P (Readings (ii) and (iii)), or concomitant with P (Reading (iv)).⁵

In order to better understand the semantics of *soo*, it may be fruitful to subject sentence (15) to a number of alterations. Inserting the adverb *halkan*, as in (16), has the effect of precluding all interpretations but (i):

- (16) *halkan soo seexo*
 here VEN sleep:IMPER.2SG
 ‘Come sleep over here!’

Presumably, the spatial immediacy inherent in the deictic *halkan* is incompatible with the time-lag implied by Interpretations (ii) to (iv). Substituting other verbs, which

refer to other types of processes, is also likely to disambiguate the contribution of *soo* to the meaning of the sentence. Thus, each of the following sentences admits of one reading only:

- (17) a. *soo fadhiso*
 VEN sit.down:IMPER.2SG
 ‘Come here and stay quiet!’
 b. *soo joogso*
 VEN come.to.a.halt:IMPER.2SG
 ‘Stop fooling around and come here!’
- (18) a. *hilib soo iibi*
 meat VEN buy:IMPER.2SG
 ‘Go buy us some meat (and come back)!’
 b. *soo hayso*
 VEN possess/keep:IMPER.2SG
 ‘Keep the book (and then come back and give it back).’

While ‘sit down’ and ‘come to a halt’ involve much the same spatial-cum-temporal immediacy intrinsic to *halkan* in (16), such processes as ‘buying meat’ and ‘keeping a book’ can hardly be construed as following immediately upon a hither-oriented motion from the far corner of the room to where the speaker is located; neither would it make much sense to view such processes as taking place concomitantly with a trip towards the speaker’s house.

Changing the mood and tense of the verb may also have a drastic effect on the interpretation of *soo*. For instance, locating the event in the past has the effect of removing it from the spatio-temporal coordinates of the speech situation, thereby ruling out any “immediate” reading along the lines of Interpretation (i) of (15):

- (19) a. *waan soo seex-day*
 FOCUS:1SG VEN sleep-PAST:1SG
 (i) ‘I took a nap before coming here.’
 (ii) ‘I took a nap on my way here (on the bus).’
 b. *bas-kuu ku soo dhex seex-day*
 bus-DEF:3MASC.SG in VEN inside sleep-PAST:3MASC.SG
 ‘He slept on the bus on his way here.’

While (19a) allows for P and M to be interpreted as contiguous or concomitant, our knowledge of the uses of buses in the real world enhances decisively the plausibility of a concomitant reading of (19b).

Pragmatic considerations also play a crucial role in the interpretation of *soo* in (20):

- (20) *waan soo seexan*
 FOCUS:1SG VEN go.to.sleep:INFIN
 ‘I am off to have some rest (and I will be back).’

Owing to the prospective/intentional value borne by the infinitival form *seexan*, the sentence typically has the illocutionary force of a promise or a commitment. This, in turn, favors the scenario of a roundtrip involving the actualization of the process away from the deictic center followed by a hither-oriented return to the addressee's location.

Example (13b), repeated here as (21), provides a useful starting-point for an analysis of *sii*:

- (21) [Speaker = Anab / Addressee = Abdi]
sii seexo
 ITIVE sleep:IMPER.2SG
 (i) 'Get some sleep (while you're travelling away from here)!'
 [face-to-face comm.]
 (ii) 'Get some sleep while I'm gone!' [face-to-face comm.]

On the first reading, the individual involved in P, i.e. the process of getting some sleep, is going to carry it out – or, rather, experience its actualization – in the course of going to a place remote from the deictic center. On the second reading, somebody who is not the individual to whom the subject refers is going to be involved in a motional event which will frame P. In other words, in contrast to *soo*, which had to satisfy *both* the Same-Subject constraint and the Space-Deictic constraint, *sii* has got to satisfy *either* what might be called for short the *Space-Antideictic* constraint *or* the *Different-Subject* constraint.

Essentially, therefore, the semantics of *sii* is defined *disjunctively* and *negatively*. This entails that pragmatic factors are going to play a crucial role in determining, in particular, the identity of the “different subject”, i.e. the individual involved in the motional event M. Interpretation (ii) of (21) represents the default case, whereby the only candidate for different-subjecthood is one of the two participants in the dialogue. However, it is clear from Interpretation (ii) of (22) that the different subject may also refer, situational context permitting, to a third party:

- (22) *wuxu sii joog-aa London*
 FOCUS:3MASC.SG ITIVE stop/stand-PRES:3MASC.SG London
 (i) 'He is in London right now and will continue on his way
 (towards a place where I am not).'
 (ii) 'He is in London, where I or somebody else is going to join him.'

In each of the following examples, the semantic import of the verb rules out the applicability of the Space-Antideictic constraint, which guarantees that only the Different-Subject constraint is operative:

- (23) *sii fadhiso*
 ITIVE sit.down:IMPER.2SG
 'Stay where you are (. . . I'll meet with you later).' [phone conversation]
 (24) *waan sii yar seexan-e*
 FOCUS:1SG ITIVE a.little sleep:INFIN-and

'I'm going to take a nap (... wake me up when you get here).' [phone conversation]

- (25) *halkan ha sii fadhi-yo*
 here OPTATIVE.MARKER ITIVE sit-OPTATIVE:3MASC.SG
 'Let him stay here (until I come back or until X comes back).'

In (26) and (27), on the other hand, *sii* essentially serves as a marker of thither-oriented motion:

- (26) *wuu iga sii bax-ay*
 FOCUS:3MASC.SG 1SG.OBJECT:from ITIVE go.out-PAST:3MASC.SG
 'He left my house (... and went somewhere).'
- (27) [Saeed (1999:227)]
baxsad-koodii bay sii wad-een
 escape-POSS:3PL FOCUS:3PL ITIVE continue-PAST:3PL
 'They kept on running away.'

When the Space-Antideictic constraint is operative, M is concomitant with P, as in Interpretation (i) of (21), or subsequent to it, as in Interpretation (i) of (22). When it is the Different-Subject constraint which is operative, M typically frames P, as in Interpretation (ii) of (21), or else is subsequent to it, as in Interpretation (ii) of (22). The nature of the medium of communication is a significant factor in these interpretative choices.

If the hypotheses formulated above are to be tested, as they must, the behavior of *soo* and *sii* needs to be analyzed contrastively.

The following minimal pairs involve the verb *joog*, which refers to a static process:

- (28) a. *wuxu soo joog-ay afar cisho*
 FOCUS:3SMASC.SG VEN stay-PAST:3MASC.SG four days
 'He stayed (there) for four days (and then he came back).'
- b. *waxay sii joog-tay afar cisho*
 FOCUS:3FEM.SG ITIVE stay-PAST:3FEM.SG four days
 'She stayed (there) for four days (and then she continued her journey).'
- (29) a. **wuxu soo joog-aa London*
 FOCUS:3SMASC.SG VEN stay-PRES:3MASC.SG London
- b. *wuxu sii joog-aa London*
 FOCUS:3MASC.SG ITIVE stay-PRES:3MASC.SG London
 'He is in London, where I or somebody else is going to join him.'

The contrast between (28a) and (28b) can be accounted for very straightforwardly by appealing to the Space-Deictic and Space-Antideictic constraints, respectively. Why exactly these constraints are no longer operative in (29) is an interesting topic, which would require further investigation. More significant, for the purposes of the present paper, is the acceptability contrast between (a) and (b). As I established above, the use of *soo* is governed by a *conjunctive* logic: both the Space-Deictic and the Same-

Subject constraints must be satisfied. As a result, the inapplicability of either of these constraints renders the sentence unacceptable. The logic governing *sii*, on the other hand, is *disjunctive*: it is enough for the Different-Subject constraint to be potentially satisfied, as it is in (29b), for the sentence to be interpretable. These different logics explain why the use of *soo* is more severely restricted than that of *sii*.⁶ They also entail that sentences with *sii* will routinely lend themselves to more interpretations than those with *soo*:

- (30) a. *la' soo fadh-i*
 with VEN sit-IMPER.2SG
 'Go spend some time with her (and come back).' [face-to-face communication]
- b. *la' sii fadh-i*
 with ITIVE sit-IMPER.2SG
 (i) 'Stay with her (until I come back with the medication).' [face-to-face comm.]
 (ii) 'Stay with her (until I get there or somebody gets there).' [phone conv.]

In the following quasi-minimal pair, *soo* and *sii* modify, yet again, a verb describing a static posture:

- (31) a. *halkan uun buu soo fadhi-yi*
 here all.the.time FOCUS:3MASC.SG VEN sit-INFIN
 (lit.) 'He sits (there looking at me) all the time.'
 'He's on my back all the time.'
- b. *halkan buu sii fadhi-yi*
 here FOCUS:3MASC.SG ITIVE sit-INFIN
 '[The child] is going to stay here while you're away.'

Soo requires both the Space-Deictic and Same-Subject constraints to be satisfied: this is possible, in (31a), as long as the hither-oriented motion denoted by *soo* can be construed as involving, if not the whole individual, at the very least a fragment of his body; the eyes would seem to be the only available option, though at the cost of a metaphorically extended interpretation of *soo*. No such extended interpretation is required in (31b), as the Different-Subject constraint is readily applicable.

In (31b), *sii* triggers an implicature of temporariness – just as it did, for example, in (13b) and (14b) above. Evidently, this is because the motional event is a roundtrip. As a result, it frames the process denoted by the verb – 'sitting' in (31b), 'sleeping' in (13b) and 'eating' in (14b) – and thereby sets it within strict boundaries. In point of fact, the implicature of temporariness triggered by *sii* is to a large extent conventionalized and frequently foregrounded:

- (32) *sii hayso gadhi-ga!*
 ITIVE possess/keep car-DEF
 'Keep the car (for a while)!'

Soo itself can take on temporal significance, though in two different ways.

First, it may trigger an implicature of “recency” (Bourdin 2002) or temporariness. This is what happens in (34), with *soo*, in contrast to (33), without *soo*:

- (33) *Cabdi baan la fadhi-yey xalay*
 Abdi FOCUS:1SG with sit-PAST:1SG last.night
 ‘I was with Abdi last night.’
- (34) *Cabdi baan la soo fadhi-yey*
 Abdi FOCUS:1SG with VEN sit-PAST:1SG
 ‘I was with Abdi (... and now I’m back).’
 Hence: ‘I was with Abdi just now.’

Both sentences are in the past tense. However, while the past situation in (33) is divorced from the present, *soo* in (34) establishes a tight connection between the two. This is an automatic consequence of the mandatory relationship of *contiguity* between P (my being with Abdi) and M (my coming here).

In (35), both *soo* and *sii* trigger an implicature of temporariness, simply because they both refer to a roundtrip which frames the process of keeping the book:

- (35) a. *soo hayso*
 VEN possess/keep:IMPER.2SG
 ‘Keep the book (and then come back and give it back)!’
- b. *sii hayso*
 ITIVE possess/keep:IMPER.2SG
 ‘Keep the book (until I come back)!’

Likewise, and for much the same reasons, temporariness is very much a factor in (36) where the directional clitics occur sequentially, i.e. in adjoining sentences:

- (36) a. *si fiican u soo naso anigu-na waan sii*
 way good in VEN rest:IMPER:2SG I-and FOCUS:1SG ITIVE
nasan
 rest:INFIN
 ‘Go take a good rest and in the meantime I will get some rest as well.’
- b. *si fiican u soo naso anigu-na waan soo*
 way good in VEN rest:IMPER:2SG I-and FOCUS:1SG VEN
nasan
 rest:INFIN
 ‘Go take a good rest and in the meantime I will go get some rest as well.’

Examples (37) and (38) highlight the implicature of recency associated with *soo* in (38a):

- (37) *wuu iga bax-ay*
 FOCUS:3MASC.SG 1SG.OBJECT:from go.out-PAST:3MASC.SG

'He left my place (and went to do some errands...).'

Hence: 'He's not here any more.'

- (38) a. *wuu iga soo bax-ay*
 FOCUS:3MASC.SG 1SG.OBJECT:from VEN go.OUT-PAST:3MASC.SG
 'He has left my place (... and he's on his way to your place).'
 Hence: 'He left my place not long ago.'
- b. *wuu iga sii bax-ay*
 FOCUS:3MASC.SG 1SG.OBJECT:from ITIVE go.OUT-PAST:3MASC.SG
 'He left my house (... and went somewhere).'

The implicature is cancelled when *soo* is deleted, as in (37), or replaced by *sii*, as in (38b). The reason for this is transparent: specifying as it does that the trip has as its endpoint the addressee's house, rather than some undetermined destination, *soo* defines a strictly bounded interval and, by implication, sets severe limits on the temporal distance involved.

The type of contribution that *soo* can make to the temporal import of the sentence is not confined to notions of temporariness or recency:

- (39) a. *toban sanadood uun baan bariis cun-ayey*
 ten years only FOCUS.1SG rice eat-PAST.PROG:1SG
 'During the ten years [that I spent there] all I ate was rice.'
- b. *toban sanadood uun baan bariis soo cun-ayey*
 ten years only FOCUS.1SG rice VEN eat-PAST.PROG:1SG
 '[Enough is enough...] For ten years now all I have been eating is rice.'

Both sentences are in the past tense. However, while the absence of *soo* in (39a) entails that the process is firmly anchored in past time, its presence in (39b) establishes a close connection with present time; *soo* does more than that, though: it forces us to conceptualize the process as spanning in its entirety an interval which started ten years ago and which is still open at speech time. In other words, *soo* here serves as an exponent of *continuation* or *persistence* up to reference time, a function which is all the more noteworthy as Somali has no formal way of encoding the perfect. We are not dealing with an implicature as such, but rather with a very familiar shift from motion in space to motion in time, specifically a hither-oriented type of motion: see Lichtenberk (1991), for a fuller account of the metaphorical process at play. Further evidence is provided by examples (40) to (42):

- (40) a. *shan sanadood baan dhig-anay-ay af Somaaliga*
 five years FOC:1SG learn-PAST.PROG-1SG language Somali
 'I have been learning the Somali language for five years.'
- b. *shan sanadood baan soo dhig-anay-ay af Somaaliga*
 five years FOC:1SG VEN learn-PST.PRGR-1SG language Somali
 'I have been learning the Somali language for five years (which is why I speak it so well, or ... and yet I haven't made much progress, etc).'

- (41) *shan sanadood baan la soo nool-aa*
 five years FOCUS:1SG with VEN be.alive-PAST:1SG
 (i) ‘I lived with her for five years (...and now I’m back where I lived before).’
 (ii) ‘I lived with her for five years (...and now I have returned to my old self).’
 (iii) ‘I’ve been living with her for five years (...but I still can’t figure her out,
 or: ... and never have I felt any regret, etc.).’
- (42) *halgan-kii iyo jihaad-kii loo soo gal-ay*
 struggles-DEF and wars-DEF INDEFINITE.SUBJ VEN wage-PAST
ilaalin-ta iyo gudbin-ta din-ta
 preservation-DEF and transmission-DEF religion-DEF
 ‘The struggles and the wars [that] one has been waging [up to now] for the
 preservation and transmission of religion...’ [Barre, M. Ciise. 1995. *Siiradii Nabiga* [‘Light on the Koran’], Weston (Ont., Canada): Sisca Islamic Recordings and Bookstore, p. 1]

Soo serves as a continuative marker in (40b), while highlighting the impact of the state of affairs on the situation obtaining at the present time. The tightness of this connection, which may be of a causal or concessive nature, is lost when *soo* is absent, as in (40a). In (41), *soo* lends itself to three distinct readings, with only the third one involving continuation up to the present moment. Interpretation (i) is of a purely spatial nature: the individual involved in P – i.e. ‘living with her’ – has now returned to his home base. Interpretation (ii) involves a psychological return to home base. Interestingly, both these interpretations place P firmly in the past. In interpretation (iii), on the other hand, *soo*’s status as a continuative marker forces a reinterpretation of the past tense, which is now construed as equivalent to a perfect marker; much the same interpretation is triggered by *soo* in (42), which is drawn from a written text.

4. The semantics of *soo* and *sii* in typological perspective

The sections above have provided no more than a broad outline of the Somali system of directional deixis. Although it falls short of the thoroughgoing description that is conditional on much needed further research, this outline, along with the attempt at a principled account sketched in §3, allows for some preliminary observations of a typological nature. These fall under three categories. First, and most obviously, it is worth examining how the Somali system responds to some of the parameters of variation which seem to structure parallel systems in language after language. Second, given the well-documented susceptibility of systems of directional deixis to mechanisms of functional extension and full-fledged grammaticalization, the ability of *soo* to behave like an aspectual marker of continuation up to reference time invites comparison with similar phenomena in other languages. Third, the *Different-Subject* logic which appears to determine one of the two core meanings of *sii* has typological implications that need

to be explored, if only because it constitutes, on the face of it, a saliently idiosyncratic feature of directional deixis in Somali.

Cross-linguistically, markers of directional deixis are subject to several parameters of variation; some of those are, broadly speaking, of the switch-on/switch-off type. For instance, some languages (Spanish, Tahitian, Japanese...) typically equate the deictic center with the realm of the speaker alone, while in others (French, Italian, English...) the deictic center additionally encompasses the position of the addressee. As example (7a) above testifies, Somali firmly belongs in the second grouping of languages.

Some languages tend to encode the return portion of a roundtrip rather than the outgoing segment. Japanese is a case in point, as are Indo-Aryan languages like Marathi and Gujarati:

- (43) [Japanese; Martin (1975:536)]
sinbun o kat-te kuru
 newspaper ACCUSATIVE buy-CONVERB come:NON.PAST
 (lit.) 'I am coming having bought the newspaper.'
 Hence: 'I'm going to go buy the newspaper.'

Likewise, one of the most typical functions of *soo* in Somali is to signal that completion of the process will be followed by motion to the deictic center. Depending on the semantics of the verb and on the medium of communication, it may or may not be the case that a roundtrip is involved. The following sentence is exemplary of a one-way, albeit complex, trajectory:

- (44) *soo mar tukan-ka o kaalay*
 VEN go.to:IMPER.2SG shop-DEF and come:IMPER.2SG
 'Go to the shop and (then) come here!' (where going to the shop may or may not be on the normal route between the addressee's house and the speaker's)

The sentence occurs in a phone conversation, and the addressee is invited to carry out a two-phase motion, involving a trip to the shop followed by a trip to the speaker's home. Interestingly, the deictic orientation of the second segment of the trajectory is encoded twice: by the clitic *soo* in the first conjunct and then by *kaalay*, the imperative form of the verb *imaw*, in the second conjunct. Although *soo* is a syntactic modifier of the verb *mar*, it has absolutely no bearing on the deictic orientation of the motion that *mar* as such refers to. Indeed, it may well be the case that the addressee lives two miles to the west of the speaker's house and that the shop is located five miles to the west of his own home. What counts is that the *final* segment of the trip has the deictic center as its goal.

This particular entailment of *soo* is highly significant, once placed in the context of recent discussion on the universality of the 'come'/'go' dichotomy. In support of their claim that there is no such thing as a prototypical 'come'-type verb across languages, Wilkins and Hill (1995:224–226) make much of the fact that in some languages, such as English, the putative 'come'-type verb indicates that the deictic center is actually

reached while in others, such as Mparntwe Arrernte, it merely signals motion in the general direction of the deictic center:

- (45) [Mparntwe Arrernte (Australian); Wilkins and Hill (1995:225)]
re petye-ke ure-werne
 3SG:SUBJECT come-PAST fire-ALLATIVE
 ‘She moved to(wards) the fire (thereby getting closer to the deictic center).’

As is shown by (44), Somali does much “worse” than letting *soo* specify motion ending short of the deictic center: it actually allows it to modify a verb referring to motion in a direction *opposite* the deictic center! This, however, is merely a consequence of *soo*’s ability to *anticipate*, cataphorically so to speak, the deictic orientation of the *subsequent* trip.⁷ It does not in any way invalidate the fact that *soo* is fundamentally and inherently ventive, specifying as it does (the ultimate occurrence of) a hither-oriented motion by the referent of the grammatical subject.

The ability of *soo* to function as a marker of *continuation* up to reference time happens to instantiate a pathway of grammaticalization which is notoriously well-traveled by ventive markers across languages (Lichtenberk 1991; Bourdin 1999b). The following examples represent a small sample of the languages that may be adduced to illustrate this particular trend:

- (46) [Tokelauan (Austronesian, Oceanic; Tokelau and New Zealand); Hooper (2000)]
e pā mai lava ki nā aho nei
 TENSE/ASPECT arrive hither INTENSIFIER towards DET.PL day now
tēia tāofi
 DEMONSTRATIVE belief
 ‘That belief continues up to the present day.’
- (47) [Thai (Daic, Tai); Howard (2000:393)]
salúpléew thamnaan ma kī pii léew há
 in total work come how many years ASPECT HONORIFIC
 ‘In total, how many years have you been working here?’
- (48) [Kiowa (Kiowa Tanoan; Oklahoma); Watkins (1984:178–180)]
ó·dé à- t^hó· -á·
 starting.from.there AGENT.2SG:OBJECT.SG- drink -DURATIVE(VEN)
-hél
 -QUOTATIVE
 ‘They say you’ve been drinking for a long time.’
- (49) [Malayalam (Dravidian; India); Asher and Kumari (1997:298)]
ṇaan etra kaalam-aayi ninne
 I how.much time-ADVERBIAL.SUFFIX you:ACCUSATIVE
nookk-ip-poor-unnu
 look.after-PARTICIPLE-COME-PRES

'How long I have been looking after you!'

- (50) [Uzbek (Turkic); Boeschoten (1998: 365)]

ikki yıl işl-âp kel-di-m
two year work-CONVERB come-PAST-1SG
'I have been working for two years.'

- (51) [Spanish; Amador (1985: 155)]

vengo desde hace día-s observ-ando-lo
come:PRES:1SG since it makes day-PL observe-GERUND-OBJECT.3SG
atentamente
attentively
'I have been observing it carefully for days.'

The 'come'-type markers involved in (46) to (51) are morphosyntactically very diverse, ranging from a particle in (46) to a serial verb in (47) and to different varieties of verbal auxiliary in (49) to (51). This point is worthy of emphasis for two reasons. First, it is a manifestation of the formal diversity which is a defining property of directional deixis as a cross-linguistic category:⁸ in this respect, I depart radically from Wilkins and Hill (1995), whose reductionist focus on verbs is inconsistent, in my view, with the inescapable reality that a significant portion of the functional load borne by the verbs *venire* and *andare* in Italian or *kuru* and *iku* in Japanese is entrusted to non-verbal items in a great many languages of which Somali and German, along with Wolof, Datog and Pero (see examples (3) to (5) above), constitute a very small sample. Second, the data in (46) to (51), along with the Somali examples (39) to (42), provide conclusive evidence, if any were needed, that grammaticalization along a particular pathway is notionally driven and largely, if not wholly, impervious to any kind of morphosyntactic constraints; it is safe to assume, in other words, that *soo* has come to specify continuation of the process up to reference time *for the sole reason* that like *mai* in Oceanic languages, *ma* in Thai, or reflexes of *kel/gel* in Turkic languages, it is fundamentally a *ventive* marker. Indeed, the very fact that those various items, irrespective of their formal status or their particular position along the lexical-grammatical gradient, exhibit a singular affinity with the same well-individuated grammaticalization pathways adds to the body of evidence that we are dealing with language-specific instantiations of a *bona fide* cross-linguistic category.⁹

That appropriateness conditions for the use of *sii* could include the *Different-Subject* constraint is just as noteworthy, typologically, as *soo*'s ability to function as a perfect-cum-continuative marker. They are so, however, for seemingly opposite reasons: whereas the mutation of ventive markers into aspectual exponents of continuation up to reference time is widespread cross-linguistically, the interplay of directional deixis and different-subjecthood is, as far as can be ascertained, a typological rarity.

To be sure, there are languages that do use directional deixis to denote motion involving an entity or individual distinct from the most salient participant in the pro-

cess, i.e. typically the referent of the grammatical subject. The following example, in Tzotzil, is a case in point:

- (52) [Tzotzil (Mayan; Mexico); Haviland (1993:43)]
t'om *tal* *volkan*
 COMPLETIVE:3ABSOLUTIVE:erupt come volcano
 'The volcano erupted (and the ashes came).'

As Haviland (*ibid.*) points out,

the motion implied by Tzotzil directionals is a matter of inference and not of syntax (...). The directional depends on the overall scene evoked by the verb and the rest of the context of situation to supply a Mover.

In other languages, the identity of the figure (Haviland's "Mover") would seem to be determined by syntactic rules:

- (53) [Jacaltec (Mayan; Guatemala); Craig (1979:36; 38)]
 a. *ch-ø-ay-toj* *naj*
 ASPECT-3MASC.SG-downwards-ITIVE he
 'He goes down in a direction opposite to where I am.'
 b. *xcin haten-ic-toj* *yul carro*
 you move.me:PAST-in-ITIVE into truck
 'You pushed me into the truck (away from you).'

While in intransitive sentences, such as (53a), the deictic center is identified with the speaker's location, Craig observes that in transitive sentences, such as (53b), it is identified with the referent of the grammatical subject; as a result, the figure in (53b) is the referent of the grammatical object. The same mechanism is attested in Arbore (Hayward 1984:310–11), a language belonging, like Somali, to the Omo-Tana branch of East Cushitic.

Jacaltec and Arbore, therefore, give grammatical sanction to a dual system of deictic anchoring. At the *exophoric* level of the speech event, the anchor, or "organizing principle", is the speaker (or the addressee); at the *endophoric* level of the sentence, namely of a predication involving a process, the anchor is the primary participant in the process – typically the referent of the grammatical subject.¹⁰ Whether it is permissible to classify the endophoric mode of anchoring under the broad heading of *deictic* anchoring is far more than a terminological issue: deixis, after all, is standardly understood to involve *solely* the coordinates of the speech event, which would *ipso facto* exclude from its scope any kind of *endophoric*, i.e. intra-sentential or intra-textual, anchoring.

Indeed, the duality of directional anchoring of which ventive and itive markers are capable in Jacaltec and Arbore is formally reflected in Hausa by a morphosyntactic duality of encoding:

- (54) [Hausa (Afro-Asiatic, Chadic; Nigeria, Niger, Chad); Abdoulaye (1996: 124)]

taa aikoo dà yaaròo
 3FEM.SG:PERFECTIVE send:VEN EFFERENTIAL boy
yà amsam ma-tà gooRòo
 3MASC.SG:SUBJUNCTIVE take BENEFACTIVE-2FEM.SG kola nut
 'She sent a boy over to take her kola nut.'

While the ventive inflection on the verb anchors the trajectory followed by the boy, exophorically, to the speaker's location, the so-called "efferential" particle anchors it, endophorically, to the location of the individual to whom *taa* refers. However, Hausa seems to be relatively idiosyncratic in encoding exophoric and endophoric anchoring by means of different morphosyntactic categories of markers (a formal duality which, incidentally, renders possible, in (54), the co-occurrence of ventive and efferential marking on the same verb¹¹). In language after language, it is markers drawn from the very same morphosyntactic set that fulfill this kind of double duty:

- (55) [Tongan (Austronesian, Oceanic; Tonga); Tchekhoff (1990: 108)]

na'à ne kole mai meia Pita 'ae huó
 PAST he ask.for VEN of Peter DET hoe
 'He asked Peter for his hoe in order to give it to me (where 'I' is the exophoric anchor).'
 'He asked Peter for his hoe in order to use it himself (where 'he' is the endophoric anchor).'

- (56) [Korean; Jo (1990: 273)]

a. *John-i ku khad-lul caki-ccokulo*
 John-NOMINATIVE DEF card-ACCUSATIVE REFLEXIVE-towards
kkul-e o-ass-ta
 pull-SUFFIX come-PAST-DECLARATIVE
 'John pulled the card towards himself (i.e., towards the endophoric anchor).'

b. *John-i ku khad-lul caki-ccokulo*
 John-NOMINATIVE DEF card-ACCUSATIVE REFLEXIVE-towards
kkul-e ka-ass-ta
 pull-SUFFIX go-PAST-DECLARATIVE
 'John pulled the card towards himself (i.e., away from me as exophoric anchor).'

- (57) [German; Vuillaume (1983: 210)]

a. *zieh die Kiste her-auf*
 pull:IMPER.2 DEF box VEN-up
 'Pull up the box upwards (towards where you are as endophoric anchor)!'

b. *zieh die Kiste hin-auf*
 pull:IMPER.2 DEF box ITIVE-up
 'Pull the box upwards (away from me as exophoric anchor)!'

There is, in fact, substantial evidence that languages are not averse to collapsing exophoric and endophoric anchoring under a unique formal system, a state of affairs which, pending further theoretical and empirical elaboration, lends legitimacy to the claim that directional deixis is best viewed as a unitary and inclusive category, encompassing both varieties of anchoring.

Somali is different. While the mechanism governing the identification of the figure does incorporate elements of the Tzotzil and Jacaltec logics, it seems to be, if the hypothesis put forward above is correct, fundamentally *sui generis* when compared to the systems of directional deixis for which sufficiently detailed descriptions are available.

As in Jacaltec, the semantic content of both *soo* and *sii* involves a dual architecture: exophoric anchoring (via the Space-Deictic and Space-Antideictic constraints) and endophoric anchoring (via the Same-Subject and Different-Subject constraints). And much as in Tzotzil, the identification of the moving figure (when it has to be non-co-referential with the grammatical subject) is pragmatically motivated, rather than syntactically determined; thus, while speaker and addressee are usually the prime candidates for figurehood in utterances drawn from dialogues, such examples as (22b), (25), (29b) and (30b) above demonstrate that there are other possibilities, depending on discursive and/or situational context.

The exceptionality of Somali has to do with its peculiar construal of endophoric anchoring, and specifically with the Different-Subject constraint governing the use of *sii*. The Jacaltec, Hausa, Tongan, Korean and German data above do not make any sense unless both the exophoric and endophoric anchors are understood to be, for the purposes of directional orientation, exclusively *locations* in physical space. The exophoric anchor in Somali has exactly the same status. Not so, however, the endophoric anchor, which can be essentially defined as the *passive* participant in the motion event: namely somebody *other* than the moving entity, but whose location is *identified* with the endpoint of the trajectory.

In other words, *sii* signals two markedly distinct types of “otherness”: exophorically, motion, by the referent of the grammatical subject, *to* a location other than the speaker’s or the addressee’s; endophorically, motion, *by* an individual other than the referent of the grammatical subject, *to* the location where that referent happens to be. This lack of parallelism between the two modes of anchoring is tellingly illustrated by the contrast between (12b) and (29b), repeated here as (58) and (59) respectively:

- (58) *subaxkasta gurig-ii Cabdi buu sii*
 every.morning house-DEF Abdi FOCUS:3MASC.SG ITIVE
mara
 drop.in:PRES:3MASC.SG
 ‘Every morning, he drops in on Abdi (on his way to work where I am not).’
- (59) *wuxu sii joog-aa London*
 FOCUS:3MASC.SG ITIVE stay-PRES:3MASC.SG London
 ‘He is in London, where I or somebody else is going to join him.’

In (58), exophoric anchoring – specifically, the Space Anti-Deictic constraint – requires motion to end up in a place distinct from where the anchor – here the speaker – is located. In (59), endophoric anchoring – specifically, the Different-Subject constraint – places two inseparable requirements on the motion event: it has to be performed by somebody distinct from the anchor (‘he’) and its trajectory has to end up where the anchor is located (London). What this entails is that (a) *sii* is fundamentally a marker of “otherness” and (b) no more than one form of otherness can be enforced at any one time. Thus, *sii* in (59) is radically and inherently unable, it would seem, to signal both that the figure is *not* the individual who is in London and that whoever is moving has as her goal a destination *other* than London. Clearly, the system disallows what might be termed “otherness overload” – a design feature which might arguably provide a principled justification for the disjunctive semantics of *sii*: when the Space Anti-Deictic constraint is operative, there must be otherness *solely* in terms of distance between the endpoint of motion and the (exophoric) anchor; when the Different-Subject constraint is applicable, there must be otherness *solely* in terms of non-co-referentiality between the moving figure and the (endophoric) anchor.

The Different-Subject constraint which I hypothesize is one of the two semantic rules governing the use of *sii* is irresistibly reminiscent of *switch reference*, a reference-tracking mechanism (Foley and Van Valin 1984:321–367; Comrie 1989) whereby certain languages, in particular of Australia and North America, use a Same-Subject marker to indicate that the subordinate clause subject is co-referential with the main clause subject, and a Different-Subject marker to indicate that they are not co-referential. Indeed, the hypothesis I have articulated above in order to account for the semantics of *sii* would be lent some typological plausibility if it could be established that ventive and/or itive markers are involved, in one way or another, in the switch reference systems of at least some languages. Overlap of encoding between directional deixis and switch reference does happen to be attested in at least two languages that are completely unrelated genetically and geographically:

- (60) [Seme (Niger-Congo, Kru; Burkina Faso); Prost (1964:376–377)]

a gyarin yè k' a ko
 it difficult:PAST they DIFFERENT.SUBJECT[<ITIVE] it do
 ‘It was difficult, they did it.’

- (61) [Mohave (Hokan, Yuman; California); Munro (1976:39)]

- a. *ʔ-su:ma:-k manʔ ʔ-iyu:-k*
 I-dream-SAME.SUBJECT[<VEN] you I-see-PRES
 ‘While I dream, I can see you.’
 b. *isva:r-m i:ma-k*
 sing-DIFFERENT.SUBJECT[<ITIVE] dance-PRES
 ‘Although he sings, she dances.’

While Seme encodes Different-Subjecthood by means of a marker which is transparently related to the itive morpheme *ke*,¹² in Mohave both the Same-Subject and

Different-Subject suffixes are derived from directional deictics – i.e. a ventive and an itive marker, respectively. According to M. B. Kendall (1975: 9), there is a straightforward cognitive motivation for the particular functional extension exemplified by Mohave *-k* and *-m* and cognate suffixes in Yavapai, a kindred Yuman language.

As long as he [the speaker] stays on the same topic or talks about the same subject he stays in one “conceptual location.” But if he switches topics or subjects, he is metaphorically moving away from the point at which he was located. The notions of *location at* and *moving away from* are certainly parts of the definitions of */-k/* and */-m/* respectively.

In other words, rather than encoding motion to a *physical location* other than the *exophoric* center, the itive marker here is encoding motion, within the *space of discourse*, to an anchor other than the *endophoric* center. So far as the respective itive markers are concerned, Kendall’s analysis could be straightforwardly conceptualized in terms of *otherness*. Just like Somali *sii*, Seme *k’* and Mohave *-m* have the ability to signal exophoric otherness, which has to do with the specification of the spatial anchor serving as endpoint, or else endophoric otherness, which has to do with the specification of the participant in the process. However, while a given occurrence of Seme *k’* or Mohave *-m* falls under the scope of *either* directional deixis *or* reference-tracking, Somali *sii* has the singular ability to instantiate both domains at the same time: this is because (a) *sii* never ceases to signal motion in space and (b) the prohibition of “otherness overload” guarantees that when otherness of participant applies, otherness of spatial endpoint cannot apply.

The partial semantic and functional convergence between *sii* and the Seme and Mohave Different-Subject-cum-itive markers raises the non-trivial possibility that even with respect to the Different-Subject constraint, directional deixis in Somali, while no doubt typologically original, is not quite as idiosyncratic as might appear at first blush.

5. Conclusion

As Mous (2001) has shown, the directional deictics *soo* and *sii* of Somali, like their counterparts in other East Cushitic languages (e.g. Arbore, Dahalo and Burunge), are part and parcel of the “selector”, a cluster of preverbal markers which is a hallmark of many Cushitic languages and which may include also, depending on the language, morphemes pertaining to subject, focus, sentence-type, case, and tense-aspect-mood marking. This speaks to the fact that *soo* and *sii* are highly integrated into the grammatical fabric of Somali, a state of affairs which, far from being instantiated only in such other languages as Wolof or German, is an endemic trend among languages of the world – which trend itself contributes significantly to validating *directional deixis* as a cross-linguistic category.

It also speaks to the fact that when the functional range of *soo* extends to the encoding of the continuative/persistent perfect, as I have shown it does, the type of functional extension being instantiated is that which takes an already grammatical, rather than lexical, marker as its starting-point. This is fully consonant with the standard definition of grammaticalization, proposed by Kuryłowicz 1965 (quoted in Lehmann 1995: 7), as a process whereby a given item “[advances] from a lexical to a grammatical or from a less grammatical to a more grammatical status” (my italics, Ph. B.).

Because it seems, on the face of it, so very singular, the behavior of *sii* as a Different-Subject marker invites a thoroughgoing reappraisal of the conceptual underpinnings of directional deixis. What this reappraisal leads to is a redefinition of ventive markers and itive markers as, essentially, exponents of *oriented motion + identity* and *oriented motion + otherness*, respectively – where ‘identity’ and ‘otherness’ refer to the relationship between one selected anchoring coordinate and either of two variables involved in the motion event, namely the location of its endpoint and the individual functioning as figure.

This schematic redefinition, which borrows loosely from some of the key conceptual tenets of Antoine Culioli’s Theory of Enunciative Operations (Groussier 2000), is both constrained and flexible enough, I believe, to allow for a principled comparison between systems which, from one language to the next, are prototypically, but not exclusively, designed to relate the endpoint of a trajectory in space to an anchor. The very concept of anchoring raises theoretical questions pertaining to the nature and scope of deicticity. I have argued, on the basis of empirical evidence, for a unitary definition, which conjoins an *exophoric* subtype – namely, *deictic* anchoring *stricto sensu* – and an *endophoric* one.

Programmatic though it may be, the theory of directional deixis that emerges is not without some significant fallout; so is the analysis, as sketched above, of its instantiation in Somali. Two components of this fallout are worth stressing.

The literature is rife with the contribution of directional deictics, widely attested across languages, to the marking of tense, aspect and mood. While the ability of *soo* to highlight the persistence of a process up to reference time demonstrates that Somali is clearly not immune to this type of functional extension, the language sanctions an even more fundamental link between directional deixis and switch reference, as a specific subsystem of the category of person. Whether or not it was at some stage the end product of processes of functional extension, it is plausible to assume that this *transcategorical* linkage is built into the grammar of contemporary Somali, to the extent that it is constitutive of the meaning of *soo* and – far more spectacularly – of *sii*. There is nothing typologically controversial in this analysis, given the propensity of ventive and itive markers to do double duty as directional deictics and person markers (Mithun 1996; Bourdin 1998). And there should be nothing too controversial, either, about the concept of *transcategoriality*, whether it be from an empirical standpoint (e.g. the nature of grammaticalization phenomena) or from a theoretical one: again, the framework developed by Antoine Culioli and his followers over the past thirty years or so (Groussier, *ibid.*) irresistibly comes to mind.

The *prohibition of otherness overload*, which has been shown to circumscribe the interpretation of *sii*, is arguably a general principle flowing from general constraints on communicative processing and as such extraneous to the meaning *per se* of itive *sii*. It is the counterpart to a much better-known *ban on identity overload*, which is applicable to the interpretation of ventive markers and is itself the direct result of obvious real-world constraints: thus, the English sentence *I'm coming* cannot possibly mean that 'I am now on my way to the place where I am' – a construal which would mechanically proceed from the triple identification of the moving entity with the speaker, the goal of motion with the speaker's location, and the time of process with speech time. This, in turn, suggests that the ability of English *come* or Somali *soo* to signal motion to the addressee's, rather than the speaker's location, is an epiphenomenal *effet de sens*, rather than an integral part of their semantics. If so, the standard dichotomy between languages which allow their ventive marker(s) to denote motion towards the addressee and those that do not should itself be viewed as a mere epiphenomenon: more fundamental is the prohibition of identity overload, which simply forces languages that do not exclusively identify the exophoric anchor with the speaker to select the addressee's location, when need be, as the default option.

This particular distinction between a "local" epiphenomenon on the one hand and the mechanical application of a general and independently motivated principle is by no means trivial, for it is difficult to see how typological research can move forward unless such distinctions are gradually brought to light.

The goal of this paper was to begin taking apart the Somali system of directional deixis so that it can ultimately be used as a benchmark in typological investigations designed to sharpen our understanding of directional deixis. I hope to have shown, at the very least, that the Somali system can be viewed as one specific translation, or configuring, of the abstract schema which defines directional deixis cross-linguistically, rather than an idiosyncratic, hence essentially "enigmatic", response to a problem that a great many languages demonstrably make it one of their structural tasks to solve, one way or another.

Notes

* With the generous and sagacious assistance of Mr. Yassin Handouleh, a native speaker of Somali, to whom I am very thankful. My gratitude also goes to Lucyna Gebert, Bernd Heine and Maarten Mous for their judicious comments and suggestions on an earlier version of this paper. Of course, all remaining errors and shortcomings are exclusively mine.

1. The following abbreviations will be used in the morpheme-by-morpheme glosses: DEF definite; DET determiner; FEM feminine; IMPER imperative; INFIN infinitive; MASC masculine; PL plural; POSSESSIVE possessive; PRES present; PROG progressive; SG singular; VEN ventive.

2. *Soo* and *sii* are, more precisely, categorized as *preverbal adverbial clitics* by Saeed (1999: 125–126), my chief reference with respect to the grammatical structure of Somali. Heine (1978: 38) has shown that cognates of ventive *soo* are to be found in the other languages (Boni, Rendille,

Jabarti) belonging to the Sam subgroup of East Cushitic, but that they behave like verbal prefixes rather than as clitics. He suggests a common Proto-Sam form (reconstructed as **soo*), and makes the important point that present-day Sam languages other than Somali offer no evidence of a corresponding itive form in Proto-Sam; itive *sii*, unlike ventive *soo*, would thus appear to have been a subsequent innovation confined to Somali.

3. While it is cross-linguistically very frequent for verbs of ‘coming’ to denote an arrival and for verbs of ‘going’ to denote a departure, it is typologically uncommon for a ‘go’ verb to alternate between the meaning of ‘depart’ and that of ‘arrive’ (Bourdin 1999a). It is not unheard of, however:

- (i) [Sinhalese (Indo-European, Indo-Aryan; Sri Lanka); Reynolds (1980:218)]

isteesama-ta giyaa-ma koocciya gihin
 station-DATIVE go:PAST-ANTERIOR train go:ABSOLUTIVE
 ‘When (we) got to the station, the train had gone.’

- (ii) [Misantla Totonac (Totonacan; Mexico); MacKay (1999:446)]

- a. *pues awi? lakaa-ču na[ɬ] ik-qn*
 well now NEGATIVE-CLITIC FUTURE 1[SUBJECT]-go
 ‘Well, now I’m not leaving.’
 b. *čaa-qn-la[ɬ] qntuhu*
 only-go-PERFECTIVE:1SG/3 over.there
 ‘He had just arrived there.’

4. German *kommen* is also severely constrained in this regard. While *die kommende Woche* is a perfectly admissible phrase, Mikame (1996:380) points to the following acceptability contrast:

*Sie gingen / *kamen langsam nach Hause. Unterwegs an der Weidendammer Brücke, fragte Pony Hütchen...* ‘They walked slowly towards the house. On the way, on the Weidenhammer bridge, Pony asked Hütchen. . .’ [E. Kästner, *Emil und die Detektive*, Hamburg, 1991]

The inherent telicity of *kommen* makes it incompatible with a context (*unterwegs*, ‘on the way’) which requires the motion event to be viewed “from the inside”, so to speak.

5. One may well ask why a fifth interpretation is disallowed, whereby Anab would be talking to Abdi on the phone and inviting him to come to her place to get some rest. This particular scenario, in fact, would require the use of *kaalay*, the suppletive imperative form of the ventive verb *imaw*:

i kaalay waad nasaan-e
 1SG[OBJECT] come:IMPER.2SG FOCUS:2SG take.some.rest:INFIN-and
 ‘Come (to my house) to get some rest.’

It may well be that there are definite limits to the trajectory that *soo*, in and of itself, can refer to, at least when the motional event is preparatory to P. What seems to be precluded is any kind of transit from one enclosed space to another. On the other hand, interpretation (i) of (15) suggests that *soo* is licensed when the trajectory takes place within the same enclosed space.

6. This might create the expectation that *soo* would occur less frequently in texts than *sii* does. In actual fact, just the opposite is true: *soo* is ubiquitous in everyday discourse, whereas *sii* is comparatively rare. I am thankful to M. Mous for bringing this contradiction to my attention. Clearly, the difference in frequency must involve other, as yet undetermined, factors.

7. Much the same mechanism of cataphoric anticipation accounts for the first occurrence of the ventive verb *tuaj* in the following Hmong example:

[Hmong (Hmong-Mien; Thailand, China, Laos, Vietnam); Bisang (1992:223)]

koj tuaj qhov twg tuaj?

you come place what come

'Via what village did you come here?'

However, the acceptability of this particular serial construction is conditional upon the village being located between the addressee's starting-point and the deictic center. As (44) testifies, Somali places no such geographical constraint on cataphoric uses of ventive *soo*.

8. This has been very explicitly articulated by Lichtenberk (1991) through the concept of *heterosemy*.

9. In this regard, I view as singularly misplaced the contention, breezily sketched by Wilkins and Hill (ibid., Footnote 9, p. 253), that Heine et al. (1993) are wrong in arguing for a link between the notional commonality shared by directional deictics across languages and their susceptibility to certain pathways of grammaticalization. This is only one of the many puzzling assumptions or statements made by Wilkins and Hill. To take one other particularly telling example, one may well wonder exactly what linguists they have in mind when they claim, as they do at the very beginning of their paper (p. 209), that "it is *commonly* assumed that *all* languages have a class of motion verbs and that this class will minimally include two forms which correspond to English "come" and "go" (my italics, Ph. B.). Such putative members of the linguistic community would have to overlook, at the very least, Vietnamese, not to mention Russian and other Slavic languages... Much in the same vein is the allusion (p. 210) to "those...[who] could seize on these subclasses in English and Yucatec as *proof* that *all* languages minimally possess a class of "basic" motion verbs" (my italics, Ph. B.): it is, to say the least, mind-boggling to think that any self-respecting typologist could possibly number among people ready to carry to such irrational extremes their defence of the universality of 'come' and 'go'! These fanciful appeals to "strawmen" – a charge which they strenuously, though unconvincingly, reject in anticipation of its being made against them (Footnote 9, p. 253) – detract somewhat from the force, if not the substance, of Wilkins and Hill's argumentation.

10. This is a greatly simplified picture, as there is ample cross-linguistic evidence that directional deictics call into play two distinct types of endophoric anchoring. While (53b) involves *sentence-based* anchoring, the following example attests that endophoric anchoring may also be *text-based*:

[Thai (Daic, Tai); Bisang (1992:365)]

din.sǎ̌ lùd paj càag myy

pen slip go out.of hand

'The pen slipped from his hand.'

The anchor here is none other than the individual who is the current topic of discourse. It is well-known that in narratives, languages routinely use ventive and itive markers in reference to the location of the main protagonist, that of the character with whom the audience is supposed to identify or, anaphorically, the location which has last been mentioned in the immediate co-text: see for instance Fillmore (1975:67).

11. Newman (2000:663–664) delineates the formal and semantic implications of such cooccurrence, as well as the restrictions to which it is subject.

12. I do not know of any evidence in the literature regarding the instantiation of switch reference in Niger-Congo languages. However, Prost's characterization of *k'* in (60) as a marker of *opposition de sujet* (ibid., p. 377) would seem to leave little doubt that we are dealing with a case of switch reference.

References

- Abdoulaye, M. L. (1996). Efferential 'verb + *dà*' constructions in Hausa. *Journal of African Languages and Linguistics*, 17 (2), 113–151.
- Amador, E. M. Martínez (1985). *Diccionario gramatical y de dudas del idioma*. Barcelona: Editorial Ramón Sopena.
- Asher, R. E., & T. C. Kumari (1997). *Malayalam*. London: Routledge.
- Bisang, W. (1992.) *Das Verb im Chinesischen, Hmong, Vietnamesischen, Thai und Khmer (Vergleichende Grammatik im Rahmen der Verbserialisierung, der Grammatikalisierung und der Attraktorpositionen)*. Tübingen: Gunter Narr.
- Boeschoten, H. (1998). Uzbek. In L. Johanson & É.Á. Csató (Eds.), *The Turkic Languages* (pp. 357–378). London: Routledge.
- Bourdin, P. (1998). Deixis personnelle et deixis directionnelle: une corrélation problématique. In B. Caron (Ed.), *Proceedings of the 16th International Congress of Linguists*, art. 0053 [on CD-Rom]. Oxford: Pergamon.
- . (1999a). Deixis directionnelle et "acquis cinétique": de 'venir' à 'arriver', à travers quelques langues. *Travaux linguistiques du CerLiCO* 12 (La référence – 2 – Statut et processus), 183–203.
- . (1999b). The grammaticalization of 'come' and 'go': a window on cognitive constructions of time? Paper presented at the 6th International Cognitive Linguistics Conference. University of Stockholm, 10–16 July, 1999.
- . (2002). The grammaticalization of deictic directionals into modulators of temporal distance. In G. Diewald & I. Wischer (Eds.), *New Reflections on Grammaticalization* (pp. 118–199). Amsterdam: John Benjamins.
- Comrie, B. (1989). Some general properties of reference-tracking systems. In D. Arnold et al. (Eds.), *Essays on Grammatical Theory and Universal Grammar* (pp. 37–51). Oxford: Clarendon Press.
- Craig, C. Grinevald (1979). Jacalteco: field work in Guatemala. In T. Shopen (Ed.), *Languages and their Speakers* (pp. 3–57). Cambridge, MA: Winthrop.
- Di Meola, C. (1994). *Kommen und gehen: Eine kognitiv-linguistische Untersuchung der Polysemie deiktischer Bewegungsverbene*. Tübingen: Max Niemeyer.
- Fillmore, C. J. (1975). *The Santa Cruz Lectures on Deixis*. Bloomington, IN: Indiana University Linguistics Club.
- Foley, W. A., & Van Valin Jr., R. D. (1984). *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Frajzyngier, Z. (1989). *A Grammar of Pero*. Berlin: D. Reimer.
- Groussier, M.-L. (2000). On Antoine Culioli's theory of enunciative operations. *Lingua*, 110, 157–182.

- Haviland, J. B. (1993). The syntax of Tzotzil auxiliaries and directionals: the grammaticalization of 'motion'. In D. A. Peterson (Ed.), *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society (Feb. 12–15, 1993): Special session on syntactic issues in native American languages* (pp. 35–49). Berkeley, CA: Berkeley Linguistics Society.
- Hayward, D. (1984). *The Arbore Language: a first investigation (including a vocabulary)*. Hamburg: Helmut Buske.
- Heine, B. (1978). The Sam languages: a history of Rendille, Boni and Somali. *Afroasiatic Linguistics*, 6 (2), 1–93.
- Heine, B. et al. (1993). *Conceptual Shift. A lexicon of grammaticalization processes in African Languages*. Cologne: Universität zu Köln, Institut für Afrikanistik [Afrikanistische Arbeitspapiere 34/35].
- Hooper, R. (2000). Aspectual uses of Tokelauan directional particles *mai* and *atu*. Unpublished ms.
- Howard, K. (2000). The notion of current relevance in the Thai perfect. *Linguistics*, 38 (2), 373–407.
- Jo, In-hee (1990). Multi-verb constructions in Korean. In B. D. Joseph & A. M. Zwicky (Eds.), *When Verbs Collide: Papers from the 1990 Ohio State Mini-conference on Serial Verbs* (pp. 265–287). Columbus, OH: The Ohio State University [Working Papers in Linguistics].
- Kendall, M. B. (1975). The /-K/, /-M/ problem in Yavapai syntax. *International Journal of American Linguistics*, 41 (1), 1–19.
- Koch, H. (1984). The category of 'Associated Motion' in Kaytej. *Language in Central Australia*, 1, 23–34.
- Kuryłowicz, J. (1965). The evolution of grammatical categories. *Diogenes*, 51, 55–71.
- Lehmann, Ch. (1995). *Thoughts on Grammaticalization*. Munich: Lincom Europa.
- Lichtenberk, Fr. (1991). Semantic change and heterosemy in grammaticalization. *Language*, 67, 475–509.
- MacKay, C. J. (1999). *A Grammar of Misantla Totonac*. Salt Lake City UT: The University of Utah Press.
- Martin, S. E. (1975). *A Reference Grammar of Japanese*. New Haven, CT: Yale University Press.
- Mikame, H. (1996). Markierte Perspektive, perspektivische Annäherung des Sprechers an das Objekt und direkte Wahrnehmung (Zur Signalisierung der psychisch-kognitiven Nähe des Sprechers zum Objekt). *Sprachwissenschaft*, 21, 367–420.
- Mithun, M. (1996). New directions in referentiality. In B. Fox (Ed.), *Studies in Anaphora* (pp. 413–435). Amsterdam: John Benjamins.
- Mous, M. (2001). Selectors in Cushitic. Paper presented at the *International Symposium: Typology of African Languages*. Sankt Augustin, Germany, May 21–24, 2001.
- Munro, P. (1976). *Mojave Syntax*. New York: Garland.
- Newman, P. (2000). *The Hausa Language: An encyclopedic reference grammar*. New Haven, CT: Yale University Press.
- Prost, A. (1964). *Contribution à l'étude des langues voltaïques*. Dakar: Institut français d'Afrique noire.
- Reynolds, C. H. B. (1980). *Sinhalese: An introductory course*. London: School of Oriental and African Studies.
- Ricca, D. (1993). *I verbi deittici di movimento in Europa: una ricerca interlinguistica*. Florence: La Nuova Italia Editrice.
- Rottland, F. (1982). *Die südnilotischen Sprachen: Beschreibung, Vergleichung und Rekonstruktion*. Berlin: D. Reimer.
- Saeed, J. (1999). *Somali*. Amsterdam: John Benjamins.

- Samb, A. (1983). *Initiation à la grammaire wolof*. Dakar: Institut fondamental d'Afrique noire.
- Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (Ed.), *Language Typology and Syntactic Description*. Vol. III: *Grammatical categories and the lexicon* (pp. 57–149). Cambridge: Cambridge University Press.
- Tchekhoff, C. (1990). Discourse and Tongan *mai, atu, ange*: scratching the surface. In J. H. C. S. Davidson (Ed.), *Pacific Island Languages: Essays in honour of G. B. Milner* (pp. 105–110). London: School of Oriental and African Studies/Honolulu, HI: University Press of Hawaii.
- Tunbridge, D. (1988). Affixes of motion and direction in Adnyamathanha. In P. Austin (Ed.), *Complex Sentence Constructions in Australian Languages* (pp. 267–283). Amsterdam: John Benjamins.
- Vuillaume, M. (1983). *Hin et her. Nouveaux cahiers d'allemand* (Université Nancy II) 1(2): 98–109 and 1(3): 204–221.
- Watkins, L. J. (with the assistance of P. McKenzie). (1984). *A Grammar of Kiowa*. Lincoln, NB: University of Nebraska Press.
- Wilkins, D. P., & D. Hill (1995). When 'go' means 'come': questioning the basicness of basic motion verbs. *Cognitive Linguistics*, 6 (2/3), 209–259.

A typology of subject and object markers in African languages

Denis Creissels

Laboratoire Dynamique du Langage (CNRS) & Université Lumière (Lyon2)

This paper proposes a typological survey of bound pronominal morphemes representing subjects and objects in African languages. Bound pronouns show variations concerning the conditions in which they are used (exclusively as pronouns, or also as obligatory agreement markers), the nature of the words they can attach to, and the importance of the morphophonological processes in which they are involved. In African languages, bound pronouns are not always easy to distinguish from free pronouns, but in fact, most African languages have bound pronominal morphemes analyzable as subject markers affixed to verbs. However, the use of subject markers as obligatory agreement markers, although relatively common, is not general, and object markers never function as obligatory agreement markers. Another interesting generalization is that the indexation patterns of typical ditransitive verbs show a strong tendency in African languages to assimilate the recipient of such verbs to the patient of typical transitive verbs.

1. Introduction

In this paper, the term ‘pronominal marker’ is applied to any bound morpheme referring to an entity that is represented elsewhere in the same clause by a noun phrase, or could be represented by a noun phrase in a clause identical in all other respects, and whose variations reflect, either certain semantic characteristics of the entity in question, or certain grammatical features of a noun phrase referring to the same entity in the same construction.

Pronominal markers typically show variations expressing distinctions that parallel those expressed by free pronouns, in particular distinctions in person and number.¹

When discussing properties of the pronominal markers, the expression ‘the corresponding noun phrase’ will be used here as an abbreviation for ‘the noun phrase referring to the same entity that is present in the same clause or that could be used to refer to the same entity in a clause identical in all other respects’.

‘Subject markers’ are pronominal markers that correspond to a noun phrase in subject function, and ‘object markers’ are pronominal markers that correspond to a noun phrase in object function.

This definition groups together several types of morphemes that may differ in some important aspects of their grammatical behavior – see Section 2. But, diachronically, the subtypes of pronominal markers can be viewed as successive stages in the evolution of former pronouns that in a first stage lose their status of autonomous words, and that subsequently may undergo additional modifications in their behavior without entirely losing the semantic properties of the pronouns they originate from.

In this connection, it is important to keep in mind that a variety of terms are applied to pronominal markers in different grammatical traditions (pronominal affixes, weak pronouns, non-emphatic pronouns, pronominal clitics, personal endings of verbs, subject/object concords, etc.), and that the choice between these different labels does not correlate in any consistent way with variations in the properties of the pronominal markers.

Even more importantly, in the current orthography of many languages, pronominal markers are written as distinct words, and in many descriptive grammars, they are not clearly distinguished from free pronouns. In other words, their status as bound morphemes is not always recognized correctly. This question will be discussed in Section 3, but it seems to me important to immediately emphasize that, in a cross-linguistic study of pronominal markers, there would be little sense in recognizing the existence of such morphemes in a language by simply relying on orthographic conventions or on the labels currently used in descriptive grammars.

2. Subtypes of pronominal markers according to their conditions of co-occurrence with the corresponding noun phrases

2.1 The three stages in the evolution of pronominal markers

Among the morphemes recognizable as pronominal markers according to the definition put forward in Section 1, three subtypes can be distinguished on the basis of their conditions of co-occurrence with the corresponding noun phrases. Diachronically, there is a considerable amount of evidence that these three subtypes represent successive stages in an evolution whose starting point is the cliticization of free pronouns.

Stage I pronominal markers are in complementary distribution with the corresponding noun phrase within the limits of the clause, and the choice between the pronominal marker and the corresponding noun phrase depends on the discourse structure of the clause: the same entity is represented by a pronominal marker or by a noun phrase depending on its degree of topicality and recoverability from the context, and the pronominal marker co-occurs with the corresponding noun phrase only if the noun phrase is topicalized in a dislocated construction; for example, modern Romance languages have pronominal morphemes (commonly termed ‘clitic pronouns’) that are

morphosyntactically bound to the verb, but that in most cases are used only to refer to an entity that is not represented by a noun phrase in the same clause.

Stage II pronominal markers are obligatory, even if a noun phrase or a free pronoun referring to the same entity is present in subject or object function, whereas the corresponding noun phrases or free pronouns are not obligatory constituents of the clause. In such situations, a given participant is obligatorily referred to by a pronominal marker; the corresponding noun phrase can be viewed as providing additional information helping to identify the referent in case the indications given by the pronominal marker and by the context are not sufficient, and the corresponding free pronoun occurs only to express emphasis. For example, in Latin, clauses do not necessarily include a noun phrase or (free) pronoun in subject function, but the argument that can optionally be encoded as the subject of a verb is obligatorily referred to by means of a pronominal marker incorporated in the verb ending.

Stage III pronominal markers share with stage II pronominal markers the property of obligatoriness, but they differ from them by not being able to represent by themselves the entity they refer to. In other words, constructions involving stage III pronominal markers must include a noun phrase or a free pronoun referring to the same entity. The English marker *-s* attached to verbs in the indicative present combined with a third person singular subject is an illustration of this type of pronominal marker: *-s* is a pronominal marker in the sense of the definition put forward here, since its presence vs. absence depends on grammatical characteristics of the subject noun phrase or on semantic characteristics of its referent, and it belongs to the third subtype, since in a clause *NP V-s*, neither the noun phrase in subject function nor the marker *-s* can be suppressed.

The pronominal nature of stage I pronominal markers is particularly clear, since they are functionally equivalent to free pronouns in the sense that, within the limits of the clause, there is no other trace of the entity referred to by means of a stage I pronominal marker. By contrast, stage III pronominal markers clearly function as pure agreement markers, and stage II pronominal markers have a mixed status, since they share with free pronouns the ability to constitute the only trace of the entity they refer to, but when a noun phrase referring to the same entity is present in the same clause, the obligatoriness of stage II pronominal markers makes it possible to consider them as agreement markers.

There tends to be a correlation between these three stages in the evolution of the syntactic properties of pronominal markers and changes in their morphophonological properties: from the morphophonological point of view, stage I pronominal markers remain generally easy to isolate as distinct segments in the morphological structure of the word they are attached to, whereas stage II or III pronominal markers often show a high degree of fusion with the other elements of the word.

2.2 Subject/object markers functioning as pure agreement morphemes in African languages

Subject or object markers functioning as pure agreement morphemes (i.e. subject or object markers whose variations refer to an argument encoded as a noun phrase in subject or object function but that cannot by themselves represent the argument they refer to) are not common in the languages of the world. Not surprisingly, examples of pronominal markers of this type are not easy to find in African languages. A possible example of stage III subject markers in an African language is the agreement of Fula verbs with their subject in gender/number, manifested by consonant alternations at the initial of verb stems – ex. (1).²

(1) FULA

- a. *debbo wari*
woman came
'The woman came'
- b. *rewbe ηgari*
women came
'The women came'
- c. *o-wari*
SM3S³-came
'(S)he came'
- d. *be-ηgari*
SM3P-came
'They came'

2.3 Discourse dependent and obligatory subject markers in African languages

Among African languages, one commonly encounters both languages with discourse dependent subject markers – ex. (2) – and languages with obligatory subject markers – ex. (3). Stage I subject markers are particular common in some language families and stage II subject markers in others, but it is difficult to say which of these two types predominates at the level of the African continent.

(2) ANYI (Quaireau 1987)

- a. *kuakú dafi*
Kuaku sleep
'Kuaku is sleeping'
- b. *ɔ-dafi*
SM3S-sleep
'He is sleeping'
- c. *kuakú díé ɔ-dafi*
Kuaku TOP SM3S-sleep
'As for Kuaku, he is sleeping'

(3) Tswana

- a. *kítsó* *ú-tsílè*
CL1.Kitso SMCL-come.TAM
'Kitso has come'
- b. *ú-tsílè*
SMCL-come.TAM
'He has come'
- c. **kítsó tsílè*

In this respect, the subject markers of some languages do not behave in a uniform way: for example, the subject clitics of French are obligatory in the first and second person but optional in the third person; Mende – ex. (4) – has subject markers in complementary distribution with the corresponding noun phrase in the third person singular, contrasting with obligatory subject markers in the third person plural.

(4) Mende (Innes 1971)⁴

- a. *i-kólíí* *lóló*
SM3S-leopard see.PAST
'He saw a leopard'
- b. *kpanà kólíí lóló*
Kpana leopard see.PAST
'Kpana saw a leopard'
- c. *tí-kólíí* *lóló*
SM3P-leopard see.PAST
'They saw a leopard'
- d. *maheisia tí-kólíí lóló*
chiefs SM3P-leopard see.PAST
'The chiefs saw a leopard'

2.4 Discourse dependent and obligatory object markers in African languages

Typical stage II object markers, i.e. object markers necessarily present in transitive constructions, even if in the presence of the corresponding noun phrase, are not uncommon in the languages spoken in certain parts of the world. For example, a number of Amerindian languages have a class of transitive verb stems that must combine with an object marker, as illustrated here by Nahuatl – ex. (5).

(5) Nahuatl (Launey 1981)

- a. *ní-c-cua* *in nacatl*
SM1S-OM3S-eat DEF meat
'I am eating the meat'

- b. *ni-c-cua*
SM1s-OM3s-eat
'I am eating it'
- c. *ni-c-cua* *nacatl*
SM1s-OMS-eat meat
'I am eating meat'
- d. **ni-cua nacatl*, **ni-cua*
(‘I am eating’, without any specification of the thing eaten, would be in Nahuatl *ni-tla-cua*, where *-tla-* is an intransitivizing morpheme)

In Africa, typical stage II object markers (obligatory object markers) are not frequent; by contrast, stage I object markers (discourse dependent object markers) are very common.

Tswana – ex. (6) – provides a good illustration of object markers that always represent topics, and are therefore in strict complementary distribution with noun phrases in object function, the choice between an object marker and a noun phrase or free pronoun in object function being always possible and pragmatically significant.

(6) Tswana

- a. *kì-χò-bídítšè*
SM.1s-OM.2s-call.TAM
'I called you' (how is it possible that you didn't hear me?)
- b. *kì-bídítšé* *wèná*
SM.1s-call-TAM you
'I called you' (and nobody else!)

However, many languages have object markers that are stage I object markers in the sense that they are not always present in transitive constructions, but that depart from the typical behavior of stage I object markers in the sense that at least in certain conditions, they must be present even if the corresponding noun phrase or free pronoun is also present. Historically, such systems can be viewed as systems in a transitional stage between stage I and stage II.

For example, Romance languages have cases of obligatory ‘clitic doubling’, in which a free pronoun or a noun phrase in object or dative function is necessarily accompanied by the corresponding clitic attached to the verb. Among African languages, Swahili illustrates a situation in which pronouns, proper nouns and definite noun phrases trigger the presence of an object marker irrespective of their discourse function, whereas no object marker accompanies indefinite noun phrases in object function – ex. (7). Note that in Swahili, definiteness is not obligatorily marked at noun phrase level, and consequently the presence of an object marker may constitute the only clue to the definiteness of common nouns in object function.

(7) SWAHILI

- a. *ni-me-ku-ona*
SM1s-TAM-OM2s-see
'I have seen you'
- b. **ni-me-ona* *wewe*
SM1s-TAM-see you
- c. *u-me-leta* *chakula?*
SM2s-TAM-bring CL7.food
'Have you brought (some) food?'
- d. *u-me-ki-leta* *chakula?*
SM2s-TAM-OMC7-bring CL7.food
'Have you brought the food?' (which I told you to bring)

2.5 Implicit reference to discursively salient entities

Free pronouns are used to refer to entities by relying maximally on the discourse feature of salience, and minimally on the mention of semantic characteristics independent of the particular speech act within which entities are referred to. This functional characterization of pronouns is shared by stage I pronominal markers, and partially by stage II pronominal markers. But when discussing certain aspects of the typology of pronominal markers according to their conditions of co-occurrence with noun phrases or free pronouns, it must be kept in mind that the use of free pronouns or pronominal markers is not the only way of exploiting the discourse property of salience to refer to an entity involved in an event. In case a noun phrase representing this entity would have the status of argument of the verb, an alternative strategy is the strategy of implicit reference, which relies on an interpretive rule according to which, if there is no explicit representation of an argument within the clause, the role of this argument must be assigned to some salient entity not mentioned in the clause.

Japanese is a good example of a language that systematically uses the strategy of implicit reference to speech act participants or other discursively salient entities. Such a systematic use of this strategy seems to be very rare in Africa. In virtually all African languages, arguments encoded as subjects, if not represented by a noun phrase or free pronoun in subject function, must at least be represented by a subject marker. In the case of arguments encoded as objects, most of the time, their total omission triggers an indefinite rather than anaphoric interpretation.

However, the descriptions of some African languages mention a regular use of the strategy of implicit reference to discursively salient entities, but in rather restrictive conditions. In contrast to Japanese, which extends the use of this strategy to any salient entity in argument function, African languages that make use of it seem to restrict it to inanimate patients.

3. Problems in recognizing the distinction between free pronouns and pronominal markers

Many descriptions of African languages do not mention the existence of bound pronominal morphemes, but in most cases the morphemes termed ‘subject pronouns’ and ‘object pronouns’ in descriptions of African languages are not really separate words and should be reanalyzed as bound morphemes.

The problem is that pronominal markers are easily recognizable as such only if the following three conditions are met:

(1) They are obligatory, so that it is relatively easy to establish a distinction between pronominal morphemes that have the same distribution as nouns (free pronouns), and pronominal elements that have a specific distribution, since they co-occur with nouns (pronominal markers); by contrast, pronominal markers in complementary distribution with the corresponding noun phrases are easy to confuse with free pronouns.

(2) Even in very short and simple sentences, their position cannot be confused with that of the corresponding noun phrase (for example, they are prefixed to the verb and correspond to noun phrases that follow the verb, or they are suffixed to the verb and correspond to noun phrases that precede the verb).

(3) From the morphophonological point of view, they show a relatively high degree of interaction with TAM or polarity markers, i.e. with morphemes that are not expected to interact with nouns or their equivalent; for example, no linguist would imagine the possibility of isolating as distinct words the subject markers that constitute the ending of Italian or Spanish verbs; by contrast, subject markers attached the initial of the verb that do not fuse with the following morphemes, like the clitic subjects of French or Piedmontese, are easily confused with words that precede the verb.

In other words, stage II pronominal markers are generally easy to identify (and most descriptive grammars identify them properly), but stage I pronominal markers are easy to confuse with free pronouns, in particular when their position is at first sight similar to that of the corresponding noun phrase.

Wolof – ex. (8) – provides a good illustration, both of the difficulties in the identification of pronominal morphemes and of the kind of observations that may help to solve this problem.

By itself, the data put forward in (8a) could suggest recognizing *ma*, *nga*, etc. and *ko* as free pronouns in subject/object function. But:

(a) In the verb tenses illustrated in (8b–c), a subject marker is obligatorily present even if there is a noun phrase in subject function, and it is often amalgamated with a

TAM or focus marker in a way that makes it difficult to isolate a segment representing specifically the subject marker;

(b) In (8b), the subject marker is very clearly suffixed to the verb, and its position cannot be confused with that of noun phrase in subject function;

(c) In (8c), the position of the object marker is clearly different from that of a noun phrase in object function.

(8) WOLOF

- a. *xale yi naan meew mi*
 child DEF.PL drink milk DEF
 ‘The children drank the milk’ (in a narrative context)
- ma-naan-ko* ‘I drank it’
nga-naan-ko ‘You SG drank it’
mu-naan-ko ‘He/she drank it’
ñu-naan-ko ‘We/they drank it’⁵
ngeen-naan-ko ‘You PL drank it’
- b. *xale yi naan-nañu meew mi*
 child DEF.PL drink-TAM.SM3P milk DEF
 ‘The children have drunk the milk’ (perfect)
- naan-naa-ko* ‘I have drunk it’
naan-nga-ko ‘You SG drank it’
naan-na-ko ‘He/she has drunk it’
naan-nañu-ko ‘We/they have drunk it’
naan-ngeen-ko ‘You PL have drunk it’
- c. *xale yi dañu naan meew mi*⁶
 child DEF.PL VFOC.SM3P drink milk DEF
 ‘The children have *drunk* the milk’ (with emphasis on the verb)
- dama-ko naan* ‘I have drunk it’
danga-ko naan ‘You SG drank it’
dafa-ko naan ‘He/she has drunk it’
dañu-ko naan ‘We/they have drunk it’
dangeen-ko naan ‘You PL have drunk it’

In cases when the position occupied by the pronominal morphemes in minimal sentences does not reveal their precise status as free pronouns or pronominal markers, two kinds of observations may help to solve the problem:

(a) In more complex constructions, free pronouns are expected to behave with the same mobility as syntactic constituents, whereas pronominal markers must remain attached to their host – for example, in the Wolof example (6a) above, it would be

possible to insert *itam* ‘too’ between a noun phrase in subject function and the verb, but not between the subject markers and the verb stem;

(b) Phonologically, free pronouns undergo only contextual changes resulting from the application of ‘post-lexical’ rules operating at word junctions; by contrast, pronominal markers have frequently allomorphs that cannot be explained as the result of phonological processes operating at word junctions and must be analyzed as the result of the interaction between morphemes belonging to the same word.

Unfortunately, morphosyntactic tests are relatively difficult to apply in African languages, due to their syntactic peculiarities. A first reason is that the contrast between the mobility of free pronouns and the lack of mobility of bound pronominal morphemes is less obvious in languages with a relatively rigid ordering of the constituents of the clause, which is the case of an overwhelming majority of African languages. A second reason is that, among the morpho syntactic tests that may help to recognize bound morphemes, the co-ordination tests are often particularly clear; unfortunately, most African languages do not have an exact equivalent of the English co-ordinating morpheme *and*, and they tend to co-ordinate noun phrases within what can be called the ‘committative strategy’, which makes it impossible to simply transpose the co-ordination tests that prove particularly useful for languages such as French or English.

By contrast, in African languages, a careful observation of the phonological data generally provides evidence that dissipates the hesitations one may have in establishing a distinction between free pronouns and pronominal markers. For example, Ewe has an object marker of third person singular consisting of a single vowel whose quality (i, e or ε) is conditioned by the last vowel of the verb stem – ex. (9), which unambiguously identifies it as a bound pronominal morpheme in spite of the fact that at first sight, it seems to occupy in the clause the same position as an object noun phrase, since Ewe is an SVO language.

(9) EWE (Schadeberg 1985)

- a. *kofi wù-i*
Kofi kill-OM3s
‘Kofi killed it’
- b. *kofi tó-è*
Kofi pound-OM3s
‘Kofi pounded it’
- c. *kofi kpɔ́-è*
Kofi see-OM3s
‘Kofi saw it’

In many languages that have ATR harmony, the fact that pronominal markers undergo vowel harmony is a decisive proof that they cannot be analyzed as autonomous words, as illustrated here with Anyi – ex. (10).

(10) ANYI (Quaireau 1907)

- a. *ɔ-dafi*
SM3s-sleep
'He is sleeping'
- b. *o-di alié*
SM3s-eat food
'He is eating'

Most African languages have tone systems, and many of them have a complex tonal morphology. An advantage of such a situation is that, once the tonal system is known, bound morphemes often turn out to undergo tonal variations that prove their phonological dependence on their host and exclude interpreting them as autonomous words. For example, in Mende, noun phrases in object function immediately precede the verb – ex. (11), so that the object markers in example (11b–c) could easily be confused with free pronouns, but their tone is always the opposite of the tone of the first syllable of the verb stem. In addition to that, in the third person, Mende has a distinction between human and non human object markers, and the third person non human object marker is clearly a bound morpheme, since in many cases it manifests itself through a modification of the initial consonant of the verb and cannot be isolated as a distinct segment – ex. (11d).

(11) MENDE (Innes 1971)

- a. *maheí kólíí lóílɔ / waáílɔ*
chief leopard saw / killed
'The chief saw / killed a leopard'
- b. *maheí ti-lóílɔ*
chief OM3P-saw
'The chief saw them'
- c. *maheí tí-waáílɔ*
chief OM3P-killed
'The chief killed them'
- d. *maheí tóílɔ*
chief OM3SNH.saw
'The chief saw it'

We are now in a position to discuss the following two questions:

- (1) What is the proportion of African languages really devoid of subject markers?
- (2) Among the African languages that have subject markers, what is the proportion of those that are devoid of object markers?

The Omotic language Maale – Amha (2001) – is a clear case of an African languages in which pronominal markers have only a very marginal status: in Maale, the verb is inflected for person and number in the imperative and in the optative only, and apart from that, all the pronominal morphemes of Maale are clearly free pronouns. But such

cases are not frequent among African languages. In particular, the pronominal morphemes of many Mande or Songhay languages may at first sight give the impression of uniformly having the behavior of free pronouns, but precise descriptions always make apparent the existence of allomorphic variations affecting at least certain pronominal morphemes in certain contexts that can be accounted for, neither as case distinctions, nor as the result of post-lexical phonological processes operating at ordinary word junctions – see for example Heath (1999) for Gao Songhay.

Many descriptions of African languages that mention the existence of several sets of pronominal morphemes clearly identify a set of free pronouns, but make no firm decision on the status of the other sets: they are written as separate words, but no evidence supporting the decision of treating them as free forms is explicitly provided. In most cases, such descriptive grammars simply do not contain the information necessary to solve the problem. But what seems to me significant is that, whenever they do, the available evidence always supports the identification of ‘weak’ or ‘non-emphatic’ pronouns as bound morphemes, as shown for example by Ikoro (1996) for Kana, by Eze (1995) for Igbo, and by Kutsch Lojenga (1994) for Ngiti.⁷ The reason why so many descriptions misidentify pronominal markers is simply that stage I pronominal markers, i.e. pronominal markers minimally different from free pronouns, are particularly frequent in African languages.

This suggests the conclusion that, contrary to what a superficial survey of the available descriptions could suggest, an overwhelming majority of African languages do have pronominal markers.

As for the second question, it seems that the vast majority of African languages have both subject markers and object markers. Oromo – Griefenow-Mewis & Bitima (1994) – is a clear case of a language similar to Latin in that it has stage II subject markers suffixed to verbs but uses exclusively free pronouns to pronominalize objects. This situation seems to be relatively common among Cushitic and Omotic languages, but rather uncommon in the other African language families.

4. Pronominal markers attached to words other than the verb

Subject/object markers attached to the verb are particularly common, but this is not the only possibility. For example, Serbo-Croatian has pronominal markers that attach to the first word or phrase of the clause; in the Amerindian language Paez – Rojas (1998), bound morphemes analyzable as the amalgam of a subject marker and of a TAM marker are suffixed to the last word of focalized noun phrases and attach to the verb by default, if no noun phrase is focalized.

These types of attachment of subject / object markers seem to be found in some Khoisan languages (Tom Güldemann, personal communication); I know of no similar case outside the Khoisan phylum, but other types of attachment of pronominal markers to words other than the verb can be sporadically found among African languages.

In Mande languages, the order of the constituents of the clause is *S (v) O V X*, where (*v*) indicates the possible presence of a grammatical word, often called ‘predicative marker’ in the descriptions of Mande languages, which expresses TAM and polarity distinctions.⁸ In such a structure, the cliticization of pronouns in subject function may create subject markers that, at least in transitive clauses, are clearly attached to a word that is not the verb.

In example (11) above, we have already encountered subject markers attached to the first word of an object noun phrase in Mende. In Dan, the subject markers have fused with the predicative marker and have become obligatory: in this language, syntactically, the subject noun phrase is clearly optional, but the verb phrase necessarily begins with a morpheme that amalgamates TAM distinctions with person-number distinctions, and in transitive clauses, this morpheme is separated from the verb by the object noun phrase – ex. (12).

(12) DAN (Doneux)

- a. SM.TAM NP_{OBJECT} V (1P and 2P omitted for brevity’s sake)
- | | | | |
|-----------|-----------|-----------|-------------------|
| <i>ā</i> | <i>yá</i> | <i>bə</i> | ‘I eat rice’ |
| | rice | eat | |
| <i>má</i> | <i>yá</i> | <i>bə</i> | ‘I ate rice’ |
| <i>ī</i> | <i>yá</i> | <i>bə</i> | ‘You SG eat rice’ |
| <i>bá</i> | <i>yá</i> | <i>bə</i> | ‘You SG ate rice’ |
| <i>yə</i> | <i>yá</i> | <i>bə</i> | ‘(S)he eats rice’ |
| <i>yà</i> | <i>yá</i> | <i>bə</i> | ‘(S)he ate rice’ |
| <i>wə</i> | <i>yá</i> | <i>bə</i> | ‘They eat rice’ |
| <i>wà</i> | <i>yá</i> | <i>bə</i> | ‘They ate rice’ |
- b. NP_{SUBJECT} SM3.TAM NP_{OBJECT} V
- | | | | | |
|--------------|-----------|-----------|-----------|-------------------------|
| <i>nə</i> | <i>yə</i> | <i>yá</i> | <i>bə</i> | ‘The child eats rice’ |
| <i>nə</i> | <i>yà</i> | <i>yá</i> | <i>bə</i> | ‘The child ate rice’ |
| <i>nə nú</i> | <i>wə</i> | <i>yá</i> | <i>bə</i> | ‘The children eat rice’ |
| <i>nə nú</i> | <i>wà</i> | <i>yá</i> | <i>bə</i> | ‘The children ate rice’ |

Another interesting case in point is Ewe. With ordinary transitive verbs, the object markers of Ewe are necessarily attached to the verb, as illustrated in example (9) above. But in the construction of transfer verbs, the noun phrase representing the transferred thing precedes the noun phrase representing the recipient, and the recipient can be represented by an object marker attached to the last word of the noun phrase representing the transferred thing – ex. (13).

(13) EWE (Felix Ameka, personal communication)

- a. *é-ná* *tsi-i*
 SM3s-give water-OM3s
 ‘(S)he gave him/her water’

- b. *é-fiá dɔ-ε*
 SM3s- how work-OM3s
 ‘(S)he taught him/her a profession’
- c. *é-fiá tepé nyui áqé-e*
 SM3s-show place good INDEF-OM3s
 ‘(S)he showed him/her a nice place’

5. Distinctions in the phonological shape of subject and object markers

Subject and object markers sharing the same semantic features may have identical phonological forms. However, in a number of African languages, even among those that have no case distinction between subjects and objects, subject markers differ from the corresponding object markers, at least in some persons.

It must be noted that, in African languages, differences in the phonological shape of pronominal markers sharing the same semantic features almost always have a straightforward explanation in terms of the traditionally recognized syntactic functions. A possible exception is Anywa – Reh (1996). This language has two sets of pronominal markers attached to verbs, but there is no one-to-one correspondence between these two sets and the syntactic functions subject and object, and Reh analyzes the correspondence as a case of split ergativity: in certain constructions, prefixes are used to represent the agent of typical transitive verbs, and suffixes represent the unique core argument of intransitive verbs or the patient of transitive verbs, whereas in other constructions, the same suffixes are used to represent the unique core argument of intransitive verbs or the agent of transitive verbs. But I have found no uncontroversial example of an African language with a system of subject and object markers conforming to the ergative pattern (in which intransitive subject markers are identical to the object markers and different from the transitive subject markers), the agentive pattern (in which intransitive verbs divide in two classes, the intransitive subject markers being identical with the subject markers of transitive verbs in one class, and with the object markers in the other), or the direct/inverse pattern (in which a given combination of persons in transitive verb morphology is encoded without taking into account the respective roles of the arguments referred to, and a distinct morpheme indicates whether the assignation of semantic roles follows a certain hierarchy of persons or violates it).

In African languages, differences in the phonological shape of subject and object markers sharing the same semantic features are more common in the 1st and 2nd person than in the 3rd person; they are more common in the singular than in the plural, and in the 3rd person singular, they are more common for pronominal markers that typically refer to human or animate entities than for those that don’t have this property.

In tonal languages, it is very common that subject and object markers have the same segmental shape but differ in their tonal properties. For example, in Tswana, subject markers divide into four sets according to their tonal behavior, and object markers show a tonal behavior that does not coincide with that of any of the sets of

subject markers; but object markers have a segmental shape distinct from that of the corresponding subject markers only in the first person singular, in the second person singular and in class 1.

It is also worth noting that very often, differences in the shape of subject and object markers are a mere consequence of the fact that subject markers tend to fuse with other types of morphemes expressing various types of semantic distinctions typically encoded through verbal morphology, whereas the same phenomenon rarely affects object markers. We have already seen – see ex. (6) above – that, in the conjugation of Wolof, each individual tense is characterized by a particular set of ‘tense-person complexes’, forms that amalgamate TAM, polarity and/or focus distinctions with person-number distinctions referring to the subject in a way that makes it very difficult to decide whether these tense-person complexes must be analyzed as monomorphemic or bi-morphemic; by contrast, as illustrated by ex. (14), object markers have the same form in all tenses and are always clearly isolable from the neighboring morphemes.

(14) WOLOF

- a. *mu- wut- ma* ‘He looked for me’ (narrative)
 SM LOOK+FOR OM
mu- wut- la ‘He looked for you SG
mu- wut- ko ‘He looked for him / her / it’
mu- wut- ñu ‘He looked for us’
mu- wut- leen ‘He looked for you PL / them’⁹
- b. *wut- na- ma* ‘He has looked for me’ (perfect)
 look+for SM OM
wut- na- la etc.
wut- na- ko
wut- na- ñu
wut- na- leen
- c. *dafa- ma wut* ‘He looked for me’
 SM OM look+for (emphasis on the verb)
dafa- la wut etc.
dafa- ko wut
dafa- ñu wut
dafa- leen wut
- d. *démb la- ma wut* ‘He looked for me yesterday’
 yesterday SM OM look+for (focus on ‘yesterday’)
démb la- la wut etc.
démb la- ko wut
démb la- ñu wut
démb la- leen wut

e.	<i>wut-</i>	<i>u-</i>	<i>ma</i>	'He did not look for me'
	look+for	SM	SMOM	
	<i>wut-</i>	<i>u-</i>	<i>la</i>	
	<i>wut-</i>	<i>u-</i>	<i>ko</i>	
	<i>wut-</i>	<i>u-</i>	<i>ñu</i>	
	<i>wut-</i>	<i>u-</i>	<i>leen</i>	

In contrast to what is observed in many languages of the world, what is remarkable in Wolof verbal morphology is that the fusion of subject markers with other types of morphemes that cross-linguistically tend to have the status of verb affixes contrasts with the total absence of any phonological interaction between the verb stem and the tense-person complex that results from the fusion of the subject marker with other types of markers. In such a system, the bound nature of the subject markers is obvious, but what is not obvious is whether the 'tense-person complex' resulting from the fusion of the subject marker and of other types of grammatical markers must be considered as a verbal affix or as a distinct word. Hausa illustrates the same situation, which seems to be fairly common in African languages.

6. Object markers and ditransitive verbs

In the preceding sections, object markers have been identified as such and discussed in reference to typical transitive verbs, i.e. verbs with two arguments to which they assign the roles of agent and patient. In this section, we examine the behavior of ditransitive verbs regarding the indexation of their arguments.

Transfer verbs can be considered as the prototype of ditransitive verbs, and the following discussion will be mainly based on the most common of them, 'give'. 'Give' has three arguments, the giver, the transferred thing and the recipient.

The observation of the indexation of the arguments of 'give' across languages having object markers in typical transitive constructions confirms that languages may organize the construction of transfer verbs in various ways: the argument assimilated to the patient of typical transitive verbs may be the transferred thing, or the recipient, but it may also happen that both the transferred thing and the recipient are represented by noun phrases whose grammatical behavior is similar to that of the object of typical transitive verbs.

6.1 Indexation systems in which object markers attached to transfer verbs can represent the transferred thing, but not the recipient

In some languages, in the construction of transfer verbs, the transferred thing receives exactly the same treatment as the patient of prototypical transitive verbs, whereas the recipient, either receives a special treatment (corresponding more or less to the traditional notion of dative), or is simply treated as an oblique. In the languages that

have systems of indexation including object markers, this type of organization of the construction of transfer verbs may be reflected in the following two ways:

(a) The transferred thing is represented by the same object markers as the patient of typical transitive verbs, whereas the recipient is represented by a special set of pronominal markers (dative markers); for example in French, in the third person, the transferred thing is represented by the same object clitics *le / la / les* as the patient of typical transitive verbs, and special dative clitics (*lui / leur*) are used to represent the recipient. The same type of indexation of the argument of transfer verbs is found in many other South-European languages (Romance languages, Greek, Basque, Macedonian, etc.), but apart from Berber languages, I know of no African language with pronominal markers functionally similar to the Romance dative markers.

(b) The transferred thing is represented by the same object markers as the patient of typical transitive verbs, whereas the recipient is pronominalized in the same way as obliques, i.e. by means of free pronouns combined with an adposition, or pronominal markers attached to an adposition; situations of the type can be found in African languages, as illustrated here by Mende—ex. (15)

(15) MENDE (Innes 1971)

- a. *kpaná lólí* → *ngi-lólí*
Kpana call OM3SH-call
'Call Kpana' 'Call him'
- b. *mbeí yéyá* → *ngéyá*
rice buy OM3SNH.buy
'Buy the rice' 'Buy it'
- c. *mbeí ve kpaná wε*
rice give Kpana to
'Give the rice to Kpana'
- d. *fe kpaná wε*
OM3SNH.give Kpana to
'Give it to Kpana'
- e. *mbeí ve ngié*
rice give 3SH.to
'Give the rice to him'
- f. *fe ngié*
OM3SNH.give 3SH.to
'Give it to him'

However, this type of organization of the construction of transfer verbs does not seem very common in African languages, particularly in the Niger-Congo phylum.¹⁰

6.2 Indexation systems in which object markers attached to transfer verbs can represent the recipient, but not the transferred thing

Mende, which has served to illustrate a type of construction of transfer verbs in which the transferred thing is assimilated to the patient of typical transitive verbs and the recipient is treated as an oblique – ex. (15), has another verb ‘give’ with a construction in which the recipient is assimilated to the patient of prototypical transitive verbs, and the transferred thing treated as an oblique – ex. (16); in this construction, the recipient, but not the transferred thing, is represented by the same object markers as the patient of typical transitive verbs.

(16) MENDE (Innes 1971)

- a. *kpanâ gɔ́ a mbeí*
Kpana give with rice
‘Give the rice to Kpana’
- b. *kpanâ gɔ́ la*
Kpana give with-it
‘Give it to Kpana’
- c. *ngi-gɔ́ a mbeí*
OM3SH-give with rice
‘Give the rice to him’

Ex. (16) illustrates a type of indexation of the arguments of transfer verbs particularly common in African languages. Two subtypes may be recognized. In most cases, there is no possibility of indexing the transferred thing by means of special pronominal markers attached to the verb. This may be the case, not only in constructions in which the noun phrase representing the transferred thing clearly has the characteristics of an oblique, as in example (16), but also in constructions in which there is no obvious indication of its oblique status, as in Swahili – ex. (17): in example (17d), the two noun phrases that follow the verb are neither case marked nor combined with adpositions, but only one of them can be indexed in the verb form.

(17) SWAHILI

- a. *ni-me-wa-ona watoto*
SM1S-TAM-OMC2-see CL2.child
‘I have seen the children’
- b. *ni-me-ki-leta chakula*
SM2S-TAM-OMC7-bring CL7.food
‘Have you brought the food?’ (which I told you to bring)
- c. *ni-me-wa-pa chakula*
SM1S-TAM-OMC2-give CL7.food
‘I have given food to them’

- d. *ni-me-wa-pa* *watoto* *chakula*
 SM1s-TAM-OMC2-give CL2.child CL7.food
 'I have given food to them'
- e. **ni-me-ki-pa* *watoto*, **ni-me-ki-wa-pa*

A less common subtype is illustrated in Shimaore. In this language – ex. (18), in the same way as in Swahili, transfer verbs include a unique object marker identical to those used to represent the patient of typical transitive verbs, and this object marker necessarily represents the recipient; but Shimaore has a third set of pronominal markers that occupy a special position at the end of the verb form and are used specifically with ditransitive verbs to represent the second object. Note that, in Bantu languages, pronominal markers of this type are a particular case of 'oblique argument markers', since formally similar morphemes occupying the same position in the verb form are commonly used to represent locative arguments.

(18) SHIMAORE

- a. *ni-tso-hu-zunguha*
 SM1s-FUT-OM2s-look+for
 'I'll look for you'
- b. *ni-tso-m-zunguha*
 SM1s-FUT-OMC1-look+for
 'I'll look for him/her'
- c. *ni-tso-li-zunguha*
 SM1s-FUT-OMC5-look+for
 'I'll look for it (cl.5)'
- d. *ni-tso-zi-zunguha*
 SM1s-FUT-OMC10-look+for
 'I'll look for it/them (cl.10)'
- e. *ni-tso-m-ba* *Haladi* *zimarke*
 SM1s-FUT-OMC1-give CL1.Haladi DEF.CL10.money
 'I'll give the money to Haladi'
- f. *ni-tso-hu-βa* *ligari*
 SM1s-FUT-OM2s-give DEF.CL5.car
 'I'll give you the car'
- g. *ni-tso-m-ba-zo* *Haladi*
 SM1s-FUT-OMC1-give-xMC10 CL1.Haladi
 'I'll give it to Haladi (the money)'
- h. *ni-tso-hu-βa-lo*
 SM1s-FUT-OM2s-give-xMC5
 'I'll give it to you (the car)'

Among languages in which transfer verbs can incorporate a unique object marker that necessarily represents the recipient, the case of Kanuri is of special interest, since this language shows a split between the case assigning properties of ditransitive verbs

and their indexation properties. In Kanuri, the object of typical transitive verbs may optionally be marked by the case marker ('accusative') *-ga*, and the complement of transfer verbs that represents the recipient obligatorily takes the case marker *-ro*, functionally similar to the 'dative case' of Indo-European languages, which suggests classifying Kanuri among the languages that have a special syntactic function typically used to encode the recipient of transfer verbs. But Kanuri has no dative markers similar to those found in the Romance languages, and transfer verbs, like typical transitive verbs, can take a unique object marker representing the recipient (i.e. corresponding to a noun phrase in the dative case), which points to an organization of the Swahili type – ex. (19). However, it must be observed that this particularity of the transfer verbs of Kanuri is consistent with the animacy properties of the arguments of transfer verbs (in unmarked situations, an inanimate thing is transferred to an animate recipient) and with the fact that Kanuri has object markers for the first and second person only.

(19) KANURI (Cyffer 1991)

- a. *shí-ga cítáko*
he-OBJ PAST.seize.MS1S
'I seized him'
- b. *agógó shí-ro cóko*
watch he-DAT PAST give.MS1S
'I gave him a watch'
- c. *nyí-ga njítáko*
you-OBJ OM2S.PAST.seize.MS1S
'I seized you'
- d. *agógó nyí-ro njóko*
watch you-DAT OM2S.PAST.give.MS1S
'I gave you a watch'

6.3 Indexation systems in which object markers attached to transfer verbs can equally represent the recipient or the transferred thing

This type of indexation of the arguments of transfer verbs occurs in so-called double object constructions. In such constructions, both the noun phrase representing the recipient and the noun phrase representing the transferred thing show some grammatical characteristics similar to those of the object of typical transitive verbs, but the two objects never show the grammatical properties of objects to exactly the same degree, and the noun phrase representing the recipient can be recognized as 'first object' in the sense that it shows more object-like properties than the noun phrase representing the transferred thing ('second object').¹¹

As regards indexation, a first possibility, illustrated here by Tswana – ex. (20) – and Wolof – ex. (21),¹² is that ditransitive verbs can simultaneously receive two object markers identical to those that are used to represent the patient of typical transitive verbs.

(20) TSWANA

- a. *kì-χò-bóǀnì*
SM1s-OM2s-see.TAM
'I've seen you'
- b. *kì-lú-rékìlè*
SM1s-OMC11-buy.TAM
'I've bought it (the lamp)'
- c. *kì-lú-χò-filè*
SM1s-OMC11-OM2s-give.TAM
'I've given it to you (the lamp)'

(21) WOLOF

- a. *Dama-y jox ganaar gi dugub ji*
VFOC.SM1s-TAM give hen DEF millet DEF
'I'm giving the millet to the hen'
- b. *Dama-ko-ko-y jox*
VFOC.SM1s-OM3s-OM3s-TAM give
'I'm giving it to it'

In Tswana, the relative ordering of the object markers is the reverse of that of the corresponding noun phrases: the object noun phrase that represents the argument whose role has the strongest affinity with the feature +animate must immediately follow the verb, and the corresponding object marker must be immediately prefixed to the verb stem. As regards Wolof, it is interesting to observe that the double object constructions of this language have no strict ordering of the two noun phrases in object function; by contrast, the ordering of the object markers is strict, but it is independent from the roles of the participants they represent and depends exclusively on the hierarchy '1st/2nd person > 3rd person plural > 3rd person singular', as illustrated by ex. (22).

(22) WOLOF

- a. *Dama-y jox xale bi mango yi* → *Dama-leen-ko-y jox*
VFOC.SM1s-TAM give child DEF mango DEF.PL 'I'm giving them to him'
'I'm giving the mangoes to the child'
- b. *Dama-y jox xale yi mango bi* → *Dama-leen-ko-y jox*
VFOC.SM1s-TAM give child DEF.PL mango DEF 'I'm giving it to them'
'I'm giving the mango to the children'

Another type of indexation of the objects in double object constructions is observed for example in Southern Sotho – ex. (23). In this type, ditransitive verbs cannot take more than one object marker at the same time, but this object marker may correspond to the first or to the second object.

(23) SOUTHERN SOTHO

- a. *hà-kí-ò-bónì*
NEG-SM1S-OM2S-see
'I don't see you'
- b. *hà-kí-lí-bónì*
NEG-SM1S-OMC5-see
'I don't see it (the broom)'
- c. *hà-kí-ò-fĩ* *liffèlò*
NEG-SM1S-OM2S-give CL5.broom
'I don't give you the broom'
- d. *hà-kí-lì-fĩ* *mòsádì*
NEG-SM1S-OMC5-give CL1.woman
'I don't give it (the broom) to the woman'
- e. **ha-kí-lí-ò-fĩ*, **ha-kí-ò-lí-fĩ*¹³

However, it must be observed that such systems of indexation of the arguments of transfer verbs are never perfectly symmetric and always show particularities that confirm the hierarchy between first and second object.

In Tswana – ex. (24), either object may be promoted as the subject of a passive form, but when the subject of a passive form represents the recipient, the transferred thing can be represented by an object marker, whereas it is impossible to introduce an object marker representing the recipient in a passive form whose subject represents the transferred thing.

(24) TSWANA

- a. *kì-filé* *bàná* *lòkwálò*
SM1S-give.TAM CL2.children CL11.book
'I've given a book to the children'
- b. *kì-lú-bà-filè*
SM1S-OMC11-OMC2-give.TAM
'I've given it (the book) to them'
- c. *bàná* *bá-filwé* *lòkwálò*
children SMC2-give.PSV-TAM CL11.book
'The children were given a book'
- d. *bá-lò-filwè*
SMC2-OMC11-give.PSV.TAM
lit. 'They were given it'
- e. *lòkwálò* *lú-filwé* *bàná*
book SMC11-give.PSV.TAM CL2.child
'The book was given to the children'
- f. **lú-bà-filwé*

In Southern Sotho (25), when both objects are pronominalized at the same time, the first object has priority to occupy the only available object marker slot, and the second object is necessarily represented by a free pronoun following the verb.

(25) SOUTHERN SOTHO

- a. *kì-fá bàsádí lifièlò*
SM1s-give CL2.woman CL5.broom
'I give the broom to the women'
- b. *kì-bá-fá lifièlò*
SM1s-OMC2-give CL5.broom
'I give them the broom'
- c. *kì-lí-fá bàsádí*
SM1s-OMC5-give CL2.woman
'I give it the women'
- d. *kì-bá-fà lóná*
SM1s-OMC2-give PROC5
'I give it to them'

6.4 The particular case of serializing languages

Serializing languages do not fit straightforwardly into the typology presented in the preceding sections, since they tend to encode events involving three participants by means of combinations of two verbs. However, in languages commonly considered as typical serializing languages in which it is possible to identify a verb 'give' in a construction that involves no other verb, this construction belongs generally to the type in which the recipient, but not the transferred thing, is assimilated to the patient of typical transitive verbs.

In serializing languages, the fact that 'give' commonly functions also as the second term of serial constructions in which it takes a unique complement representing a recipient or a beneficiary can be viewed as an evidence of the predominance of the complement representing the recipient in the construction of 'give' as a ditransitive verb – ex. (26) & (27).

(26) KPOSO (Eklo 1987)

- a. *kúmá á-ká kɔ́kú ítùkpá*
Kuma SM3s-give Koku goat
'Kuma gave Koku a goat'
- b. *kúmá á-jɔ́ ítùkpá ká kɔ́kú*
Kuma SM3s-take goat give Koku
'Kuma gave Koku a goat'
- c. *kúmá á-ufè ègà ká kɔ́kú*
Kuma SM3s-lend money give Koku
'Kuma lent money to Koku'

(27) YORUBA

- a. *Òjọ́ fún ìyá ní owó*
Ojo give mother PREP money
'Ojo gave mother money'
- b. *Òjọ́ fún-un ní owó*
Ojo give-OM3s PREP money
'Ojo gave her money'
- c. *Òjọ́ rà ìwé fún ìyá*
Ojo buy book give mother
'Ojo bought a book for mother'
- d. *Òjọ́ rà-á fún-un*
Ojo buy-OM3s give-OM3s
'Ojo bought it for her'

Conclusion

The main tendencies observed in African languages regarding subject/object indexation can be summarized as follows:

(1) An overwhelming majority of African languages have bound morphemes analyzable as subject markers; in many cases, their obligatoriness or their tendency to fuse with TAM or polarity markers facilitates their analysis as bound morphemes, but stage I subject markers difficult to distinguish from free pronouns are relatively common in Africa.

(2) Languages that have subject markers only (i.e. languages that can pronominalize subject by means of bound morphemes but use exclusively free pronouns in object pronominalization) are relatively uncommon in Africa, but obligatory agreement of transitive verbs with their object is not a common phenomenon in African languages, and in many languages, the distinction between object markers and free pronouns in object function is even more difficult to establish than the distinction between subject markers and free pronouns.

(3) 'Exotic' patterns of subject / object indexation on transitive verbs (ergative, agentive, direct / inverse, etc.) are very rare in African languages, which confirms a general tendency of African languages towards types of morphosyntactic organization in which the traditional notions of subject and object can be recognized in a relatively unproblematic way.

(4) The indexation of the arguments of typical ditransitive verbs confirms that African languages show a strong tendency to assimilate the recipient to the patient of typical transitive verbs. In 'double object constructions', object markers identical to those that refer to the patient of typical transitive verbs, when attached to transfer

verbs, can indistinctly refer to the transferred thing or to the recipient; but in most cases, object markers identical to those that refer to the patient of typical transitive verbs, when attached to transfer verbs, necessarily refer to the recipient, not to the transferred thing.

Notes

1. In the first person plural, a distinction between ‘we including you’ and ‘we excluding you’ occurs sporadically in several groups of African languages. As a rule, additional distinctions in the third person are encountered in languages with a gender system in which identical distinctions are involved in the agreement between nouns and modifiers. With gender systems of the Niger-Congo type (traditionally referred to as ‘noun class systems’), gender distinctions are found in the third person only. In gender systems based on the sex distinction, gender distinctions may be found in the second person too. Note that the correlation between ‘nominal gender’ and ‘pronominal gender’ is not absolute: one may encounter languages, either with gender-like distinctions in pronouns and/or pronominal markers only, or languages with gender distinctions manifested at the level of the relation between the noun and its modifiers that do not extend to pronouns and/or pronominal markers. For example, Wolof has noun class distinctions at the noun phrase level, but these distinctions do not manifest themselves in the variations of free pronouns or of subject and object markers. Conversely, Zande is devoid of any gender distinction at the noun phrase level, but in the third person, the free pronouns and the subject markers of Zande have different forms for masculine human, feminine human, non-human animate, and inanimate.

2. Historically, it seems reasonable to analyze these consonant alternations as the reflex of the presence vs. absence of an ancient prefix **n-*. Note that Fula has also stage I subject markers, illustrated here in example (1c–d), which are in complementary distribution with NPs in subject function.

3. Abbreviations used in the glosses:

CLX	class X	PREP	preposition
DAT	dative	PROCX	pronoun class X
DEF	definite	PSV	passive
FUT	future	SG	singular
INDEF	indefinite	SM	subject marker
NEG	negation	SMCX	subject marker class X
OBJ	object	SM1s	subject marker 1st person singular
OM	object marker	SM3P	subject marker 3rd person plural
OMCX	object marker class X	SM3s	subject marker 3rd person singular
OM1s	object marker 1st person singular	SM3SM	subject marker 3rd person singular masculine
OM2s	object marker 2nd person singular	TAM	tense-aspect-mood marker
OM3s	object marker 3rd person singular	TOP	topic
OM3SH	object marker 3rd person singular human	VFOC	focalization of the verb
OM3SNH	object marker 3rd person singular non-human	XMCX	oblique argument marker class X
PL	plural	1s	1st person singular
		3SH	3rd person singular human

4. Mende is an SOVX language with subject markers attached to the first word of the verb phrase (i.e. the first word of the NP in object function, if any) – see Section 4.
5. The subject markers of Wolof always have the same form in 1st person plural and 3rd person plural.
6. In the presentation of this example, I consider that *dañu naan* is a compound verb form with the subject marker included in a word that can be viewed as an auxiliary. This analysis, which simplifies the formulation of the rules accounting for the attachment of subject and object markers, is supported by the invariability of the verb stem, which in Wolof suffers no exception. An alternative analysis would be to analyze the ‘tense-person complex’ *dañu* as prefixed to the verb stem, but what is important here is that both analyses recognize that the pronominal morphemes occurring in (6c) are pronominal markers rather than free pronouns.
7. Note however that, even in cases when the relevant facts are established in a precise and complete way, difficulties in identifying the exact status of pronominal morphemes may persist, due to the fact that clear evidence supporting the identification of pronominal morphemes as bound morphemes may appear only in certain conditions; for example, the clearest evidence that the ‘weak object pronouns’ of Hausa are in fact verb suffixes is that they undergo a tonal alternation conditioned by the tone of the preceding syllable, but this alternation operates only with certain verb classes. It may also happen that within the same set of pronominal morphemes, some show more clearly than others the behavior of bound morphemes; for example, in Yoruba, the bound nature of the object marker suffixed to the verb is much more obvious in the third person singular than in the other persons, since the object marker of third person singular has no stable segmental form and is realized as a copy of the preceding vowel; in Manding, the phonological behavior of the ‘weak pronouns’ of first person singular and third person singular provides much more evidence supporting their identification as bound morphemes than in the other persons.
8. At least in some languages, there is a relation between the choice of this element and morphological variations of the verb. Historically, at least some of these ‘predicative markers’ may originate from auxiliary verbs, but synchronically, most of them show no evidence of a verbal status. It is also worth noting that sometimes (but not always) their phonological interaction with the context suggests analyzing them as bound to the last word of the subject noun phrase, or to the first word of the verb phrase; but this is not directly relevant to the present discussion.
9. In Wolof, the object markers of 2nd person plural and 3rd person plural are identical.
10. The case of Hausa may be interesting to mention here. In this language, the verb *baà* ‘give’ has very clearly a construction in which the recipient is treated exactly in the same way as the patient of typical transitive verbs:

<i>yaa</i>	<i>baà</i>	<i>Audu</i>	<i>àbinci</i>
SM3SM.TAM	give	Audu	food
‘He gave food to Audu’			
<i>yaa</i>	<i>baa-nì</i>	<i>àbinci</i>	
SM3SM.TAM	give-OM1s	food	
‘He gave me food’			

By contrast, the other verbs of transfer have a construction currently analyzed as a construction in which the recipient is treated as the complement of a preposition *wà* ~ *ma-*:

yaa kaawoo wa Audù àbinci
 SM3SM.TAM bring PREP Audu food
 'He brought food to Audu'

yaa kaawoo mi-nì àbinci
 SM3SM.TAM bring PREP-1s food
 'He brought food to Audu'

But according to this interpretation, the Hausa verbs of transfer would have very strange properties, since they would be separated from their 'direct' object by a prepositional object. The lack of mobility of the 'preposition' involved in this construction suggests reanalyzing it as a verbal suffix, which would lead to reanalyze this construction as a 'normal' double object construction similar to that of *baà*. Note however that Newman (2000) provides some evidence against this reanalysis.

11. In the discussion of double object constructions, the traditional terms of direct / indirect object are particularly misleading, since both objects are in some sense 'direct', and the one that fully assimilates to the unique object of typical transitive constructions is not the one traditionally recognized as 'direct'.

12. In the 3rd person, Tswana has 12 different object markers according to the class of the corresponding noun, whereas Wolof has only 2 (singular and plural), but these languages have in common that, with ditransitive verbs, two object markers of the same paradigm can be attached to the same verb.

13. In Southern Sesotho, 'I don't give it (the broom) to you' can only be *hà-kí-ò-fí lònà*, with the free pronoun *lònà* representing the transferred thing (see below).

References

- Amha, A. (2001). *The Maale Language*. Leiden: Center for Non-Western Studies.
- Cyffer, N. (1991). *We Learn Kanuri*. Köln: Rüdiger Köppe Verlag.
- Doneux, J. (n.d.). *Exercices de langue dan*. Lyon: Afrique et langage.
- Eklo, A. A. (1987). *Le kposso de Tomegbe (Togo)*. PhD Dissertation, Grenoble: University of Grenoble.
- Eze, E. (1995). The forgotten null subject of Igbo. In A. Akinlabi (Ed.), *Theoretical Approaches to African Linguistics* (pp. 59–81). Trenton: Africa World Press.
- Griegenow-Mewis, C., & Bitima, T. (1994). *Lehrbuch des Oromo*. Köln: Rüdiger Köppe Verlag.
- Heath, J. (1999). *A Grammar of Koyraboro (Koroboro) Senni*. Köln: Rüdiger Köppe Verlag.
- Ikoru, S. M. (1996). *The Kana language*. Leiden: Research School CNWS.
- Innes, G. (1971). *A Practical Introduction to Mende*. London: SOAS.
- Kutsch Lojenga, C. (1994). *Ngiti, a Central Sudanic language of Zaire*. Köln: Rüdiger Köppe Verlag.
- Launey, M. (1981). *Introduction la langue et la littérature aztèque*. Paris: L'Harmattan.
- Newman, P. (2000). *The Hausa Language: An encyclopedic reference grammar*. New Haven, CT: Yale University Press.
- Quaireau, A. (1987). *Description de l'agni*. PhD Dissertation, University of Grenoble.
- Reh, M. (1996). *Anywa Language, Description and Internal Reconstructions*. Köln: Rüdiger Köppe Verlag.

- Rojas Curieux, T. (1998). *La lengua páez*. Bogotá: Ministerio de Cultura.
- Schadeberg, T. C. (1985). *A Short Sketch of Ewe*. Köln: Institut für Afrikanistik.

Head marking, dependent marking and constituent order in the Nilotic area^{*}

Gerrit J. Dimmendaal

Universität zu Köln

*A language can do whatever it wants to with
whatever material it has to hand, if it wants to*
Goddard's Law

Central to the present study is the Nilotic branch within Nilo-Saharan. It is argued first (in Section 2) that the Nilotic branch is somewhat exceptional for Nilo-Saharan as a whole in that many of its members are verb-initial with a strong tendency towards head marking, rather than dependent marking, as clausal strategies. As argued next (in Section 3), there is evidence for a historical drift in Nilotic towards loss of dependent marking at the clausal level. In Section 4 it is shown how one group of closely related Eastern Nilotic languages has redeveloped dependent-marking strategies while maintaining head marking at the clausal level. As further shown in the same section, a similar combination of strategies is attested in neighboring Surmic languages, which also belong to Nilo-Saharan; these shared typological properties must be the result of areal diffusion. As argued next (Section 5), the convergence in the Nilotic-Surmic twilight zone discussed in Section 4 is the outcome of long-term multilingualism in the area. Section 6 presents some additional thoughts on the kind of data needed in order to deepen our understanding of areal types and areal diffusion in these languages.

1. Head-marking and dependent-marking languages

It is a well-known fact, ever since Sapir's cross-linguistic survey of morphosyntactic coding mechanisms (Sapir 1921:120–146), that languages may differ widely in their strategies for expressing syntactic relations. Cohesion at the clause level may be expressed on the head of the clause, i.e. the verb, on dependent categories, or on both types of categories; alternatively, when neither head marking nor dependent marking occurs, languages may use fixed constituent order as a coding strategy for syntactic and semantic relations. Nichols (1986:104) has argued, on the basis of a cross-linguistic survey, that "...the head-marked clause pattern favors verb-initial word order...the

types having a strong dependent-marked component in their grammar favor verb-final pattern. . .” As further observed by Nichols (1986:71), the dependent-marking languages show more dispersion with respect to these alternative coding strategies than the head-marking languages, in that many otherwise dependent-marking languages have verbal agreement with one or even two arguments, while few of the strongly head-marking languages apparently show any analogous preference for a particular dependent-marked pattern.

Corroborating evidence for Nichols’ claims concerning coding strategies and constituent order can be found in verb-final and verb-initial languages of northeastern Africa. In Maale, an Omotic (i.e. Afroasiatic) language of Ethiopia, there is only dependent marking in main clauses. An example from Azeb Amha (2001:59):

- (1) *ʔííni fooc'-atsí-m goys'-ó daww-é-ne*
 3M:SG:NOM guest-M:ABS-DAT road-ABS show-PERF-A:DCL
 ‘He showed the road to the guest’

Most other Omotic languages, however, tend to have pronominal subject markers on the verb in addition to the dependent-marking strategies.¹ (For a general survey of the syntactic structure of Omotic languages see Hayward 1990.)

In verb-initial Nilotic (i.e. Nilo-Saharan) languages such as Kipsikiis, spoken in Kenya, one finds a highly developed head-marking pattern at the clausal level (data from Toweett 1979:163):

- (2) *kii.kāt-cíínéecín-äänéèùun làakwéet*
 INF.greet-DAT:DAT-VEN:VEN child:ABS
 ‘To pass greetings for the child to him/her as one moves towards the speaker’

At the same time Kipsikiis and other verb-initial Nilotic languages do have a certain degree of dependent marking, in that postverbal subjects are distinguished from other constituents (such as objects) by way of case marking. In sharp contrast with Omotic, however, most verb-initial Nilotic languages lack peripheral case marking. (For the distinction between core and peripheral case, see Blake 1994.) What appears to be crucial for a typological classification in terms of syntactic coding strategies therefore is the formal expression of peripheral semantic roles such as direction, location, or instrument, either through peripheral case marking, as in Omotic, or by way of verbal extensions, as in many Nilotic languages. Both language types may use these alternative strategies in tandem with a third strategy, namely adpositions.

2. Areal types in northeastern Africa

2.1 The verb-final bond and dependent marking

In his typological survey of verb-final south Asian languages, Masica (1976) suggested that Ethiopia might form an extension of the “Indo-Altaic” verb-final bond.

Table 1. Dependent-marking in Nilo-Saharan

Language group	Constituent Order	Periph. Case*	ProSu	ProOb
Saharan	V-final	<i>yes</i>	<i>yes</i>	<i>yes</i>
Maban	V-final	<i>yes</i>	<i>yes</i>	<i>yes</i>
Fur	V-final	<i>yes</i>	<i>yes</i>	<i>no</i>
Kunama	V-final	<i>yes</i>	<i>yes</i>	<i>yes</i>
Eastern Sudanic				
Nubian	V-final	<i>yes</i>	<i>yes</i>	<i>no</i>
Tama	V-final	<i>yes</i>	<i>yes</i>	<i>no</i>
Nyimang	V-final	<i>yes</i>	<i>no</i>	<i>no</i>

* Peripheral case: Dative, Instrument, Locative, Ablative, Genitive

The core of this area in Ethiopia is formed by Cushitic and Omotic languages; neighboring Ethiopian Semitic languages are known to have converged historically towards their genetically distant relatives (compare, for example, Leslau 1945). As argued by Tosco (2000), the notion of an “Ethiopian language area” as such is false, but the author agrees that SOV constituent order is a good example of an areal feature “...attested...well outside Ethiopia as far south as Tanzania by Iraqw” (p. 344). As observed by Heine (1976), there are also various Nilo-Saharan languages mainly west of this area, and extending all the way towards Chad, with a verb-final clause structure plus the usual concomitant features. Today, these various languages belonging to, what Heine has called, the type-D languages do not constitute a geographically contiguous area. It is important to note, however, that virtually no other languages or language types are spoken between these at times geographically isolated Nilo-Saharan and Afroasiatic languages. The relative isolation of these languages in particular in the central and eastern Sahel region most likely is an outcome of the gradual desertification of the region over the past 5,000 years, a process which forced people to retreat towards more mountainous regions where there was still water available, such as the border area between Sudan and Chad. Historically, then, the verb-final Afroasiatic and Nilo-Saharan languages in the area in all likelihood constituted a large, geographically contiguous convergence zone.

Similar to verb-final Afroasiatic languages in eastern Africa, the verb-final Nilo-Saharan languages in regions further towards the west show a strong propensity for dependent marking at the clausal level, with a certain degree of head marking (for pronominal subjects, and, occasionally, objects), as Table 1 helps to show.

As we shall see next, a number of Nilo-Saharan groups, spoken mainly towards the south of this verb-final bond, deviate from this pattern, in that they either have a verb-initial syntax, or verb-second properties.

2.2 Verb-initial languages and head marking

The most extensive bond of verb-initial (or, in terms of Heine’s 1976 typology, type-C) languages in the area is formed by Southern and Eastern Nilotic languages. A num-

ber of neighboring Surmic languages also have a verb-initial syntax. Surmic, a genetic unit with representatives in Sudan and Ethiopia, probably is Nilotic's closest relative, as argued by a number of authors (Ehret 1983; Dimmendaal 1988; Ehret 2001). The neighboring Kuliak languages of Uganda are also verb-initial. They have been classified as Eastern Sudanic by Greenberg, the larger genetic grouping within Nilo-Saharan to which also Nilotic and Surmic belong; similarly, Ehret (2001:89) also assumes these languages belong in one subgroup, which he calls Eastern Sahalian, rather than Eastern Sudanic. Further towards the south in northern Tanzania, the Hadza language (classified as Khoisan by Greenberg, but possibly constituting a linguistic isolate according to other scholars) also has a verb-initial syntax. In addition, the Kadu languages in the Nuba Mountains of central Sudan, have a verb-initial syntax. The Nuba (Kordofan) mountains probably constituted an area of refuge for a considerable period of time, as attested by the great number of languages and language families in this area, as well as the typological divergence between them; for example, whereas the Kadu languages are verb-initial, the Nilo-Saharan language Nyimang is verb-final.

As this short survey helps to show, verb-initial languages in (north)eastern Africa are more dispersed geographically than verb-final languages in the region. Given the fact that the former are also interspersed with a variety of distinct genetic groupings which are also typologically more diversified, it is more difficult to prove that these verb-initial languages together constituted a typological zone at one point.

It is common for verb-final Eastern Sudanic language groups such as Tama and Nubian, or more distantly related Nilo-Saharan language groups such as Kunama and Fur, to leave the syntactic subject morphologically unmarked (i.e. Nominative being expressed through zero marking) and to mark the object with (Accusative) case. Nilotic and Surmic languages deviate from these remaining Nilo-Saharan groups in their case-marking systems, in that in the former two groups the subject is marked for (Nominative) case, whereas the object lacks case inflection (or takes 'zero case' marking). Such case-inflected subjects in Nilotic and Surmic languages always follow the verb. The same languages using postverbal case inflection for subjects either have a verb-initial syntax, or the verb occurs in second position (allowing for OVS or Topic VS as a basic structure). This case inflection for post-verbal subjects may be a shared innovation of Nilotic and Surmic (rather than being due to area diffusion), given the fact that cognate case suffixes are involved (cf. Dimmendaal 1998:41). In some of these Nilotic and Surmic languages subjects of transitive verbs follow the latter in basic sentences, whereas subjects of intransitive constructions precede the verb in utterances not marked for any discourse context. Since only postverbal subjects receive morphological case marking in these languages, they may be argued to have ergative properties. (See, for example, Andersen 1988 on Pāri; Reh 1996 on Anywa; Last & Lucassen 1998 on Chai; Miller & Gilley 2001 on Shilluk.) But, whereas the Surmic languages involved also have peripheral case marking, most Nilotic languages with case-marking only distinguish between Nominative (or Ergative) case and Absolutive case; the latter case form is used for (indirect) objects, preverbal subjects, complements of prepositions (with few exceptions), as well as for nouns or noun phrases (and pronouns) in isola-

tion. As argued next, there is historical-comparative evidence that this more reduced system of dependent marking in Nilotic presents an innovation.

3. The drift towards head marking in Nilotic

Nilotic case systems belong to either the Nominative-Absolutive or the Ergative-Absolutive type, depending on whether transitive and intransitive predicates behave differently from each other, as pointed out above. The formal case marking of postverbal subjects is attested in all three primary subgroups of Nilotic, as Table 2 shows.

The subgrouping for each of the three primary branches of Nilotic is based on historical-comparative studies by Vossen (1982) for Eastern Nilotic, Rottland (1982) for Southern Nilotic. Their typological (as against their genetic) classification, based on the order of meaningful units, would seem to be non-controversial. The classification as V2- languages (i.e. as verb-second) languages is based on the distributional fact that different constituents may precede the verb in a number of these languages. In Western Nilotic Dinka, for example, sentences tend to start with a topic, with the verb occurring in second position, followed by the subject or some other constituent; verb-initial structures are also possible, as shown by Andersen (1991).² For a number of other Western Nilotic languages, all spoken in the southern Sudan, it has been argued that OVS constitutes the basic order (cf. Andersen 1988 on Pāri; Reh 1996 on Anywa; Miller & Gilley 2001 on Shilluk); alternatively, we find SVO order, as in Acholi, Lango or Luo. What appears to be excluded in these Western Nilotic languages is a syntactic order whereby both the subject and the object follow the verb, i.e. VSO

Table 2. Case marking in Nilotic

LANGUAGE GROUP	CONST.ORD	NOM/ERG POSTV. SU	ABS OB	PERIPH.CASE*	PROSU	PROOB
Western Nilotic:						
Anywa	V2/OVS	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>yes</i>
Dinka	V2	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>no</i>
Pāri	OVS	<i>yes</i>				
Shilluk	V2/OVS	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>no</i>
Luo	SVO	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>yes</i>
Eastern Nilotic:						
Bari group	SVO	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>
Lotuxo	SVO	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>
Ongamo-Maa	V-initial	<i>yes</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>yes</i>
Southern Nilotic:						
Kalenjin	V-initial	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>yes</i>
Datooga-Omotik	SVO	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>yes</i>

*Peripheral case: Dative, Instrument, Locative, Ablative, Genitive

or VOS; similar verb-second properties are attested in neighboring Surmic languages (Dimmendaal 1998).

In Nilotic languages where the subject occurs post-verbally in basic sentences not marked for any discourse context these subjects take morphological case. Preverbal subjects on the other hand are not inflected for case, or phrased differently, preverbal subjects take the morphologically unmarked case form, the Absolutive; this case form is used with nouns (or noun phrases, as well as pronouns) functioning as objects, as complements of prepositions, or nouns in isolation. In all three primary branches of Nilotic, there are language with SVO order, e.g. the Eastern Nilotic Bari group, the Southern Nilotic Omotic-Datooga group, and Western Nilotic languages such as Acholi or Luo (compare also Table 2). As may be expected, these languages lack case marking for subjects (since only postverbal subjects take morphological case). This situation must be due to more recent innovations, for a number of reasons. First, according to a well-attested universal, verb-initial languages allow for SVO-order as an alternative. The latter order accordingly can be explained as a natural drift from an erstwhile verb-initial syntax. Moreover, this way one also has a natural explanation for the absence of case marking in these Nilotic SVO languages, since only post-verbal subjects take morphological case in Nilotic. Once this latter strategy disappears, case marking also vanishes as a coding device. It is noted further that cross-linguistically SVO languages do not necessarily allow for a verb-initial structure as an alternative; consequently, if one assumes an original constituent order whereby subjects preceded the verb, one would need to explain separately why different Nilotic languages developed into verb-initial languages. Also, there is no principled reason why such post-verbal subjects rather than objects should be inflected for case. Nevertheless, all verb-second and verb-initial Nilotic language allowing for (case-marked) post-verbal subjects do have this property, a feature therefore best explained as a shared innovation, and following from the fact that the subject could follow the verb in their common ancestral language. As shown for the Eastern Nilotic language Turkana in Dimmendaal (1983:261), there is language-internal evidence that the tonal inflection for the Nominative goes back to a suffix. Western Nilotic languages like Anywa or Pāri in fact do have such case suffixes used with postverbal subjects; in the very same languages, a few nouns are inflected for Nominative/Ergative case through tonal inflection. As is common across the African continent, inflection or derivation by way of tone usually is due to the loss of segmental structures. And here a third argument may be found for a hypothesis which takes the case inflection for postverbal subjects in Nilotic to be old, probably even predating Proto-Nilotic: The case markers for postverbal subjects in Nilotic appear to be cognate with case suffixes for postverbal subjects as found in the closely related Surmic languages.

All these factors point towards a shared historical innovation of Surmic and Nilotic involving the placement of the syntactic subject after the verb, whereby the same subject was inflected for case. Preverbal subjects, on the other hand, were not inflected for case, i.e. they took Absolutive case, i.e. the same case form as nouns or noun phrases functioning as objects, complements of prepositions or nouns and noun phrases oc-

Table 3. Nominative/Ergative case markers in Surmic and Nilotic

Absolutive Nominative/Ergative		
Surmic:		
Baale	unmarked	-(j)ε, -(j)i (sg)
Tennet	unmarked	-ε (sg), -i (pl)
	-a (pl)	
Murle	unmarked	-ε (sg), -i (sg)
	-a (pl)	
Majang	unmarked	-ε
Nilotic:		
Anywa*	unmarked	-ε, -(C)I
Päri	unmarked	tone, -ε, -i

*Reh (1996) treats these suffixes as definiteness markers

curing in isolation. This system is still found today in all Surmic and Nilotic languages involved. An alternative explanation for the marking of postverbal subjects in these languages in terms of areal diffusion from Surmic into Nilotic or vice versa cannot be excluded, but would appear to be extremely unlikely, since there is no evidence either for lexical or for grammatical borrowing in general between Nilotic as a whole and Surmic. Moreover, it would leave the other common properties discussed above unaccounted for.

Whereas Nominative (or Ergative) case marking must be an archaic property of Nilotic and Surmic languages, only the Surmic group has a productive system of peripheral case marking (similar to other Nilo-Saharan groups, as pointed out above). Nilotic languages in general lack peripheral case marking, i.e. case marking is not used in order to express semantic roles such as location, direction or instrument. Interestingly, however, remnants of Locative case marking are found throughout this family, thereby providing conjectural evidence that such a system existed in pre-Nilotic. Thus, in Western Nilotic Nuer, the locative form of specific nouns – usually for singular nouns only – is distinct from the so-called citation form, which is used for subject and object (data from Crazzolara 1933:29):

- (3) Citation Locative
- | | | |
|------------|-------------|----------|
| <i>lêp</i> | <i>lêb</i> | ‘tongue’ |
| <i>lôc</i> | <i>lôci</i> | ‘heart’ |

With certain other nouns in Nuer, a suffix -ɔ appears when such nouns have a locative meaning.

Similarly, in Eastern and Southern Nilotic languages remnants of Locative case marking are found in a few singular forms of nouns typically associated with the homestead and its surrounding areas (e.g. the words for ‘house’, ‘field’, ‘mountain’, ‘well’). Compare the following examples:

		Nominative/Absolutive	Locative
Eastern Nilotic			
Maasai:	<i>enk-aŋ</i>	<i>aŋ</i>	‘home, boma’
Turkana:	<i>a-mana</i>	<i>mana</i>	‘field’
	<i>a-kar(ε)</i>	<i>karε</i>	‘well’
Southern Nilotic			
Kipsikis	<i>kɔ</i>	<i>ka</i>	‘house’

Eastern and Southern Nilotic languages, and to a lesser extent Western Nilotic languages, use verbal strategies, i.e. head marking, in order to express semantic roles such as direction, location, or instrument. Compare the following examples from Eastern Nilotic Maasai (data from Tucker & Mpaayei 1955), and Nandi (data from Creider & Creider 1989)

Maasai:

- (4) *a-dót-ú nkójít*
1SG-pull-VEN grass:ABS
‘I pull out grass’
- (5) *á-írrág-áá Náròk*
1SG-sleep-IT Narok:ABS
‘I sleep at Narok’
- (6) *á-ból-ókì papá ólbéné*
1SG-open-DAT father:ABS basket:ABS
‘I open the basket for father’
- (7) *á-dúŋ-íé enkálém*
1SG-cut-INST knife:ABS
‘I cut it with a knife’

Nandi:

- (8) *ii-nyòðr-u*
2SG-find-VEN
‘So that you (sg.) may find’
- (9) *kee-toor-tá*
INF-push-IT
‘To push’
- (10) *ky-áápír-cí ceeroono ceepeet*
PAST-1SG-hit-DAT cherono:ABS Chebet:ABS
‘I hit Cherono for Chebet.’

These morphosyntactic properties of verbs are common in Eastern and Southern Nilotic languages, but also in a series of Western Nilotic languages, although here such valence changes often involve vertical morphology and consonant alternation.

In other words, peripheral semantic roles such as location, direction or instrument are coded on the verb rather than on the adjunctival constituents themselves in these languages. It is impossible within the scope of the present contribution to show to what extent these various verbal extensions, e.g. for dative, ventive, itive, instrument as well as a number of other valence-changing verbal markers are the result of shared innovations in Nilotic or the result of historical drift. The fact that so many of these verbal suffixes are cognate between the various Nilotic languages suggests that we are dealing with a development going back to the earliest stages of Nilotic. We thus observe a historical drift or slant in Nilotic towards the loss of peripheral case marking as a dependent-marking strategy, accompanied by the growth of head-marking strategies on the verb in particular in Eastern and Southern Nilotic languages, but also in Western Nilotic. Head marking at the clausal level is particularly strong in Eastern and Southern Nilotic. Given the correlation established in Nichols (1986), namely that the head-marked clause pattern favors verb-initial word order, it seems logical to assume a correlation between the emergence of a more rigid verb-initial syntax in these Nilotic languages and extensive head marking on the one hand, and the virtual loss (with few lexical exceptions) of dependent-marking on the other.

Languages usually change their constituent order in situations where their speakers get into contact with speakers of typologically different languages. As noted above, Southern and Eastern Nilotic languages border on the Kuliak languages (whose speakers are widely assumed to have lived in the area before the intrusion of Nilotic groups from the north. Whether it was this contact that triggered the change in these Nilotic languages, or whether these languages were once part of a larger verb-initial area (with Hadza in Tanzania being another representative) remains an open question.

Nilotic languages may combine valence-changing strategies on the verb with the use adpositional phrases as productive strategies for the expression of peripheral semantic roles. As may be expected, these alternatives do not simply constitute “notational variants”. Adpositions may be used in order to further specify the search domain of some entity (in/up/next to etc.); the verbal “incorporation” strategy tends to be used for a number of other reasons, e.g. when such nominal complements are definite, or when the event expressed by the verb constitutes important information.

According to a widespread tendency, verb-initial languages tend to be prepositional, and verb-final languages tend to be postpositional. Verb-final Nilo-Saharan groups such as Saharan, Fur, Maban, Nubian, make extensive use of postpositions. Occasionally, this strategy is combined with the use of a small number of prepositions. This is also the system found in Nilotic’s closest relative, Surmic. Although none of the Surmic languages are verb-final, they use the same postpositional strategy in combination with a few prepositions. Thus, in Tannet there are postpositions specifying the search domain for some object (up, underneath, behind etc.) as well as a few prepositions used for expressing more abstract grammatical meanings, as shown in Randal (1998). Note, however, that regardless of whether a prepositional or postpositional strategy is used in these Nilotic languages, such adpositional phrases always follow the

Table 4. Adpositions in Nilotic

	Constituent Order	Postpositions	Prepositions
Eastern Nilotic:			
Turkana	VSO	<i>no</i>	<i>yes</i>
Maasai	VSO	<i>no</i>	<i>yes</i>
Southern Nilotic:			
Kalenjin	VSO	<i>yes</i>	<i>yes</i>
Western Nilotic:			
Dinka	V ₂	<i>yes</i>	<i>yes</i>
Päri	OVS	<i>yes</i>	<i>yes</i>
Luo	SVO	<i>no</i>	<i>yes</i>

verb in these languages. In Nilotic, the situation is somewhat more diverse, as shown in Table 4.

Given the presence of postpositions both in Southern and in Western Nilotic, as well as in the closely related Surmic languages, this strategy probably goes back to their common ancestor. The development of a prepositional strategy in verb-initial Eastern Nilotic languages could be the result of a drift or slant towards consistent head-modifier relations in these languages. In historical terms, the strict verb-initial structure of Eastern Nilotic languages appears to constitute an innovation. By abandoning postpositions in favor of prepositions in this Nilotic branch a more consistent head-initial pattern emerged in these languages.

The verb-initial Kuliak languages (Uganda) manifest a similar mixture of typological properties. In, for example, Ik the verb occurs in initial position in basic sentences, but in adpositional phrases, which follow the verb, the adposition may either precede or follow the complement noun (König 1999: 246):

- (11) *kínyomu* *ǵwarí* or *ǵwarí-ed-a* *kínyom-í*
 seed:OBL top:NOM top:POSS-NOM seed-GEN
 ‘The top of the seed’

Interestingly, Tosco (2000) has argued for a mirror-image situation with respect to head-final traits and the behavior of adpositions in modern Ethiopian Semitic languages. Whereas in these languages the verb has shifted to sentence-final position as a result of a gradual convergence towards neighboring Cushitic languages, some Ethiopian Semitic languages (e.g. Tigre) still have prepositions as in Arabic or other Semitic languages of the Middle East; other Ethiopian Semitic languages (e.g. Chaha) have postpositions, whereas still other languages, e.g. Amharic, combine the two strategies. As argued by (Tosco 2000: 358), these other head-final traits appear to be driven by the tendency to build a consistent (head-final) syntactic pattern “...very possibly in order to ensure parsing efficiency. But this tendency is, nevertheless, an internal development and, as such, is subject to the internal constraints of the language”.

Proto-Nilotic was not necessarily a verb-initial language. But in all likelihood it did allow for clauses with post-verbal subjects, the latter being inflected for case, as well as for verb-initial sentences (as in Western Nilotic Dinka, Southern Nilotic Nandi, and Eastern Nilotic Maasai). In Eastern and Southern Nilotic verb-initial structures were generalized as the basic sentence type, thereby allowing both subjects and objects to follow the verb; in Western Nilotic languages either the subject or the object may appear post verbally. Case marking for postverbal subjects functioned as a main coding device for syntactic relations. Viewed from a wider Nilo-Saharan perspective, the Nilotic branch has a highly reduced case system; at the same time, however, we may observe a strongly developed head marking system at the clausal level, most dramatically so in the Southern and Eastern Nilotic branch. There is, however, one cluster of closely related Eastern Nilotic languages, known as the Teso-Turkana group, where peripheral case marking re-appeared. The question of how and why this situation may have come about is central to the following section.

4. A Twilight Zone: The Nilotic-Surmic borderland

4.1 Areal adaptation at the Nilotic side of the border

As we saw above, a typical Nilotic language does distinguish between Nominative (or Ergative) case and Absolutive as core case distinctions, and lacks peripheral case marking; the latter can be shown to be due to loss historically. There is, however, one group of closely related Eastern Nilotic languages or dialects, known as the Teso-Turkana dialect cluster, where head marking (on the verb) is combined with a system of dependent marking. These languages, which are spoken in the border area between Ethiopia, Kenya, Sudan and Uganda, have a fully productive system of case marking for core constituents and peripheral constituents. They distinguish between Nominative case (for postverbal as against preverbal subjects) and Absolutive case (which is used for objects, pre-verbal subjects, and pronouns, nouns and noun phrases in isolation). This, of course, is the common system for Nilotic. However, unlike other Nilotic languages, Teso-Turkana languages also distinguish between core case and peripheral case, by way of gender-sensitive nominal *prefixes*. Gender marking on nouns is found as a derivational category in a number of Western Nilotic and Southern Nilotic languages. Gender marking as an inflectional category of nouns, however, is an Eastern Nilotic innovation, as argued by Heine and Vossen (1983). Following Vossen (1983), Eastern Nilotic languages are usually divided into a Bari group and a Non-Bari group. Gender marking in the Bari group is covert, i.e. nouns are either masculine or feminine, their inherent gender being manifest only indirectly, by way of gender-sensitive agreement marking on nominal modifiers. The so-called Non-Bari languages, consisting of Lotuxo-Maa and Teso-Turkana, have obligatory marking of gender on head nouns as well. Historically, these politics or prefixes (depending on the language) developed out of demonstratives, as argued by Heine and Vossen (1983). In the Lotuxo-Maa group

within the Non-Bari group the form of these gender markers is not affected by case. These languages distinguish between Absolutive and Nominative, following the more common Nilotic pattern, by way of tonal inflection. Whether a noun or noun phrase functions as a core constituent or as a peripheral constituent syntactically is irrelevant for the actual shape of the gender marker in Lotuxo-Maa. Compare the following examples from Maasai:

- (12) *e-iput-a e-moti n-ε-pík en-kíma*
3-fill-PA F:SG-pot:ABS SUBS-3-put F:SG-fire:ABS
'(S)he filled the pot and put it on the fire'

Note that the variation for the feminine gender prefixes between *e-* (in 'pot') and *en-* (in 'fire') is determined by morphophonemic, rather than, morphosyntactic, rules. Prepositions may be used in Maasai and other Eastern Nilotic languages in order to specify the search domain for some entity. From a historical-comparative point of view, the important point is that there is no oblique (peripheral) case marking in Maasai, or Lotuxo-Maa in general.

In Teso-Turkana, however, the gender markers are sensitive to case. The distinction between nominative and absolutive case is rendered by way of tonal inflection, as in Lotuxo-Maa. But the actual form of the gender prefix depends on whether the noun or noun phrase functions as a core constituent (i.e. a subject, a primary or secondary object), or peripheral constituent. With respect to the latter, a distinction occurs between locative and instrumental case. Compare Turkana, where the palatal nasal of the feminine gender prefix, found in more conservative northern dialects as well as in traditional songs, reflects the more archaic singular form still found in the closely related Toposa language. Note also that the Instrumental case in Teso-Turkana involves tonal inflection of the noun, whereas the gender prefixes are identical to those used for the core syntactic functions of subject and object (cf. Dimmendaal 1983:259–269 for a description).

	Object	Postverbal subject	Adjunct phrases		
	ABS.CASE	NOM.CASE	LOC.CASE	INSTR.CASE	
F:SG	(<i>n</i>)a-beru	(<i>n</i>)a-beru	<i>nà-beru</i>	(<i>n</i>)a-berù	'woman'
M:SG	(<i>n</i>)e-kile	(<i>n</i>)e-kile	<i>lò-kile</i>	(<i>n</i>)e-kilè	'man'
F:PL	<i>ηa-bèr(u)</i>	<i>ηa-ber(ù)</i>	<i>na-ber(ù)</i>	<i>ηa-ber(ù)</i>	'women'
M:PL	<i>ηi-kilyok</i>	<i>ηi-kilyòk</i>	<i>lo-kilyok</i>	<i>ηi-kilyòk</i>	'men'

In their comparative study of gender marking in Eastern Nilotic, Heine and Vossen (1983) have argued that the original gender markers for this Nilotic branch go back to demonstratives **na* (feminine singular) and **lo* (masculine singular); in the plural, a morpheme **ku-* preceded these gender markers. These functions have been retained as such in the Bari group within Eastern Nilotic (Heine & Vossen 1983:255–256). In the Non-Bari languages, the gender-sensitive demonstratives became cliticized onto the noun (following certain Greenbergian stages of definiteness marking). Heine and

Vossen (1983:262) further observe that the original forms **na* and **lo* are still attested as such in Teso-Turkana as demonstratives expressing proximity and as relative clause markers. But, we may now add, they were also retained as such when nouns (or noun phrases) functioned as peripheral syntactic constituents expressing location. The distinct forms of the gender prefixes required when nouns function as subject or object are due, according to the same authors (p. 262), to the fact that a pre-prefix **ni* was added. There is a widespread singular demonstrative with the same form (*ni*) in Nilotic; Rottland (1982:225), for example reconstructs a [+proximate] demonstrative *ni* for Proto-Kalenjin, with a corresponding plural form *cu*. Cognate morphemes occur in the other primary sub-branch of Southern Nilotic, Omotic-Datooga (cf. Rottland 1982:252, where a demonstrative marker **i* (singular), **cu* (plural) are given).

Heine and Vossen discuss developments in the gender system for singular nouns in Teso-Turkana, for which the following scenario is proposed (p. 262):

Proto-Eastern Nilotic	Proto-Teso-Turkana	Teso-Turkana	
		Toposa	Teso and Turkana
M:SG <i>*lo</i> >	<i>*ni-lo</i> >	<i>je-</i>	<i>ε-, e-</i>
F:SG <i>*na</i> >	<i>*ni-na</i> >	<i>ja</i>	<i>a-</i>

Additional, supporting evidence comes from the fact that in conservative Teso-Turkana varieties like Toposa, there is also an allomorph *jo-* for the masculine singular, used when the first root vowel is a back vowel (as in *jo-tuko* 'zebra'). There is no rounding harmony for prefixes in Toposa otherwise; the form *jo-* accordingly is best explained as a reflex of a former back vowel in the prefix **lo-*.

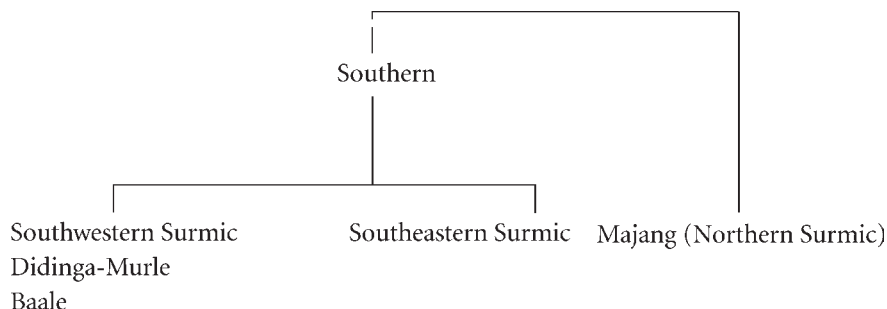
The reconstructed Proto-Teso-Turkana prefixes used with nouns functioning as core constituents are virtually identical to the [+proximate] demonstratives in the Bari group (*nyi-lɔ* (m.) and *nyi-na* (f.)). These forms may therefore have been added to nouns stems in Proto-Teso-Turkana in order to mark definiteness or specificity, once the erstwhile gender markers **na-* and **lo-* had lost this function. The corresponding [+proximate] demonstrative for the plural in the Bari group are *ku-lɔ* (m.) and *ku-ne* (f.). Whether they formed the etymological base for the plural gender markers *ji-* (m.) and *ja-* (f.) for nouns functioning as core constituents in Teso-Turkana remains to be seen.

We thus conclude that the Teso-Turkana languages developed a distinction between core case and peripheral case through a modification of the prefixal gender forms used for syntactic subjects and (primary and secondary) objects as against adjunctival forms; this modification resulted in syncretism, in that gender prefixes became sensitive, not only to gender and number, but also to case. The same gender prefixes were/are used in Teso-Turkana in order to introduce instrumental nouns; in addition, however, the tonal structure of such nouns is modified.

As we shall see next, this combination of head marking on the verb and dependent marking on peripheral syntactic constituents in Teso-Turkana is also used in neighboring Surmic languages.

4.2 Areal adaptation at the Surmic side of the border

Until recently, very little was known about the Surmic branch within Nilo-Saharan. This group of approximately twelve languages is “squeezed in” geographically between Omotic (Afroasiatic) languages to the north and east, and Nilotic languages mainly towards the south and west. Surmic consists of a relatively isolated Northern Surmic language Majang, and a Southern Surmic group consisting of Southwestern and Southeastern Surmic according to Unseth 1988; Dimmendaal 1998).



The first in-depth historical-comparative study of Surmic by Moges Yigezu (2001), using classical Neogrammarian methods, has confirmed the validity of this subgrouping. First sketches of the various members of the Surmic language family may be found in Dimmendaal and Last (1998). Here we shall concentrate on the Didinga-Murle languages, first, because a number of detailed typological studies exist for these Southwestern Surmic languages. Second, because the areal picture described for Nilotic above allows us to make some additional observations on developments in the Teso-Turkana group that would remain somewhat enigmatic otherwise.

Surmic languages use dependent-marking (case) systems (as is common in other Nilo-Saharan branches), as well as a certain degree of head marking, though usually not for peripheral semantic roles, with the exception of one Southwestern Surmic group, Didinga-Murle. These four languages combine core and peripheral case-marking strategies with a set of valence-increasing operations marked on the verb. Compare a typical language such as Tannet, where oblique roles marking location or instrument require a case suffix (data from Randal 1998):³

- (13) a. *k-a-tángû* *anná* *kween-a*
 1-PERF-sleep 1SG:NOM mat-OBL
 ‘I slept on the sleeping-mat’
 b. *k-a-tangu-óí* *anná* *kween*
 1-PERF-sleep-OR 1SG:NOM mat:ABS
 ‘I slept on the sleeping-mat’
- (14) a. *k-a-kat-a* *anna* *taang illa-w-a*
 1-PERF-spear-SG 1SG:NOM cow spear-EP-OBL
 ‘I speared the cow with a spear’

- b. *k-a-kat-oi* *anna* *illa*
 1-PERF-spear-OR 1SG:NOM spear:ABS
 'I speared with the spear'

As these examples illustrate, Tennet, is a verb-initial language (as are the other members of the Didinga-Murle cluster, which thus are similar to the neighboring Nilotic languages belonging to the Teso-Turkana cluster).

Whereas in general languages tend to use either head marking or dependent marking, if one of these techniques is in fact used as a morphological strategy, the Didinga-Murle languages use both. The case-marking system for peripheral roles reflects a wider Nilo-Saharan pattern, and thus probably constitutes a retention. As is common in Nilotic and elsewhere in Nilo-Saharan, there are a few prepositions (expressing direction and accompaniment) as well as postpositions (or postnominal modifiers) specifying location (compare, for example, Randal 1998: 223–224 on Tennet).

With respect to the verbal valence-markers the following may be observed: The dative in Didinga-Murle (involving a suffix *-k*) may be cognate with the dative in Nilotic (probably **-ki* in Proto-Nilotic). Other verbal valence-markers, however, e.g. affixes for instrumental, ventive, or itive) do not seem to be cognate with the corresponding forms in Nilotic. This suggests that they are the result of separate and independent historical developments. But the emergence of these valence markers exactly in those Surmic languages, which border on the Nilotic Teso-Turkana languages presumably is not a coincidence. First, there is ample lexical evidence for lexical borrowing between Didinga-Murle and Teso-Turkana (Dimmendaal 1982). Second, the Didinga-Murle languages are verb-initial, as are the neighboring Teso-Turkana languages. Note that Baale, which is closely related to Didinga-Murle, as well as other Southern Surmic languages, are not verb-initial. Third, the Teso-Turkana and Didinga-Murle languages use similar idiomatic expressions. Compare:

- (15) Turkana (Nilotic)
k-à-jam-it' *ayɔŋ'* *akòrò*
 PA-me-eat-AS 1SG:ABS hunger:NOM
 'I am hungry (lit. hunger is eating me)'
- (16) Tennet (Surmic)
á-dáh-ha *anét* *mágíz*
 IMPF-eat-me 1SG:ABS hunger:NOM
 'I am hungry (lit. hunger is eating me)'

Moreover, in both language groups attitude markers are used expressing the speaker's evaluation of a propositional act; in both groups this interactional concept is expressed by particles following the verb.

To these various morphosyntactic properties shared by Teso-Turkana and Didinga-Murle we may now add another property: Whereas the Teso-Turkana languages can be shown to have developed case marking for peripheral constituents (by way of pre-fixes), the Didinga-Murle languages (which already had such a system of case suffixes)

must have innovated an extensive system of verbal valence-marking, parallel to the Teso-Turkana languages. This morphosyntactic convergence between the two groups resulted in a typologically somewhat odd situation, whereby these two neighboring language groups use both head marking and dependent marking at the clausal level; cross-linguistically, it is more common for languages to exploit one of the two coding strategies.

Still, there is also a difference between the two groups, the importance of which needs to be evaluated. Whereas the Teso-Turkana group uses prepositions, the neighboring Didinga-Murle languages use postpositions. Note, however, that postpositional phrases in Surmic languages follow the verb (plus subject and object); prepositional phrases in Teso-Turkana and other Nilotic languages groups also follow the verb (plus subject and object). Compare Tennenet, which uses postpositions expressing location, in combination with a few prepositions expressing direction or accompaniment (data from Randal 1998):

- (17) a. úk Loú dó cééz-a écitó
 PERF:go Loudo:NOM house-OBL inside
 'Loudo went into the house'
- b. úk Loú dó *rok* Júba
 PERF:go Loudo:NOM PREP Juba
 'Loudo went up to Juba'

Here, then, is an area where Didinga-Murle and Teso-Turkana apparently did not converge. As noted for Ethiopian Semitic above, this typological "inconsistency" is also attested elsewhere. It would therefore be a useful research question to ask oneself why languages allow for such patterns. First, it may be asked how frequent adpositions are in normal discourse. Second, prepositional and postpositional phrases in Nilotic and Surmic follow the verb phrase; their prepositional or postpositional nature in fact only becomes clear once the adpositions are used in combination with a complement, and so the question arises how frequent bare adpositions as against adpositional phrases are. The research question raised above can only be answered, it would seem, once extensive texts are available for the various languages involved.

5. Goddard's Law

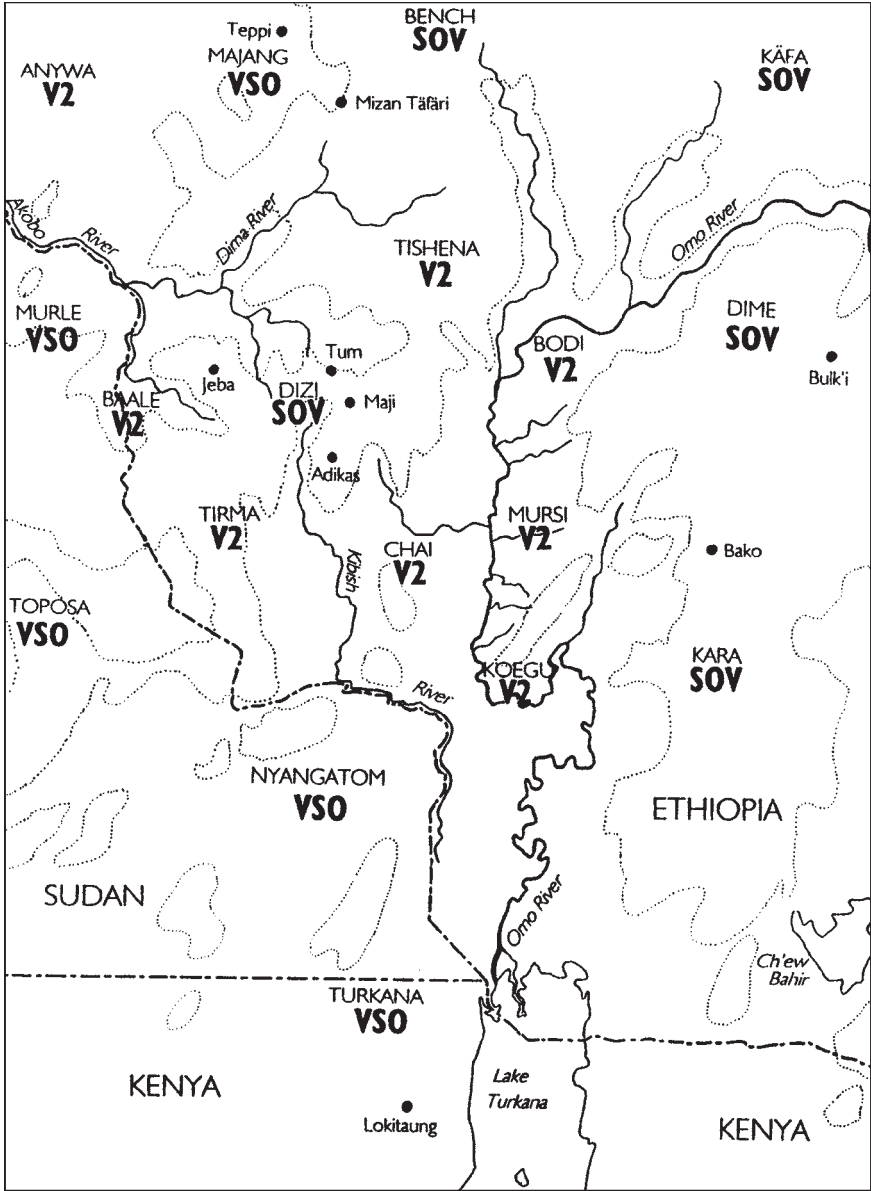
Initial evidence for areal contacts between the Surmic Didinga-Murle group and the Nilotic Teso-Turkana group was based on the identification of mutual lexical borrowing (Dimmendaal 1982). It has become clear ever since that the mutual influence between these language groups has been much more profound, affecting not only constituent order, but also the coding of syntactic relations. The Didinga-Murle languages are closely related to Baale, with which they form the Southwestern branch of Surmic. But, whereas the Didinga-Murle languages show many structural similarities with the neighboring Teso-Turkana group, Baale has converged towards the distantly

related Southeastern Surmic languages Tirma and Chai (Moges Yigezu & Dimmendaal 1998; Dimmendaal 2001a). Such linguistic restructuring or metatypy usually is the outcome of profound cultural contact and adaptation with (compound) bilingualism being the norm. The Baale form an ethnic unit with the Tirma and Chai, known as Suri or Surma. Speakers of languages such as Teso, Toposa, Nyangatom or Turkana, or Surmic groups such as the Didinga, Murle, Narim, or Tennenet each form their own distinct ethnic groups. In spite of the divergent situation in terms of ethnicity, the Baale identifying themselves with neighboring groups such as the Chai and Tirma, the Didinga-Murle groups each keeping their own distinct ethnicity, the languages of all of these Southwestern Surmic groups converged towards neighboring languages. Whereas the Baale appear to be on friendly terms with the Tirma and Chai, the Didinga-Murle groups often were not on friendly terms with Teso-Turkana groups; in fact, there have been numerous reports of clashes also in more recent times. Still, this latter situation did not impede the development of multilingualism and compound bilingualism triggering the transference of grammatical features between the Didinga-Murle and Teso-Turkana languages.

At the same time, one may observe sociolinguistic situations where languages spoken in adjacent areas do not converge. Majang, a verb-initial Surmic language surrounded mainly by verb-final Omotic languages (see Map), shows virtually no signs of convergence towards the latter. The social setting of the Majang has been described by Stauder (1971). From this source and from the present author's own observations, the Majang appear as a relatively close-knit group with few relationship links to speakers in other groups; intermarriage with neighboring groups, for example, appears to be rather restricted. Consequently, there appear to be or have been relatively few long-term multilingual settings in the Majang community that could have lead to structural interference from typologically distinct languages.

Southern Nilotic groups have been in close contact with speakers of Bantu languages for many generations. This resulted in convergence, in various parts of their grammars, e.g. the development of tense marking on verbs (see Dimmendaal 2001b for a description). But again, when traveling further south towards northern Tanzania one comes across languages like Hadza, either an isolated language or a member of the Khoisan phylum, which must have been surrounded by Bantu (or "Swahili people", as the Hadza call them), and possibly other families such as Cushitic and Southern Nilotic such as the Datooga for a considerable period of time. Here, however, there appears to be no evidence for structural convergence between Hadza and these latter languages. And here is where the final part of "Goddard's law" (quoted in Watkins 2001:60) appears to be coming in: "A language can do whatever it wants to..., if it wants to".

Speech communities may be close-knit internally but they may still be open towards others, as was the case historically presumably for the Teso-Turkana and Didinga-Murle groups, whose languages were discussed above. This sociolinguistic situation would create a situation where convergence (or metatypy in the sense of Ross 2001) may take place. If, on the other hand, the external boundaries are also relatively



Areal spreading of word order type.

close-knit, either out of free choice or because one constitutes a somewhat stigmatized social group, compound bilingualism and convergence is less likely to develop. This, it seems, is what happened in the case of Hadza or the Surmic language Majang.

6. A research program: Understanding the “Bauplan” of a language

Some two decades ago, when the present author presented a description of Turkana (Dimmendaal 1983), this Eastern Nilotic language appeared to manifest a typologically odd mixture of syntactic strategies. Whereas better known verb-initial Nilotic languages such as Maasai (first described in detail by Tucker & Mpaayei 1955), showed a clear propensity for head marking at the clausal level, the closely related Turkana language appeared to use this strategy in tandem with a dependent-marking strategy. Much more is known today about Turkana’s neighbors, the Surmic languages. From an areal point of view, Turkana and other closely related Eastern Nilotic languages belonging to the Teso-Turkana cluster form an almost perfect transition with respect to the expression of cohesion at the syntactic level in terms of the head-marking/dependent-marking parameter, and so today the Turkana system makes sense from an areal point of view. This areal knowledge, combined with knowledge about the historical development of the genetic group of which it forms part, Nilotic, as well as language typology as a heuristic and controlling device, allows the interested scholar to understand the system as it manifests itself in all its diversity and complexity. This approach, whereby intra-genetic and inter-genetic comparisons are combined, of course is known as “the method of dynamic comparison” (a method probably first proposed by Greenberg 1978). Greenberg did not include areal linguistics as a dimension, but by including the latter, we may arrive at an even better understanding of the Bauplan of a language.

In the case of Nilotic, historical comparison with genetically related groups suggests a decrease in peripheral case marking and an increase in head marking at the clause level. Language typology teaches us that head marking is common in verb-initial languages. Similarly, language typology teaches us that such languages tend to use prepositions, rather than postpositions. The observed changes from postpositions to prepositions in Eastern Nilotic accordingly may be interpreted as being the result of self-organizing dynamics of the system, resulting in a consistent head-modifier order in languages involved. Areal contacts are known to play an important role in language change on the African continent. What is needed next, in order to arrive at a deeper understanding of these languages in the area from a general linguistic and historical-comparative as well as areal point of view, are in-depth studies of individual languages. With respect to the languages and their alternative clausal strategies discussed above, it is crucial to know what the consequences are in terms of information packaging, when languages use verbal valence-markers rather than case-marked adjuncts? Moreover, what consequences do historical changes in constituent order have for the discourse structure of these languages? Or, what is the role played by adpositions in terms of information packaging? It is only through such dynamicization of sub-typologies that

one may arrive at an ultimate understanding of the Bauplan of individual languages in the area as well as of important differences in their organizational structure.

Abbreviations

1	First person	INST	Instrumental
2	Second person	IT	Itive
3	Third person	M	Masculine
A	Affirmative	NOM	Nominative
ABS	Absolutive	OBL	Oblique
AS	Aspect marker	OR	Oblique raising
DAT	Dative	PA	Past tense
DCL	Declarative	PERF	Perfective
EP	Epenthetic consonant	PL	Plural
F	Feminine	POSS	Possessive
GEN	Genitive	SG	Singular
IMPF	Imperfective	SUBS	Subsecutive
INF	Infinitive	VEN	Ventive

Notes

- * I would like to thank members of the audience at the International Symposium for their critical comments and suggestions. Thanks are also due to Monika Feinen for the map included in the present contribution.
1. As shown in the detailed description by Azeb Amha (2001), subjects and objects can in fact follow the verb in Maale, given specific discourse contexts. For a number of distributional reasons, however, Maale is best analyzed as a verb-final language at the structural level (cf. Azeb Amha 2001:235–250).
 2. It is not entirely clear whether the post-verbal subject marking system in Dinka is cognate with that found in languages such as Pāri and Anywa, or with Eastern and Southern Nilotic Nominative case marking. The reason for this uncertainty is the fact that such case-inflected post-verbal subjects in Dinka are formally identical to the Genitive (Andersen 1991); consequently, the former (Nominative case) in Dinka could have developed out of a genitive phrase marking the agent in a sentence. This formal identity between the Nominative and the Genitive case form is not attested in other Nilotic languages with postverbal case marking; consequently, postverbal case marking for subjects in Dinka could be the result of an innovation, after the original Nominative had been lost historically; its reintroduction (through an extended use of the Genitive case form) may have been triggered through areal contact with neighboring Surmic languages (which also mark postverbal subjects).
 3. Randal (1998) has argued that in a number of cases not only post-verbal but also pre-verbal subjects in Tennen appear to take Nominative case. It should be noted, however, that such pre-verbal subjects in Tennen are always preceded by certain particles, as Randal calls them; most

probably, these particles are auxiliary verbs. If correct, Tennet would confirm to a common pattern in Surmic and Nilotic languages in which only subjects following verbs (whether auxiliaries or main verbs) are inflected for case.

References

- Andersen, T. (1988). Ergativity in Pāri, a Nilotic OVS language. *Lingua*, 75, 289–324.
- — —. (1991). Subject and topic in Dinka. *Studies in Language*, 15, 265–294.
- Azeb Amha (2001). *The Maale Language*. Leiden: Center for African, Asian and Amerindian Studies.
- Blake, B. (1994). *Case*. Cambridge: Cambridge University Press.
- Crazzolaro, J. P. (1933). *Outlines of a Nuer Grammar*. Vienna: Anthropos.
- Creider, C. A., & J. T. Creider (1989). *A Grammar of Nandi*. Hamburg: Helmut Buske.
- Dimmendaal, G. J. (1982). Contacts between Eastern Nilotic and Surma groups: Linguistic evidence. In J. Mack & P. Robertshaw (Eds.), *Culture History in the Southern Sudan* (pp. 101–10). Nairobi: B.I.E.A.
- — —. (1983). *The Turkana Language*. Dordrecht: Foris.
- — —. (1988). The lexical reconstruction of Proto-Nilotic: A first reconnaissance. *Afrikanistische Arbeitspapiere*, 16, 5–67.
- — —. (1998). A syntactic typology of Surmic from an areal and historical-comparative point of view. In G. J. Dimmendaal & M. Last (Eds.), *Surmic Languages and Cultures* (pp. 35–81). Cologne: Rüdiger Köppe.
- — —. (2001a). Areal diffusion versus genetic inheritance: An African perspective. In A. Aikhenvald & R. M. W. Dixon (Eds.), *Language Diffusion and Genetic Linguistics* (pp. 358–392). Oxford: Oxford University Press.
- — —. (2001b). Language shift and morphological convergence in the Nilotic area. *Sprache und Geschichte in Afrika*, 16/17, 83–124.
- Dixon, R. M. W. (1997). *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.
- Ehret, C. (1983). Nilotic and the limits of Eastern Sudanic: Classificatory and historical conclusions. In R. Vossen & M. Bechhaus-Gerst (Eds.), *Nilotic Studies* (pp. 375–421). (Two Volumes). Berlin: Dietrich Reimer.
- — —. (2001). *A Historical-Comparative Reconstruction of Nilo-Saharan*. Cologne: Rüdiger Köppe.
- Greenberg, J. H. (1978). Some methods of dynamic comparison. In J. Puhvel (Ed.), *Substance and Structure of Language* (pp. 147–203). Berkeley, CA: University of California Press.
- Hayward, R. J. (Ed.). (1990). *Omoti Language Studies*. London: School of Oriental and African Studies.
- Heine, B. (1976). *A Typology of African Languages Based on the Order of Meaningful Units*. Berlin: Dietrich Reimer.
- — —. (in collaboration with Rainer Vossen). (1983). On the origin of gender in Eastern Nilotic. In R. Vossen & M. Bechhaus-Gerst (Eds.), *Nilotic Studies* (pp. 245–268). (Two Volumes). Berlin: Dietrich Reimer.
- König, C. (1999). Kasus im Ik: Wortarten, Kontext und Diskurs. Habilitationsschrift, Johann-Wolfgang Goethe University, Frankfurt/Main.

- Last, M., & Lucassen, D. (1998). A grammatical sketch of Chai, a Southeastern Surmic language. In Dimmendaal & Last (Eds.), *Surmic Languages and Cultures* (pp. 375–436). Cologne: Rüdiger Köppe.
- Leslau, W. (1945). The influence of Cushitic on the Semitic languages of Ethiopia: A problem of substratum. *Word*, 1, 59–82.
- Masica, C. P. (1976). *Defining a Linguistic Area – South Asia*. Chicago: The University of Chicago Press.
- Miller, C. L., & Gilley, L. G. (2001). Evidence for ergativity in Shilluk. *Journal of African Languages and Linguistics*, 22, 33–68.
- Moges, Y. (2001). *A Comparative Study of the Phonetics and Phonology of Surmic Languages*. Ph.D. dissertation, Université Libre de Bruxelles.
- Nichols, J. (1986). Head-marking and dependent-marking grammar. *Language*, 62, 56–119.
- Randal, S. (1998). A grammatical sketch of Tannet. In Dimmendaal & Last (Eds.), *Surmic Languages and Cultures* (pp. 219–272). Cologne: Rüdiger Köppe.
- Reh, M. (1996). *Anywa Language: Description and Internal Reconstructions*. Cologne: Rüdiger Köppe.
- Ross, M. (2001). Contact-induced change in Oceanic languages in northwest Melanesia. In A. Aikhenvald & R. M. W. Dixon (Eds.), *Language Diffusion and Genetic Linguistics* (pp. 134–166). Oxford: Oxford University Press.
- Rottland, F. (1982). *Die südnilotischen Sprachen. Beschreibung, Vergleichung und Rekonstruktion*. Berlin: Dietrich Reimer.
- Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Company.
- Stauder, J. (1971). *The Majangir: Ecology and Society of a Southwest Ethiopian People*. Cambridge: Cambridge University Press.
- Tosco, M. (2000). “Is there an “Ethiopian language area”?” *Anthropological Linguistics*, 42 (3), 329–365.
- Toweett, T. (1979). *Kalenjin Linguistics*. Nairobi: Kenya Literature Bureau.
- Tucker, A. N. & J. Tompo Ole Mpaanyi (1955). *A Maasai Grammar with Vocabulary*. London: Longmans, Green and Co.
- Unseth, P. (1988). The validity and unity of the ‘Southeast Surma’ language grouping. *Northeast African Studies*, 10, 151–163.
- Vossen, R. (1982). *The Eastern Nilotes: Linguistic and Historical Reconstructions*. Berlin: Dietrich Reimer.
- Watkins, C. (2001). An Indo-European linguistic area and its characteristics: ancient Anatolia: areal diffusion as a challenge to the comparative method. In A. Aikhenvald & R. M. W. Dixon (Eds.), *Language Diffusion and Genetic Linguistics* (pp. 44–63). Oxford: Oxford University Press.

Agent phrases in Bantu passives

Axel Fleisch

Universität zu Köln/University of California, Berkeley

Passive constructions have been discussed extensively and from various angles in the literature on Bantu languages. In the present paper, I wish to deal with the ways in which the semantic role of an agent may be encoded in passive sentences. It is relatively easy to gain information on the formal marking of a passive construction in a Bantu language from grammatical descriptions. Often passives are formed by means of a productive verbal extension indicating the passive use of a transitive verb. Since verbal extensions are a typical feature of Bantu languages, most descriptive grammars treat them in relatively detailed ways. Consequently, this also covers the passive.

The use of a passive construction usually implies that the subject of a corresponding active transitive sentence is not encoded as the subject of the derived sentence. For the sake of comparability, I will depart from prototypical transitive verbs, because these are the most likely to be able to undergo passivization. Therefore, although restrictions of applicability concerning passive morphology may exist, they are not likely to affect this particular group of verbs. The typical subject of these verbs is an agent exerting some action on something or someone else. It follows that in a passive sentence, the agent is not encoded as the syntactic subject. It may, however, be encoded in other ways depending on the particular languages. In this paper, three basic claims are made about the nature of the agent encoding in passive sentences in Bantu languages: (1) Agent encodings are among the most variable features concerning the passive voice in Bantu; (2) agent encoding is not an independent grammatical category; (3) although agent encodings are linked to the phenomenon of the passive voice, they are not an integral part of this grammatical domain.

The ways in which agents are construed in passive sentences play a minor role in the literature on the functions of passive morphology in Bantu verb phrases (cf. Kimenyi 1981; Givón 1976; 1986; Givón & Kawasha 2001). Yet, if one looks into grammatical descriptions of Bantu languages, they are often treated in connection with the formation of passive sentences. It is important to note that I do not intend to argue that agent encoding and passivization are two entirely independent grammatical phenomena. On the contrary, they are functionally related. Although this appears to be in opposition to the claims made above, I will try to show that it is not. This paper is

therefore not to be misunderstood as advocating the viewpoint that agent encodings are not to be dealt with in connection with passives. I simply wish to bring to the fore in exactly which ways both domains are linked with each other.

In order to analyze the nature of agent phrases with regard to passive constructions, I will proceed along the following lines. At first, some general information about the construction and meaning of passives in Bantu languages will be given. This will be followed by an overview of the different agent encoding strategies in Bantu languages. This information is needed in order to find convincing arguments for the claims made above concerning the relationship between passive constructions in the verbal domain and the encoding of nominal arguments in general and the agent noun in particular. Several arguments will be outlined in the following section that show the relative independence of agent encoding mechanisms and passive constructions. In a brief conclusion, the major points are summarized and implications that may reach beyond the realm of the Bantu languages will also be outlined.

1. Passive constructions in Bantu languages

Following the rather general characteristics of passives mentioned in the introduction, it is necessary to look into typical passive constructions in Bantu in order to delimit the subject of this paper. Several important points will not be raised in this paper, because they do not immediately bear on the issue of how to express an agent in a passive sentence. Therefore, it can be summarized rather briefly that the main concern here is with prototypical transitive verbs with an agent noun encoded as the subject of an active sentence, whereas the patient, which is affected by the action, is represented in the form of the direct object. Bantu languages differ considerably as to the extent to which they allow the use of passives. Thus, in many languages less typical passives occur, such as those with verbs which do not assign an agent role, passives on the basis of intransitive verbs, etc. Therefore, it should be borne in mind that passives in Bantu languages, notwithstanding their formal similarity, may differ as to their functional range from one language to another. Given these differences across language boundaries, it is important to make sure that one talks about the same phenomenon when talking about a passive construction. As a matter of fact, there seems to be little doubt that the underlying phenomenon is the same no matter how much it may diverge in applicability from one language to another. This is partially because of formal resemblance across Bantu languages. The most widespread passive extension is *-(ib)w-/-(ig)w-* which is used to form passive verb stems. Apart from formal aspects, however, there is also a functional common denominator concerning the use of passives: whenever there is a formal passive, it is always applicable to prototype actions with an agent subject and a patient object in the active sentence (Keenan 1985:247).

It is important to note, however, that passives in Bantu do not necessarily follow this neat picture. Most Bantu languages have other extensions some of which may show a certain functional overlap with the passive. Among the languages dealt with in this

article attention should be drawn to Ndonga, because in this language the extension *-ik-*, which usually expresses neuter-passive meaning in most languages, has been attested in typical passive contexts (cf. Fivaz 1986). With regard to the extension *-ik-*, one often finds – instead of ‘neuter-passive’ – designations such as ‘potential’ or ‘stative’ which are used due to additional functions taken over by the respective construction in a given language. There is an important difference between a passive and the other grammatical functions mentioned here. Whereas in a typical passive the agent noun is removed from its prominent subject position, but is still understood to be at work, in the case of neuter-passives, the agent is not only syntactically removed, but also deleted from the semantic arrangement of roles assigned by the verb. No agent is involved in situations expressed by neuter-passive constructions. Therefore, the latter constructions are only relevant here to the extent that they must be clearly distinguished from the “true” passives which will be dealt with here, because only they possibly allow an agent phrase to be expressed explicitly.

Other Bantu languages make use of a different morphological means of expression altogether. This is the case for many of the SW Bantu languages, which lack an overt passive marker in the form of a verbal extension. Instead they make use of a verb form marked for a (non-referential) third person plural. One might of course argue that these constructions are not passives at all. Lacking any morphological device on the verb phrase, from a formal perspective they could simply be regarded as active sentences with a non-referential pronominal marker, which serves to overcome the lack of a proper passive. The issue is not so simple, however, and there are good reasons to assume that such constructions should indeed be regarded as passives in many languages (cf. Givón 1976). One of the reasons is the fact that these sentences in addition to the third person plural marking require a specific syntactic structure. Interestingly, some of these structures do occur with agent phrases, which indicate desemanticization of the subject marking for third person plural.

(1) Kimbundu [H.21] (Givón 1981:182)

Nzua a mu- mono kwa meme
 John SUBJ.3PL- OBJ.3SG see.PAST by PRON.1SG
 ‘John was seen by me’

A question which ensues, and which will be taken up in the discussion of the relationship between passive constructions and agent encoding strategies, is the following: may it not be the respective path of grammaticalization which may or may not leave room for an agent construction? Given that passives themselves often derive from different backgrounds, this would be a plausible explanation – at least if there was a systematic (inter-)dependence of agent marking strategies and passive formation. Haspelmath (1990) mentions several different paths of grammaticalization that may lead to passives. These are basically the use of generalized subject constructions, the extension of reflexives to cover passive meaning, similarly the extension of causatives, and the use of inactive auxiliaries (ibid. 37–46). The Bantu suffixes often have a verbal origin. Two plausible candidates that may have come to be used as auxiliaries in passive

constructions are **-gual/*-bua* ‘fall’ and **-(i)kala* ‘stay, sit, remain’ leading to *-(ib)w-* (passive) viz. *-ik-* (neuter-passive). As mentioned above, another strategy found in the Bantu languages pertains to the field of generalized subject constructions. In many SW Bantu languages the semantically empty third person plural marking should rather be considered a passive marker.

2. Construal of agent phrases in passive sentences

A few general remarks have to be made about agent phrases in passive sentences. From typological evidence it appears that the occurrence of agent phrases in typical passive sentences is not a necessary component of passives. There are many languages with a morphological passive, which do not even allow the agent of a passive verb to occur in the passive sentence. Although the agent is understood to be present in the grid of semantic roles as assigned by the verb, no overt agent noun may be used in passive sentences in these languages. Often also in those languages which show a morphological way of encoding agents as oblique arguments, their occurrences in passive sentences are not frequent. Thus, one can find a cline from languages which have a passive construction but do not allow agents to co-occur to those languages which allow agents to be mentioned in passive sentences rather freely with different degrees of agent phrase acceptance in between.

This raises another interesting question: is there anything clearly discernible that would account for possible restrictions concerning the occurrence of agents in passive sentences? Recalling what was stated before, one might think of the passive type, e.g. its diachronic background implying a certain conceptual make-up. This is of course a line of investigation that will be taken up in the following sections. There is, however, another point here. Agent encoding strategies are far from being uniform. There are several formal ways in which agents can be encoded. These need to be presented in some detail in order to gain more insight into the question how agent encodings and passive constructions interact. Interestingly, whereas the passive constructions in Bantu seem relatively homogeneous from a formal point of view, this is not at all the case for agent marking devices.

If one takes a closer look at the mechanisms which may serve to express an agent in a passive sentence, it will be noted that many Bantu languages indeed do not allow making explicit mention of an agent. This seems to be the case for Cokwe. In this language, an agent phrase cannot be present if the verb is derived with the passive extension *-w-*. If the agent is to be mentioned, an active sentence must be used. The object may be topicalized—in which case it stands sentence-initially—but the syntactic functions of the nominal arguments are left untouched as compared to a simple active sentence (dos Santos 1962:175–176). Other languages, as stated above, show special agent encoding devices. Notwithstanding, explicit mention of the agent tends to be avoided in a sentence with a passive predicate. In some cases, it is difficult to assess whether a language allows an overt agent in a passive sentence on the basis of descrip-

tive grammars, because if no explicit mention of this topic is made, the reason may be that agent phrases do not occur in passives. However, it may of course also be the case that the phenomenon has simply been omitted by the respective author. For the same reason, it is in some cases not possible to identify how agents are marked, even if their use may be totally acceptable among the speakers of a particular language. In the case that an overt agent phrase may occur freely in a passive construction, it is usually, however, easier to receive information about the way in which the agent phrase is marked.

What strikes me as important is the fact that no matter what restrictions may exist in the domain of the passive voice in Bantu, this does not bear consequences for the formal marking of agent phrases where these are allowed to occur. There is no evidence that any formal way of expressing an agent in a passive sentence tends to be more or less subject to restrictions than the others. Before discussing such issues in more detail, the different ways of marking agents as oblique arguments are presented in the following.

2.1 Comitative marking on agent phrases

First of all, comitative marking is often used to express agents. This strategy may be paraphrased in English as ‘sth. has been done WITH agent’. This is of course rather well-known and is briefly illustrated by the examples from Swahili and Shona, although this strategy is much more widespread. The morphological marker *na-* introduces the agent phrase; in the case of Shona the vowel is subject to regressive assimilation depending on the class prefix of the agent noun. So what is translated in English as ‘by’ should more literally rather say ‘with’:

(2) Swahili [G.42]

a- me- pig -w -a na wa-tu
 3S PERF beat PASS FV COM 2-people
 ‘He has been beaten by the people’

The original comitative meaning is illustrated in the following example:

(3) Swahili [G.42]

a- me- rudi na rafiki yake
 3S PERF return COM 9.friend 9.his
 ‘He has returned with a friend’

It is particularly difficult to find examples where the comitative is used as the clearly discernible source for the agent marking. It is striking that in many cases, there is only one polysemous marker that covers both comitative and instrumental meaning. An example for such a language in which both domains intersect is Shona:

(4) Shona [S.12] (O’Neil 1935:55)

nyoka y.aka- uray- iw- a no mu.komana u.pi
 snake 9.PAST kill PASS FV COM 1.boy 1.which
 ‘The snake was killed by which boy?’

The following examples illustrate that the same marker serves to express comitative and instrumental meanings. The formal difference (*nó-/ná-/né-*) is due to regressive morphophonological assimilation depending on the class prefix of the following agent noun:

- (5) Shona [S.12] (Fortune 1939:399)
- a. *enda na baba*
go-IMP COM father
'Go with father!'
 - b. *nd- a- roh- w- a ne-shamu* [*<na+i.shamu*]
1SG PERF beat PASS FV INSTR-5.cane
'I was beaten with a cane'

It must be noted that in example (5b) the last noun phrase is not an agent phrase, but refers to the instrument with which the action is done. An agent is not made explicit in this sentence, although it is understood to exist. What is interesting in this regard is the fact that the appearance of the instrument noun seems to hinder the explicit mentioning of an agent. This may be regarded as further evidence for the fact that agent marking, comitative and instrumental are facets of only one grammatical category in Shona, if one assumes that the occurrence of several adverbial phrases covering a similar range of meaning (COM ~ INSTR) is avoided for pragmatic reasons.

2.2 Locative

A second strategy is the one based on a locative notion. The agent is expressed as the place at which the passive situation takes place: 'sth. is done AT agent'. Therefore, the agent phrases which are marked in this way show one of the noun class prefixes indicating the locative classes, usually either class 16 or 17, cf. the following example from Tonga [M.64]

- (6) Tonga [M.64] (Collins 1975: 54)
- u.aka- jayig- w- a a.Joni*
3SG.REM.PAST kill PASS FV 16.John
'He was killed by John'

The same prefix *a-* is the marker for the locative noun class 16 as shown in the example *kolanga ansi* [*<a_{LOC.16}- (i)nsi*] 'look on the ground' (Collins 1975:55).

Probably, the locative concept as a source of agent marking in passive sentences should rather be treated as a group of related strategies. There are other languages which also base the agent marking on a locative concept. They differ, however, from the example in (6) in that they show additional morphology, not the simple class prefix. This is clearly the case in Ila. Whereas in Tonga the prefix of noun class 16 is added directly to the agent noun, in Ila the agent marker is composed of the noun class prefix – in this case class 17 – and the genitive marker. Thus the agent phrase would literally

translate as ‘at of God’; the closest English equivalent would be something like ‘at God’s [place]’.

- (7) Ila [M.63] (Smith 1907:138)

i.nshi y.aka- bumb- w- a ku- a.Leza ku- ku.kank- a
 9.earth 9.PAST create PASS FV LOC GEN.God LOC 15.begin FV
 ‘The earth was created by God in the beginning’

The third locative example which is from Luba combines the locative class prefix with a copula, i.e. with the defective verb *-di*. The best way to convey the underlying concept is by paraphrasing these constructions as ‘where there is (the owner)’.

- (8) Luba [L.31] (Burssens 1939:162)

mu.ivi u.di u.kwit- i:bu- a ku.de mu.ana mbui
 1.thief 3SG.PROG 3SG.catch PASS FV 17.be 1.owner 9.goat
 ‘The thief is caught by (< where there is) the owner of the goat’

There may be conceptual differences between divergent locative strategies. In English, the preposition ‘by’ which is used to introduce an agent phrase rests on a rather static spatial association which implies a concept of control requiring physical closeness. On the contrary, in German a directional notion is involved in the use of *von* ‘from’. The underlying notion which allows directionality to be coupled with agentivity is based on a concept of an agent as a spatial source from which the action is induced (Heine et al. 1993: 8–9). Concerning the Bantu languages, it is problematic to attempt discerning between both notions, since both locative class 16 and 17 are used to mark agents. They are relatively vague as compared to the prepositions in the above-mentioned Germanic languages. It is noteworthy, however, that no Bantu language considered for this study uses class 18, which usually conveys the meaning of being inside a place.

2.3 Instrumental

A third strategy which may be paraphrased as ‘sth. has been done BY MEANS OF agent’ is illustrated by the example from Venda. This language uses the same device for agent marking which is also used to express instruments with which something is done. So, Venda *nga-* can be translated more literally as ‘by means of’.

Since I wanted to show an unequivocal link between instrument and agent marking, I chose examples from Venda, because in this language, the respective marker goes back to neither a locative nor a comitative marker: the element *ngá-* does not show links to the comitative *ná-*, nor to any of the locative noun classes (*fha-*, *ku-*, *mu-*). Example (9b) illustrates that this marker denotes basically an instrumental meaning.

- (9) Venda [S.21] (Poulos 1999:173–174)

a. *ma.fhi a.do- nw- iw- a nga- tshi.mange*
 6.milk 6.FUT drink PASS FV INSTR 7.cat
 ‘The milk will be drunk by the cat’

- b. *vha.ya mu.shumoni nga-modoro*
 2.go LOC.18.work INSTR-3.car
 ‘They travel to work by car’

In the preceding section on the comitative strategy it has been mentioned that morphological markers often cover various meanings. It should therefore be remembered that in many languages instrumental and comitative meaning are expressed by the same marker. Such marking devices which are not clearly assignable to one of the strategies are dealt with in more detail in Section 2.6 below.

2.4 Copulative

Examples (10) to (12) illustrate another strategy that serves to express the agent in a passive sentence. In Sotho-Tswana and Nguni, the copula is used to introduce the agent. Thus, the agent phrases can be paraphrased as ‘sth. is done, IT is agent’. In languages from the Sotho-Tswana group the copula is always *ke*, as illustrated in the following example from Northern Sotho.

- (10) Northern Sotho [S.32] (Louwrens et al. 1995:60)

- a. *nôga ê- bôn- w- a ke- mo.sadi*
 9.snake 9- see PASS FV COP 1.woman
 ‘The snake is being seen by a/the woman’
 b. *Maepa ke- mo.rutiši*
 Maepa COP 1.teacher
 ‘Maepa is a teacher.’

It must be noted that the agent marking is invariable, although the formal shape of the copula depends on person and number. The copula that served as the formal source for the agent marker is the one from the third person singular as illustrated in (10b).

In languages from the Nguni cluster the copula is variable as well. In contrast to Northern Sotho this is also reflected in its use as an agent marking device: in the examples from Xhosa and Zulu we find both an element *ngu-/nga-* and *y-* depending on the noun class of the following agent noun.

- (11) Xhosa [S.41] (McLaren 1955:100–102)

- in.qwelo y.a- tsal- w- a nga- ama.hashe / yi- in.kabi*
 9.wagon 9.PAST draw PASS FV COP 6.horse COP 9.ox
 ‘The wagon is drawn by horses / by an ox’

- (12) Zulu [S.42] (Taljaard & Bosch 1991:68)

- a. *uku.dla ku.dl- iw- a nga- aba.ntwana*
 15.food 15.eat PASS FV COP 2.child
 ‘The food is eaten by the children’
 b. *in.dlu i.fulel- w- a ngu- u.baba*
 9.house 9.thatch PASS FV COP 1a.father

'The house is thatched by father'

- c. *i.bhola li.dlal- w- a y izin.sizwa*
 5.ball 5.play PASS FV COP 10.young man
 'Soccer is played by young men'

Since at first sight, this strategy may appear to be less common, I would like to emphasize that in addition to languages from the Nguni and Sotho-Tswana clusters the same strategy is also found in several other languages which are not too closely related, such as Herero from Namibia and several languages from Malawi and Mozambique.

2.5 Without any morphological marking 'sth. has been done (Ø) agent'

Some languages do not require a morphological marker to be added to the agent phrase. Luganda and Haya are languages which simply move the agent phrase to a slot after the verb phrase when a direct object is promoted to the subject slot in the passive sentence. Example (13b) illustrates that the absence of a morphological marker indicates that the noun phrase which immediately follows the verb is understood to be the subject of a corresponding active sentence. If the spear is understood to be the weapon by means of which someone committed the act of killing, it requires the comitative *na-*.

- (13) Luganda [E.15] (Ashton et al. 1954:337)

- a. *y.a- kub- ibw- a Mukasa*
 3S.PAST strike PASS FV Mukasa
 'He was struck by Mukasa'
 b. *y.a- tt- ibw- a (na-) ffumu*
 3S.PAST kill PASS FV COM spear
 'He was killed by (with) a spear'

- (14) Haya [E.22] (Duranti & Byarushengo 1977:47)

- ebi.tooke bi.ka- cumb- w' omu-kâzi*
 8.banana 8.PAST cook PASS 1.woman
 'The bananas were cooked by the woman'

Interestingly also in Tonga which I have mentioned before, the locative marker which often precedes the agent phrase as in example (6) is not obligatory. Thus also in Tonga sentences may be found in which the agent and the patient simply appear to change their position when the verb phrase is at the same time marked for passive.

2.6 Combinations

The catalogue of different strategies used in the creation of agent marking devices should not lead to the wrong impression that these strategies are always clear-cut and easily distinguishable. As a matter of fact, the contrary often seems to be true. Different

strategies may be intertwined in different ways. Some of these will be outlined in the following paragraphs.

First of all, in many languages comitative and instrumental meaning are expressed by the same morphology. Therefore, in these languages, it may be a fruitless task to try to find out whether instrumental or comitative serves as the immediate source for the agent marker. As described above, Shona is a language which presents this particular property, cf. examples (4) and (5).

Also the locatives may come into play in connection with the comitative strategy. If someone does something together with someone else, a locative form is often used to express spatial proximity. Therefore, it is not unusual to have comitative, instrumental and locative expressed by the same morphological markers, as in Tonga where the locative prefix *a-* (cl. 16) serves to denote all three meanings, cf. example (6).

Apart from these examples, another point here is that copulae may often be used in combination with locative prefixes. An example has been given above, cf. example (8) from Luba. It is important to note, however, that the use of the copula is deemed secondary here. Often in these languages, the combination of a locative prefix and a copulative element has given rise to a preposition which does not preserve any of the verb-like features which may have been associated with the underlying copula, as in Luba:

- (15) Luba [L.31] (Burssens 1939:182)
ndiyí dié.lá kùdì mfùmù
 COP order from chief
 'It is an order from the chief (where the chief is)'

As an implication, what has been termed here the copula strategy for the expression of an agent in passive sentences is restricted to those cases, where the agent-marking device clearly and solely derives from an existential copula. Although these elements serve as predicative nuclei, they bear little verb-like features. For example, copulae tend not to be tense-marked. Therefore, one will also hardly find agent markers specified for tense. Nevertheless, there are examples which demonstrate unambiguously that the copula does not necessarily lose all of its morphosyntactic properties when it is used as an agent marker. This is shown by the Nguni languages who preserve the different allomorphs of the copula also in its use as an agent marking device.

Yet another case is that of Tsonga, in which the verbal copula *hi* (cf. 16a) is identical with the instrumental marker (16b). Also the agent in a passive sentence is expressed by means of this morphological marker *hi* (16c).

- (16) Tsonga [S.53]
 a. *ti.homu le.ti hi- le.ti nene.*
 10.COW DEM.10 COP DEM.10 good
 'These cows are the good ones' (Ouweland 1964:66)
 b. *ndzi.ta- famba hi- mi.lenge*
 1.SG.FUT walk cop 4.foot

'I will walk on foot' (Ribeiro 1965:488)

- c. *a.lum- iw- ile hi nyoka*
 3SG.bit PASS PERF COP snake

'He was bitten by a snake' (Ribeiro 1965:307)

The ultimate conceptual source of agent marking appears to be the verbal copula. A question which is difficult to answer is whether the instrument marking preceded the use of the copula as an agent marker diachronically, and may thus be regarded as having cleared the path towards this development. In Tsonga, this does not appear unlikely, since the functions of the copula reach even further: the copula *hi* may, e.g., also express cause, as in *a file hi chirhami* 'he died from cold' (Ribeiro 1965:488). One has to be careful, however, to assume that these functions are necessarily steps along the same path of grammaticalization, because they may rather turn out to be co-evolutions. It should be remembered in this regard that there are several languages in which the agent marker is formally identical with the copula, but instruments (as well as cause, comitative or any other function) are marked differently (cf. the examples from Nguni above).

The situation in Ronga is quite similar to that in Tsonga. The Ronga agent marker is formally identical with the instrument marker, both going back to a copula. An interesting phenomenon that has also been observed for Tsonga (Ouwehand 1964:114) can be found in this language. There is a syntactic construction which is reminiscent of the passive, because the patient is found in the subject slot and the agent is marked by the morphology as in a typical passive sentence, i.e. by the copula *hi*. However, the verb phrase is not obligatorily marked with the passive marker, cf. the contrast between (17a) and (17b):

(17) Ronga [S.54] (Junod 1896:165)

- a. *ši.yentš- iw- i hi mine*
 7.do PASS PAST COP 1SG.PRO
 'This was done by me'
- b. *ši.yentš- i hi mine*
 7.do PAST COP 1SG.PRO
 'This was done by me'

This particularity about Ronga in the use of active forms instead of passives without any further change in the constituent order and argument status of the noun phrases represents – so-to-speak – the mirror-image of the case which I have outlined for Luganda, Haya and Tonga, in which passive marking occurs, whereas the agent phrase does not receive any additional morpheme. Here, on the contrary we have a case in which the verb is not marked for passive, but still the agent is marked as such. The fact that for the above given passive sentence in (17a) the corresponding active verb form can be used interchangeably leaves a much heavier load to the agent marking device than in any other of the examples presented so far. Since the verb lacking *-w-* is in terms of voice ambiguous, it appears that the presence of an agent phrase alone suf-

fices to mark the entire sentence as passive. Of course, one would have to go into more detail here. There are important questions that should be tackled in fieldwork. First of all, one would need to know whether in these cases the agent phrase is obligatory. Another question that should be asked concerns the impact of specific verbs in passive sentences. It may be the case that for predications such as 'so. opens the window', passive marking is optional since the actual assignment of semantic roles appears so natural. However, other verbs which may be reciprocal could turn out to result more easily in ambiguous sentences. Therefore, it would not come as a surprise if passive marking on the verb would be more rigorously used in such cases, i.e. in connection with verbs like 'wash', 'beat', etc. It must be noted that these questions do not solely depend on specific (groups of) verbs. What matters here is rather a complex contextual information. In this sense a verb like 'hit' may be part of an unambiguous assignment pattern (as when someone hits an inanimate object), but it is more likely to become ambiguous when both agent and patient are willfully acting human beings. Therefore, it is possible that the optional use of the passive marker in (17) is due to the distinction between an animate agent and an inanimate patient.

To summarize this section, it should be noted that agent marking morphemes usually derive from elements which carry other grammatical functions. Generally, it is rather easy to pin down their conceptual origin and the respective grammaticalization paths. As a matter of fact, one may even wonder whether one should describe this phenomenon in terms of diachronic changes and grammaticalization, because the extension of a former meaning which apparently also comes to cover agent marking differs from many other instances of grammaticalization chains in one important respect: the extended meaning 'agent marking' does never become fully autonomous.

No example of a morphological marker was found which fulfils exclusively the function of agent marking in a passive sentence (the only possible exception being Ndonga, see below). This may be due to an important difference with regard to other well-attested grammaticalization chains in which original meanings of source concepts tend to bleach and often, original functions are at a certain stage discontinued. In the case of agent marking, one apparently rather deals with an extended meaning of a specific source concept – be it locative, comitative, instrumental, etc. – which can never be totally detached from the latter.

This might lead to a rather radical proposal. It could be suggested to dispose with agent marking as a grammatical category altogether. Since there are examples in which agent markers carry a functional load that goes clearly beyond that of their respective underlying source concepts, this proposal is rejected here. Ronga represents a language that should be borne in mind with regard to this issue. Since the morphological marking of the verb phrase is at least in some cases optional, the morphological agent marker appears to serve as the only explicit morphological means of characterizing the respective sentence as passive. An even stronger argument would, of course, be the existence of an independent agent-marking device in Ndonga (cf. 3.3). Unfortunately, the case is not entirely clear yet and deserves more attention. Cases like these, however,

suggest that one should not simply disregard agent marking as a potential grammatical category in Bantu.

3. The relationship between passive and agent marking

On the base of the information provided so far, it will be argued in this section that passive marking in the verb phrase and the marking of an agent noun phrase are phenomena which are not too closely related. There are different arguments for this opinion which will be outlined in the following three sections.

3.1 Diachronically independent developments in the domain of passives and agent expression

In the preceding section it has been shown that the formal ways of encoding an agent phrase may derive from a number of strategies which fulfill grammatical functions other than agent marking. It could be hypothesized that a given type of passive bears restrictions on the agent encoding mechanism that may or may not apply in a specific case. It will be shown, however, that such links are only relatively loose. In spite of the fact that the need for a morphologically more marked agent encoding is actually triggered by an inherent passive property, namely the demotion of the agent from the subject function, there is still a broad variety of formal ways of construal which are not simply determined by the type of passive. In other words, the diachronic origins and grammaticalization of passives are largely independent from that of agent markers.

If there was a close functional relationship, one would suspect a more stable development in the realm of the agent phrase encodings in Bantu languages, or at least some kind of co-variation involving both passive and agent marking. What can be observed, however, is a notable variation concerning the agent markers, whereas the formal aspects of passive constructions are rather stable and uniform across Bantu languages. Even if one takes into consideration functional variations of the passives in Bantu languages (e.g. the applicability to non-typical transitive action verbs which is not as uniform from one language to another), these phenomena do not, however, co-vary systematically with the phenomena to be observed in the expression of the agent phrase. Even among languages whose passive constructions differ formally from those of the bulk of Bantu languages, we do not find evidence for a close relationship between the ways in which a passive comes into being and the agent marking. Cases in which a passive is an obviously more recently grammaticalized form are important in this regard. Looking at the SW Bantu languages lacking a basic morphological passive in *-(ib)w-*, it is evident that the agent phrase can only be used after the third person plural pronoun has lost some of its semantic contents. Once it has desemanticized to the extent that it solely serves as a desubjective marker, an agent phrase may occur. At this point, however, there does not seem to be anything inherent in the respective

passive construction which would determine which specific way of encoding the agent should be used. The passive predication in itself – even in the case of e.g. Kimbundu where the construction is still relatively transparent – does not impose any restrictions in the sense of a preference towards or against a specific formal marking of the agent phrase.

If one looks into those languages making use of the basic morphological passive extension, there is further evidence for the assumption that links between passive constructions and agent marking are rather loose. In most Bantu languages, the passive is a completely grammaticalized, basic morphological construction. If one wanted to explain the acceptance or non-acceptance of agent phrases and the selection of a specific strategy as depending on the kind and degree of grammaticalization of a passive construction, the evidence from the basic morphological passives in Bantu hint at the opposite direction. The assumption turns out to be wrong that the relative age of a passive construction is proportional to the acceptance of overt agent phrases. If one looks at those examples where a passive is constructed by the use of the verbal extension *-(ib)w-* with a supposedly shared origin among the Bantu languages, one still finds a broad variety of ways in which the agent may be construed. Thus, the underlying conceptual structure of the erstwhile more transparent construction of the passive does not seem to interfere with the selection of a particular agent marking strategy.

3.2 Synchronic variation of agent marking and areal features

It has been stated that the agent marking strategies are much more varied than the formal marking of passives. If one looks into the distribution of the different strategies used to express agent nouns, one will notice that there is very little correspondence between agent marking and genetic family ties. This reflects the high variability of agent marking devices. As argued in the preceding section, they do not come as part and parcel with the formal and functional features of a specific passive construction. Rather it appears that they are much more susceptible to area influences. One will find similar strategies in contiguous areas, often disregarding genetic language boundaries. It is noteworthy that the impact of area contact may be manifest in different ways. Whether or not agents are allowed to stand in passive sentences may be an areal feature, since we find areas in which such a restriction is widespread and cross-cuts genetically determined language groups. The same holds for particular encoding strategies that may be found in languages belonging to various sub-groups being geographically close. It is important to note that the same strategy in different languages may be represented by cognate morphology; this, however, is not necessarily the case. If e.g. the copulative strategy is dominant in a given linguistic area, different languages may use their respective copulae which can be formally distinct. As a matter of fact, this is indeed the most common case. On the contrary, it is noteworthy that no example of a borrowed agent marking device has been encountered in the languages considered here. This is another piece of evidence showing that agent marking is a widespread functional requirement, but not a prominent grammatical category. Apparently the formal devices

which are used to express an agent phrase are not primarily understood to represent such a weak grammatical category as agent marking. At the same time, this does not keep the underlying mechanism of how an agent can be expressed from becoming a shared feature in a linguistic area.

A good example for such area influence is Shona which has been mentioned as an example for the comitative strategy. It must be noted, however, that there is considerable dialect variation in Shona and some of the regional varieties use strategies other than the one given in example (4). This leads to another point: languages do not necessarily rely on only one strategy to express an agent in a passive sentence. In Nguni for example the prevailing marking strategy is that of the existential copula. Notwithstanding locative constructions may also occur. Other examples are those of the facultative use of the locative prefix in Tonga, cf. (6).

This is an important example in the sense that this language represents one of the few cases in which a grammaticalization chain leading to zero agent marking can be demonstrated. Looking into other languages, in which zero agent marking does not alternate with another strategy, it is hard to find evidence that zero-marking is – so-to-speak – the final stage of a grammaticalization process concerning morphology, because one might suspect purely syntactic case assignment patterns to be responsible for the association of the agent role with a noun phrase in a specific syntactic position. For instance, for Haya and Luganda it could be argued that agent marking is entirely syntactic and does not have anything to do with morphology, cf. examples (13) and (14).

On the contrary, in Tonga zero-marking is certainly due to the disappearance of a former obligatory marker. If the language should once reach the stage of total loss of the respective marker, it might become difficult to distinguish such a case from those outlined for Luganda (and other languages from that region).

3.3 The use of agents in non-passive sentences

So far, all the arguments listed can be grouped under two headings. At first, diachronic aspects have been dealt with in order to show that passives and morphological devices to encode agents in passive sentences develop independently. After that, arguments have been presented which mainly deal with synchronic aspects of variation and areal features concerning the agent marking devices in order to show that they are rather independent from the passive constructions with which they may be used and that agent marking is a relatively weak grammatical category. Another aspect which has not been mentioned so far also serves as evidence arguing that agent marking and passive constructions are largely independent grammatical domains.

With relatively few exceptions the formal construction of passive in Bantu languages has been shown to be relatively homogeneous, whereas agent marking is more variable: languages seem to be more idiosyncratic in this regard. They choose from a broader variety of formal constructions. The respective conceptual sources have been described in detail, and it has been argued that agent markers are never fully de-linked

from the respective source meanings. Nevertheless, there is evidence that agent marking may develop into a substantial property of these markers. Despite the narrow bond between agent marking and a grammatical source domain, once agent markers have come into being, they can apparently extend to contexts other than the primary passive constructions which trigger their existence.

The first example below is from Luba. As shown in example (8) above, this language has a basic morphological passive with the extension *-w-*. In addition to this, there is an alternative passive which is based on a generalized subject construction involving the third person plural marker. Also with this supposedly more recent passive strategy, whose introduction may be due to areal influence from neighboring Bantu languages to the south-west, an overt agent phrase may be used. Its formal means of construction is the same as in the more frequently used morphological passive: the locative preposition *kùdì* derived from a copulative verb.

- (18) Luba [L.31] (Burssens 1939:182)

bà.ká.mú.sùm.á kù-dè nyòká

3PL.PERF.3SG.bite.FV 17.be snake

‘He has been bitten by the snake (literally: they have bitten him where the snake is)’

Another example is Ndonga. In this language, one encounters a phenomenon that calls for a special explanation. Ndonga has a morpheme *ká-* which introduces the agent in passive sentences. Interestingly, this does not only hold for passive sentences with a verb form containing the basic morphological passive morpheme *-w-*; instead, also sentences containing a neuter-passive verb extended by *-ik-* may take an agent phrase:

- (19) Ndonga [R.22] (Fivaz 1986:111)

ókinó ndjoká o.y.a.tál.ik.á

ká.á.ntu

a.yéhe

‘9.film DEM.9 AFF.9.PFV.watch.NTPASS.FV INSTR.2.person 2.many

‘This film is seen by many people’

Such an example poses a number of problems, which are difficult to solve on the basis of the published material on this language. There are, however, some aspects that should be taken into consideration. One of these is the fact that Ndonga is possibly a counter-example to the claim made above that no exclusive agent markers occur in Bantu languages. All the sources known to us so far from which an agent marker may derive do not seem to apply in Ndonga. The element *ká-* does not show links to the comitative *ná-* which is also used to mark instruments, nor to any of the locative noun classes (*pu-/pa-*, *ku-*, *mu-*). Looking into the range of copulae used in Ndonga, there is also no formally similar element which may have given rise to this particular agent marker. In this sense, the Ndonga case seems indeed noteworthy, because the sole function of *ká-* seems to be that of agent marking. A possible explanation is that the marker is actually a loan from neighboring languages in which it goes back to a copula. The closest attested occurrence of such an element as an agent marker is

Tswana (Sandilands 1953). A morphological loan from Tswana to Ndonga does not seem likely, however, and another explanation should be sought. An element *ká-* with a rather wide distribution in the SW Bantu languages has a copulative function, associating a nominal entity with a place. Sometimes it may also be used in order to single out an individual from a mass. Several ethnonyms apparently make use of this element (*Kaluvale* ‘a person of the Luvale group’ [the ethnic designation goes back to a toponym]). It is at present not possible to make a safe statement about whether there is a link between these elements.

4. Conclusion: Some general implications

In the introduction to this article three claims are made about the character of agent phrases in Bantu passives. First of all, the broad variety of different strategies that may be used in order to encode an agent as an oblique argument in a passive sentence has been illustrated. Among the various Bantu languages all different source concepts for agent marking devices that are known to us seem to be applied in one language or another. It is thus noteworthy that despite the relative homogeneity on the formation of passives in different languages, agent marking is a functional requirement that is relegated to the “responsibility” of each particular language. From this it follows that even within a given language, some variation can often be found in the selection of one of the possible strategies. At the same time, similar strategies are often found in neighboring languages irrespective of genetic family ties. It seems therefore safe to state that agent marking is as likely to be an areal feature as it is to be an inherited property of a given language.

Some characteristics, however, are common to all Bantu languages. The second claim made about agent phrases concerns such a general tendency found among all Bantu languages: its relative weakness as a grammatical category. Agent marking is in most cases a peripheral use of constructions which predominantly serve to express other meanings. No language has a marker whose function is restricted to that of marking an agent phrase in a passive sentence. The diachronic origin of passive markers is usually easy to trace, since the respective morphological elements serve their original functions as well. The only possible exception is presented by those languages in which no morphological element is used in order to mark the agent phrase in a passive sentence. It must be noted, however, that in these languages different types of oblique arguments may usually be used in a sentence without the necessity of having them marked in any special way as shown by many languages of the Great Lakes region. Thus, the hypothesis receives considerable support that agent marking is indeed not an autonomous grammatical category. From the facts in some languages (e.g. Ronga, Ndonga) it appears nevertheless that agent marking as a grammatical category cannot altogether be done away with. In any case – borrowing expressions from Kemmer (1993; 1994) – it is appropriate to call its status as a grammatical category less funda-

mental than that of cognitively more primary categories such as comitative, locative and instrumental.

The validity of the third claim made at the beginning of this article results from a combination of the other two hypotheses outlined above and for which support has been adduced. Agent encoding is not an integral part of the grammatical domain of passive as a voice phenomenon. Despite formally and functionally similar passives, languages differ with regard to the acceptance of agent phrases in these constructions. Neither the acceptability nor the specific way in which an agent phrase may be encoded depend in any way on the specific passive construction present in a given Bantu language. Furthermore, agent constructions – despite their relative weakness as a grammatical category – have in some cases been recorded to occur in constructions other than passives. For these reasons, also the third claim receives strong support: agent phrases and passive voice phenomena are largely independent grammatical domains in the Bantu languages.

Notes

1. I am grateful to a number of participants of the symposium in Sankt Augustin for their valuable comments: Denis Creissel, Talmy Givón, Suzanne Kemmer and Thilo Schadeberg. In addition to this I wish to thank Yvonne Treis for her comments on a written draft version of this article. Of course, none of these persons can be held responsible for any possible shortcomings or mistakes, which are my responsibility.
2. Interestingly, the original translation by Burssens makes use of a passive sentence in order to render the meaning in Dutch: *Het is een bevel dat door den chef werd gegeven (waar de chef is)* ‘This is an order that was given by the chief (where the chief is)’. The internal structure, however, is closer to the English translation as given above.
3. A possible caveat is that the optional omission of the passive extension may also depend on the semantic roles involved in the respective sentence as described in the previous paragraph.
4. The claim made here is that the diachronic origin of a passive construction is not relevant to the selection of the agent encoding strategy. Interestingly, even a stronger claim seems to hold. One might suspect that the passive type may at least determine whether or not an agent phrase may co-occur at all. As will be shown in some more detail later, this is obviously also not the case in Bantu.
5. Notwithstanding, there is probably also syntactic evidence in the sense that, in general, the association of semantic roles and syntactic functions of noun phrases seems rather variable in the latter languages (cf. Kimenyi 1988). This is not true for a language such as Tonga.

References

- Ashton, E. O., Mulira, E. M. K., Ndawula, E. G. M., & Tucker, A. N. (1954). *A Luganda Grammar*. London: Longmans, Green and Co.

- Burssens, A. (1939). *Tonologisch Schets van het Tshiluba (Kasayi, Belgisch Kongo)*. Antwerpen: De Sikkel [Kongo-Overzee Bibliotheek 2].
- Collins, B. (1975 (1962)). *Tonga Grammar*. Lusaka: Neczam.
- dos Santos, E. (1962). *Elementos de gramática Quioca*. Lisbon: Agencia-Geral do Ultramar.
- Duranti, A., & Byarushengo, E. R. (1977). On the notion of 'direct object'. In E. R. Byarushengo, A. Duranti, & L. M. Hyman (Eds.), *Haya Grammatical Structure* (pp. 45–71). Los Angeles: University of Southern California [Southern California Occasional Papers in Linguistics 6].
- Fivaz, D. (1986). *A Reference Grammar of Oshindonga*. Windhoek: The Academy [African Studies of the Academy 1].
- Fortune, G. (1955). *Analytical Grammar of Shona*. London: Longmans, Green & Co.
- Givón, T. (1976). Topic, pronoun and grammatical agreement. In C. N. Li (Ed.), *Subject and Topic* (pp. 149–188). New York, NY: Academic Press.
- . (1981). Typology and functional domains. *Studies in Language*, 5 (2), 163–193.
- . (1986). Prototypes. Between Plato and Wittgenstein. In C. Craig (Ed.), *Noun Classes and Categorization. Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983* (pp. 77–102). Amsterdam: John Benjamins.
- , & B. Kawasha (2001). Indiscrete grammatical relations: The Lunda passive. Manuscript: University of Oregon, Institute of Cognitive & Decision Sciences, Technical Report No. 01-05, pre sented at the International Symposium in Sankt. Augustin, May 2001.
- Haspelmath, M. (1990). The grammaticization of passive morphology. *Studies in Language*, 14 (1), 25–72.
- Heine, B. et al. (1993). *Conceptual Shift. A lexicon of grammaticalization processes in African languages*. Köln: Institut für Afrikanistik [Afrikanistischen Arbeitspapiere 34/35].
- Junod, H. A. (1896). *Grammaire Ronga suivie d'un manuel de conversation*. Lausanne: Georges Bridel.
- Keenan, E. (1985). Passive in the world's languages. In T. Shopen (Ed.), *Language Typology and Syntactic Description*, Vol. 1: *Clause structure* (pp. 243–281). Cambridge: Cambridge University Press.
- Kemmer, S. (1993). *The Middle Voice*. Amsterdam: John Benjamins [Typological Studies in Language 23].
- . (1994). Middle voice, transitivity, and the elaboration of events. In B. Fox & P. J. Hopper (Eds.), *Voice: Form and function*. Amsterdam: John Benjamins [Typological Studies in Language 27].
- Kimenyi, A. (1981). Linkless Clauses in Bantu. *Studies in African Linguistics, Supplement*, 8, 85–89.
- Kimenyi, A. (1988). Passives in Kinyarwanda. In M. Shibatani (Ed.), *Passive and Voice*. Amsterdam: John Benjamins [Typological Studies in Language 16].
- Louwrens, L. J. et al. (1995). *Northern Sotho*. München: Lincom Europa [Languages of the World Materials 30].
- McLaren, J. (1955). *A Xhosa Grammar*. Cape Town: Longmans, Green and Co.
- O'Neil J. (1935). *A Shona Grammar. Zezuru Dialect*. London: Longmans, Green and Co.
- Ribeiro, A. (1965). *Gramática Changana (Tsonga)*. Caniçado: Missão de S. Vicente de Paulo.
- Sandilands, A. (1953). *Introduction to Tswana*. Lobatse: Bechuanaland Book Centre.
- Smith, E. W. (1907). *A Handbook of the Ila Language*. London: Oxford University Press.
- Taaljard, P. C., & S. E. Bosch (1991). *Handbook of Isizulu*. Pretoria: Van Schaik.

Grammaticalization of switch reference*

Motivation and means

Zygmunt Frajzyngier

University of Colorado

The aim of this paper is twofold: The first is to demonstrate that grammaticalization can be motivated by the internal properties of a grammatical system. This motivation does not figure in most contemporary studies of grammaticalization. The second aim is to carry out a specific study of the grammaticalization of switch-reference coding. The system-internal motivation is the principle of functional transparency, which states that if a language has grammaticalized a given functional domain, the role of relevant elements in a clause must be transparent with respect to this domain. If a language has grammaticalized the category 'subject', there must be means to identify not only which element of the clause is the subject but also what is the referent of the subject. In certain types of clauses the identification of the subject as the same or as different from that of the preceding clauses is obligatory. The present study demonstrates how morphological switch-reference markers grammaticalized from two types of pronouns.

o. The aim of the paper

I take the functions of same reference-switch reference systems to indicate "whether a subject or other argument of a following verb has or has not the same referent as that of the verb preceding" (Matthews 1997:365). This definition is much broader than the one usually accepted in Amerindian literature, where switch reference is limited to specific means of coding switch reference, viz. markers added to verbs. Although many of these markers are actually clausal conjunctions (Mithun 1999), they are taken by some to be the canonical cases of switch reference (Finer 1985; Stirling 1993; Broadwell 1997). Grammaticalized switch-reference coding through conjunctions has also been claimed to exist in Papuan languages and Polish (Haiman 1983; Frajzyngier 1987).

The motivation for the existence of switch reference coding has been stated by Haiman and Munro as follows "The function of the switch-reference is to avoid ambiguity of reference' (1983:xi). This motivation, which could be extended to a number

of other domains in language structure, does not explain why some languages have grammaticalized means to code switch reference and other languages have not.

The first aim of this paper is to explore the motivation for the grammaticalization of switch-reference marking. It is shown that its motivation is language internal, viz. to satisfy the principle of functional transparency. The principle states that if a given domain has been grammaticalized in a language, the function of a relevant item for this domain must be transparent.

The second aim of the present study is to demonstrate that which function a given formal means is used for does not depend on the properties of the formal means but rather on the interaction of this means with other formal means in the language. This goal is achieved by showing that the same formal means can be used in one language to code coreference and in another language to code switch reference. The way coreference and switch reference are coded depends on whether the language has a distinction of class or gender (or some other system, such as an honorific system) and on how many means to code participants the language has.

Throughout this study I use the term 'preceding clause' and 'following clause' as in Matthews 1997, to refer to preceding and following clause constructions that require transparency with respect to identity of the argument.

0.1 Motivation for coreference and switch-reference coding

The present study is limited only to reference to subject. Reference to object is discussed in study in progress. Moreover, I do not include here the coding of coreference through logophoric pronouns or the coding of disjoint reference through other pronouns, those subjects having been amply discussed in other papers (cf. Frajzyngier 1997 and references there, and Stirling 1993).

0.2 The hypothesis

I propose the following hypothesis with respect to coreference–switch-reference coding: If the language has grammaticalized the category 'subject', then not only the item that has the function of the subject must somehow be marked, but also the referent of the subject must be transparent. This hypothesis is more specific than the hypothesis by Haiman and Munro (1983) and it explains why some languages do not have the grammaticalized means to code switch reference. I propose two types of evidence for the hypothesis: affirmative evidence, showing that languages that have the category 'subject' also have grammatical means to code the reference to subject; and negative evidence, consisting of the analysis of one language that does not have the category 'subject' and, in agreement with the hypothesis, does not have the grammaticalized means to code switch reference.

1. The affirmative evidence

The evidence for this hypothesis is provided by analyses of three languages, each of which codes coreference and switch reference in a different way: English, which has not grammaticalized morphological switch-reference markers and codes switch reference through the use of nouns; Polish, which has grammaticalized pronouns as switch-reference markers; and Mina (Central Chadic), which has grammaticalized two switch-reference markers.

1.1 Coding switch-reference through nouns

In English, reference coding means include pronouns and nouns (determined in various ways) but not reference coding through agreement. Pronouns in English have no inherent value with respect to the coreference – switch-reference distinction and can be used in constructions coding both functions. If the preceding and the following clause differ in the person of the subject, the pronoun in the following clause codes switch reference (all examples, unless indicated otherwise, come from the London-Lund Corpus. The examples from the London-Lund corpus retain their original numbering for the ease of reference):

1_1_0 <177 B> ^now [i] if !these papers come . by the :twenty-ninth of J\une#
 1_1_0 <178 B> and you ^send them through to m\e# .
 1_1_0 <179 B> ^in L/oughton# -

If the pronouns of the preceding and the following clause are the same, they code coreference. If the pronouns of the preceding and the following clause are different, they code switch reference. If the subject of the preceding clause is nominal and the subject of the following clause is a pronoun having the same value for gender and number as the nominal subject, the pronoun codes coreference. All three possibilities, viz. the same pronouns, different pronouns, and the coreference of the nominal and pronominal subject, are illustrated in the following example:

1_4_0 <712 B> ^I s=aid#
 1_4_0 <713 B> well I ^don't !\actually 'do it like th/is#
 1_4_0 <714 B> and ^he said you kn/ow#
 1_4_0 <715 B> that . [ko:p] ^how
 1_4_0 <716 A> *((^how else 2 sylls))*
 1_4_0 <715 (B> will 'students* j\ustify#
 1_4_0 <717 B> ^staying in L\ondon#
 1_4_0 <718 B> ^spending m\oney you _see#
 1_4_0 <719 B> if they ^don't get *t\ought#*
 1_4_0 <720 A> ^y\es#*

The absence of an overt subject codes coreference with the preceding subject whether first, second, or third person:

1_1_0 <91 B> be^cause I'm !going to Madr\id# .
 1_1_0 <92 B> on the ^t\enth#
 1_1_0 <93 B> and ^coming back on the twenty-n/inth# -
 1_1_0 <50 B> ^you give* them the :l\ot {^you s=ee#}# *.
 1_1_0 <51 B> ^that's the **p\oint##
 1_1_0 <52 B> ((and)) ^make sure that there's :s\omething#
 1_1_0 <53 B> [:@] ^fairly :cl\osely rel/ated#

If a construction requires a subject, as is the case in embedded and concessive clauses, the pronominal subject that has the same value for number and gender as the subject of the preceding clause codes coreferentiality:

1_1_0 <115 B> I'll be at ^h/\ome#
 1_1_0 <116 B> and al'^though I'll be doing CS/C_stuff#
 1_1_0 <117 B> and ^that kind of th=ing#
 1_1_0 <118 B> ^I can always 'put it on one *s/ide#*
 1_1_0 <119 B> and ^get on with the p\aper#

Coreference may also be coded by a full noun, but those are cases where the use of a pronoun would be considered disrespectful:

2_4_1 <257 A> *^yes \I'm* - I'm pre^pared to m\et ^meet my :m\aker#
 2_4_1 <258 ALL> (- laugh)
 2_4_1 <259 A> ^whether my !m\aker *is* pre'pared
 2_4_1 <260 B> *I'm*
 2_4_1 <259 A> for this /ordeal#
 2_4_1 <261 A> is an^other qu\es+tion#+

How switch reference is coded depends on the subject of the following clause. If the subject of the following clause is the first or second person, it may code switch reference without any additional marking. With a third-person subject having the same features for number and gender, switch reference may be coded only by full nouns, with or without determiners. Pronouns may not be used in the following clause:

1_1_0 <172 A> [:@:m] . the ^last meeting of ((the)) Council
 Com:mission is about the middle of Jul/y# -
 1_1_0 <173 A> ^and - [:@] therefore the candidates wouldn't be
 able to re!c\ieve their cer_tificates#

In narratives with two participants with the same gender, number, and person, a clause has at least one of the participants coded by a full noun:

This is a story of Uncle Rabbit and the coyote. The rabbit came to a big rock, and there he deceived the coyote. He was leaning on the rock when the coyote came by.

“What are you doing, brother?” the coyote asked the rabbit.

“Come here quickly, brother, the sky is falling down on top of us. Lean against the rock and hold it up while I go for a stick. We’ll prop it up with that,” said the rabbit to the coyote.

“All right,” said the coyote and began holding it up with all his might. Since the coyote was so stupid, he did exactly what the rabbit told him to. The rabbit had said that he was going to get a stick, but he went and left the coyote holding up the rock. When the rabbit didn’t return the coyote shouted: (http://www.folkart.com/latitude/folktale/tale_1.htm)

Notice that in the above text even if a pronoun is used for one participant, the other participant is still coded by a full noun in the same clause; hence any potential ambiguity with respect to the antecedent for a pronoun can be resolved through a simple computation, by ruling out the participant coded by the full noun. Were the full nouns in the above narrative to be replaced by pronouns, the text would be incomprehensible:

This is a story of Uncle Rabbit and the coyote. He came to a big rock, and there he deceived him. He was leaning on the rock when he came by.

“What are you doing, brother?” he asked him.

“Come here quickly, brother, the sky is falling down on top of us. Lean against the rock and hold it up while I go for a stick. We’ll prop it up with that,” he said to him.

“All right,” he said and began holding it up with all his might. Since he was so stupid, he did exactly what he told him to. He had said that he was going to get a stick, but he went and left him holding up the rock. When he didn’t return he shouted:

Although English does not have morphological means to code switch reference, it has the syntactic means, viz., the use of the full nouns to code switch reference.

1.2 Pronouns as switch-reference markers

If a language codes participants through full nouns, pronouns, and coding on the verb (agreement), then it has one more means of coding reference than a language without an agreement system. However, which means is used to code what function has to be empirically tested. Consider Polish, which codes the subject on the verb in all tenses and aspects. There is an important difference with respect to tenses in that in the past and future tenses the language codes gender for all persons and numbers, whereas in the present tense gender is not coded on the verb.

Full nouns are not used to code coreference or switch reference. If only one noun is used, its function is that of topicalization. If two or more nouns are used, their func-

tion is that of setting the participants for the subsequent discourse. The agreement system and the subject pronouns function in the coding of coreference and switch reference. As shown in Frajzyngier 1997, the coding on the verb ('agreement') in literate and non-literate varieties of Polish is a means to code coreference with the immediately preceding subject. Here is an example from literate language with nominal antecedents in bold face (unless otherwise indicated, examples from literary Polish come from sources to Kurcz et al. 1990):

- (1) *Hrabia Zygmunt nie mieszka już w Toronto.*
 count Zygmunt NEG live:3SG:PRES already in Toronto
 'Count Zygmunt does not live in Toronto anymore'
Przeniósł się na wieś pod Riwerton.
 move:PAST:3M:SG REFL PREP country near Riverton
 'He moved into the countryside near Riverton.'
Nabył tam farmę niedaleko jeziora Winnipeg.
 buy:PAST:3M:SG there farm not far lake:GEN Winnipeg
 'He bought a farm there not far from Lake Winnipeg.'
Pisze, że krajobraz bardzo mu przypomina nasz
 write:PRES:3SG:M COMP landscape very 3M:DAT remind:PRES:3M our
Dembowiec
 Dembowiec
 'He writes that the landscape reminds him very much of our Dembowiec.'

Here is an example with a pronominal antecedent:

- (2) *Ach, on jest tak mało gadatliwy.*
 ah 3M:SG be:3SG so little chatty:M:SG
 'Ah, he is not a chatty fellow.'
nie umie się nawet pochwalić
 NEG can:PRES:3SG REFL even praise:PRF
 'He cannot boast, even a little bit.'

Pronouns in Polish code switch reference with respect to the immediately preceding subject. The necessary condition for the use of the pronoun is a previous mention in discourse or the presumed knowledge of the antecedent by the addressee.

- (3) *Miała dziewczka kawalera; vumar, nie żył długo.*
 have:PAST:3F girl boyfriend die:3M NEG live:M long
 'A girl had a boyfriend; he died, didn't live long.'
to wzieni go i zanieśli do kościoła
 COM take:PAST:3PL:M 3M CONJ carry:3PL:M to church
 'So they took him and brought him to a church.'
i ona go barz żałowała
 CONJ 3F:NOM 3M:ACC very grieve:PAST:3F
 'And she grieved for him.'

Un sie nie kazał pochować

3M:SG REFL NEG order:3M bury

'He asked not to be buried' ...

I ona tam chodziła do niego zawdy do kościoła

and 3F there go:PAST:3F to 3M always to church

'And she always was going to church to [be with] him'

a un fstawał s trumny, a ona sy staneła na kazalnicy

CONJ 3M get-up:3M from casket CONJ 3F REFL stand on pulpit

'And he would get up from the casket and she would stand on pulpit' (Nitsch 1960: 148)

(4) *on jej przejście daje, a ona się cofa*

3M 3F:DAT passage give:PRES:3SG CONJ 3F:SG REFL retreat:3SG:PRES

'He lets her pass, but she retreats.'

In the present tense, as in the above clause, the omission of the third-person pronoun would code coreference. But even if the subject of the following clause is first-person, and is coded on the verb, contrast must be coded by subject pronouns. The omission of the first-person pronoun *ja* in the following example would make it an ungrammatical sentence:

(5) *Nie. On nie. Ja cię będę oskarżał*

no 3M:SG NEG 1SG 2SG:ACC be:FUT:1SG accuse:M

'No, not him. I will be accusing you'

Switch reference may be the only function of the pronoun, without the contrastive focus function:

(6) *stan ten nazywamy absolutnie pierwotnym narcyzmem. Trwa on*

stage DEM call:PRES:1PL absolutely primary narcissism lasts 3M

tak długo,

so long

'This stage we call "absolutely primary narcissism." It lasts so long ...'

In the following example, the third-person masculine pronoun *on* does not refer to the immediately preceding masculine subject *Bóg* 'God' but rather to another masculine singular antecedent in a preceding discourse:

(7) *zwierzęta, zwierzętami, Panie administratorze, Bóg z*

animal:PL:NOM animal:PL:INSTR Sir:VOC administrator:VOC God with

nimi.

3PL

'Who cares about animals, Sir Administrator, God be with them.'

Ale dlaczego on ludzi drażni tymi otwartymi
 But why 3M:SG people:ACC tease:3SG 3PL:INSTR open:PL:INSTR
klatkami.

cage:PL:INSTR

‘But why does he provoke people by keeping those cages open?’

First- and second-person pronouns are used as switch-reference markers in situations where the subjects change. In Polish the situation is additionally complicated by the existence of the formal second-person addressee form *Pan* and its feminine and plural variants. This form of address and reference can be followed by a proper name or a title. Such a form of address behaves similarly to pronouns with respect to the coding of reference. In the following example the subject of the first clause is coded only on the verb. In the complement clause the honorific form of address *Pani* ‘Madam’ is used:

- (8) *mam nadzieję, że nie ma Pani nic przeciw temu*
 have:1SG hope:ACC COMP NEG have:3SG Madam nothing against this
 ‘I hope that you have nothing against it.’

In the next sentence the 1st-pers. subject is coded by the first-person pronoun (rather than any coding on the verb) because the speaker is setting up a contrast with another participant:

- (9) *ja. nie. ale on napewno miałby.*
 1SG NEG but 3M:SG certainly have:3M:SG:HYP
 ‘I don’t. But he certainly would have.’
nie, niepotrzebnie się Pani męczy, ja i tak nie
 NEG unnecessarily REFL Madam tire 1SG CONJ so NEG
uwierzę.
 belief:FUT:1SG
 ‘You, Madam, tire yourself unnecessarily. I won’t believe you anyway.’
Pani mnie nie cierpi,
 Madam 1SG:ACC NEG suffer:3SG
 ‘You cannot stand me.’
wiem. nic zresztą dziwnego, ja Pani też nie
 know:1SG nothing anyhow strange:GEN 1SG Madam:ACC also NEG
znoszę
 support:PRES:1SG
 ‘I know. Nothing strange about it. I can’t stand you either’ (Sources)

Pronouns can be bound by an immediately preceding noun when the noun is an adjunct. In each case the pronoun is still a switch-reference marker with respect to the subject of the preceding clause:

- (10) *nie myśl tyle o nim.*
 NEG think:2SG so much about 3M:INSTR
 ‘Do not think about him that much.’

pobrudzisz tylko materiał i oberwiesz od Leontyny.
 dirty:FUT:2SG only cloth CONJ get:2SG from Leontine
 ‘You will dirty the cloth and will get punishment from Leontine.’

ona ma dzisiaj bardzo ciężką rękę.
 3F have:3SG today very heavy:F hand
 ‘She has a very heavy hand today.’ (Sources)

- (11) *poczekaj, skoczę do sąsiadki po radę.*
 wait! jump:FUT:1SG to neighbor:F PREP advice
 ‘Wait, I will run to the neighbor to get advice.’

ona ma do interesów dobrą głowę.
 3F have to business good head
 ‘She has a good head for business.’ (Sources)

If the antecedent is a numeral larger than one, the pronoun may have an antecedent in the set referred to by the numeral:

- (12) a. *staliśmy z twoją matką przy oknie i*
 stand:PAST:1PL:M with 2:SG mother:INSTR by window and
patrzyliśmy na dwoje młodych szczęśliwych ludzi,
 look:PAST:1PL:M at two:N young:PL happy:PL people
idących ulicą przed naszymi oknami.
 walk:PART:PL street:INSTR in front 1PL:PL window:PL
 ‘We were standing with your mother at the window and looked at two young happy people walking on the street in front of our windows.’
Ona niosła na ręku małe zawiniątko.
 3F:SG carry:3SG:F on hand small package
 ‘She carried a small package.’ (Sources)

Omission of the third-person feminine pronoun *ona* in the above clause would result in a grammatical but uninterpretable sentence because the subject of the second clause could not be identified.

- (12) b. *?niosła na ręku małe zawiniątko.*
 carry:3SG:F on hand small package
 ‘She carried a small package.’ (Sources)

2. Two sources of switch-reference markers

2.1 The reference coding system in Mina

Mina (Hina), a Central Chadic language spoken in the Far North Province in Cameroon, has two morphological switch-reference markers. In order to situate these markers within the grammatical system of the language here are some basic facts about

the grammar. Mina codes grammatical relations through configuration, in the order SVO. Subject pronouns distinguish between first, second, and third person, and two numbers, singular and plural, the singular being the unmarked number. The language does not have gender distinction anywhere in its grammatical system. There is a rich system of determiners coding proximate and remote previous mention, as well as coding a noun as known, and therefore outside of the domain of inquiry on the part of the addressee. In addition to nouns, pronouns, and the absence of subject, Mina has two pronouns that serve as switch reference markers. Each pronoun codes a different type of switch reference. In order to understand the function of these pronouns, one has to understand the function of other elements in the coding of reference.

A noun can serve as the subject without any determiners when it is non-topicalized (nominal subjects are boldface):

- (13) *í nd rà í nd rà vàη wà kà dā*
 3PL walk D.HAB 3PL walk D.HAB rain start INF fetch
 ‘While they were walking the rain started to fall.’
- (14) *ngwáy skàn-yū vlàgám rà dāhà*
 PL.ADDRESSEE thing-PL talk D.HAB exist
 ‘Hey, there is something talking there!’

Although pronominal subjects of the first and second person have to be used whenever the subject is first or second person, the function of the third-person subject pronouns differs considerably from the function of the third-person subject pronouns in either English or Polish. The function of the third-person subject pronouns has important implications for the coding of coreference and switch reference.

One function of the third-person subject pronouns, singular *a* (with low or high tone depending on the aspect and mood) and plural *i*, is to code the subject when the nominal subject is topicalized. The construction has the form noun – topic marker – pronominal subject – verb. The topicalization marker is the phrase-final form of the last element of the noun phrase. If the phrase-final form is a demonstrative or a pronoun, it receives the suffix *n*: The pronoun *míndí* ‘another’ has the phrase-final form *míndéη*:

- (15) *míndéη à ndí ləm bíη*
 another 3SG HAB build house
 ‘One builds a house.’
- (16) *híd-yū wécín í tètè nfád*
 man-PL DEM 3PL.3 PL four
 ‘There were these four men.’

The second function of the third person pronominal subject is as follows: In the embedded clause and in the apodosis clause it codes coreferentiality with the preceding subject:

- (17) *hìdì wèhín à zá ván à n kà dā á gār kà*
 man DEM 3SG COMP rain 3SG PREP INF fall 3SG want INF
nd-á-k kàsəm skù
 touch-OBJ-1SG body NEG
 ‘This man said, “Rain when it falls will not touch me.”’

The third function of the third-person subject pronoun is to code the coreferential subject in a clause in the perfective aspect formed through the reduplication of the verb. When the verb is in the reduplicated form, the subject must be coded. So even if the subject has already been mentioned in previous discourse, it is coded by a pronoun:

- (18) *gáw zá á tá-n déy sà n kà vlánj tàn*
 hunter COMP PREP GEN-1SG also 1SG PREP INF CROSS DEM
 ‘The hunter said, “I will also cross it.”’
bàt á bát gādéd ngən bál bál bál á bál á nà
 take 3SG take arrow 3SG shoot shoot shoot 3SG shoot PREP PREP
lāk wát wà cūr rá
 river DEM straight (Ful.) D.HAB
 ‘He took his arrows and shot them straight into the river.’

Coreferentiality of subjects in sequential clauses is coded by the absence of a pronominal or a nominal subject:

- (19) *séy čáp á māl vlámbáy wà kà dāp ndád ká n*
 then chap! 3SG catch stick DEM down only lay down down PREP
skən ngən bát
 thing 3SG take
 ‘Then she chap!, caught the stick, put it down, took her thing’
- (20) *mà ləm bínj rá’ dřiš ngád dřiš*
 REL build house dig mud mix mud
 ‘The one who builds the house digs the mud, mixes the mud’
ləm bínj vlá hámás nd-á hàvl ká wán kà nà mən
 build house cut straw go-DIST thatch AFF lie inside PREP L.ANAPH
 ‘Builds the house, cuts straw, thatches the roof, lies inside it.’
- (21) *mà pèdák njũl bát pèdák á pèdák-á nástà ngən nà*
 REL split grass took split 3SG split-DIST enter (Ful) 3SG PREP
mən
 L.ANAPH
 ‘The one who splits grass, split a stalk of grass and entered it.’
tsú ngən kà nà mən ván kà mbàlém dá skù
 enter 3SG inside PREP L.ANAPH rain INF touch exist NEG
 ‘He entered it [the grass], rain did not touch him.’

2.2 Switch-reference coding in Mina

Switch reference in Mina is coded by two forms: the third-person singular pronoun *a* and the independent third-person pronoun *mbí*.

If the first clause has a nominal subject, the use of the third-person pronominal subject in the next sequential clause (but not in the next sentence) codes switch reference to another subject previously mentioned in discourse. The evidence that what follows is another clause is provided by the low tone on the verb *tàŋ* ‘go’:

- (22) *kwáyàŋ tàŋ à zá s kà dál tséy zà*
 squirrel go 3SG COMP 1SG INF do finished AUX
 ‘The squirrel went. He [the monkey] said, “I finished.”’

The following fragment illustrates the use of *a* as the switch-reference marker and the absence of the pronoun for the same-reference subject:

- (23) *séy bàhámàn wurtə páláh à zá ndè séytiinà bá dàp*
 then Bahaman leave (Ful.) out 3SG COMP go ‘call’ again
 ‘Then Bahaman went out. She said to him, “Go make that call again.”’
 (*séytiinà* ‘name in Fula of muezzin’s call in the morning’)
bàhámàn nd-á gār
 Bahaman go-DIST stand
 ‘Bahaman went and stood’
díyà séitin go wàcín syì
 start ‘call’ call DEM COM
 ‘He started to make the call.’

So the same form that in complement clauses and in the perfective aspect coded through configuration codes coreferentiality with the preceding subject, in sequential clauses codes switch reference with respect to the preceding subject. This fact indicates that the grammatical functions of a linguistic form, in this case opposite functions, coreference and switch-reference coding, do not necessarily depend on the inherent properties of the form, but rather on the structure of the particular language and the interaction of various coding means within a given functional domain.

The third-person pronoun *mbí* is realized as *mbə* in the phrase-internal position, and as *mbén* in the phrase-final position. This form is glossed as ANAPH because the minimal condition for the use of this form is to have an antecedent in the previous discourse. The phrase-internal form *mbə* is the unmarked form. It codes contrastive focus on switch-reference subjects. The subject of the preceding clause can be nominal or pronominal. When the subject of the following clause is third person and it is non-topicalized, the switch-reference marker is the only subject marker in the clause:

- (24) *tsáy kwáykwáy mə ndəv-á-y zà bitiríd á káyàk*
 so hyena REL fall-DIST-STAT AUX heavily PREP earth
 ‘So, hyena_i fell down, heavily on the ground.’

tséy mbi fát fát bákùl tà kwáykwáy wàcín
 SO ANAPH skin skin hide GEN hyena DEM
 ‘Then he_i skinned that hyena ...’

- (25) *báy wílè á dāmù mbi nd-á bàt àdá*
 chief still PREP bush ANAPH go-DIST take food
 ‘The chief_i is still in the bush. He_i came to take the food.’

- (26) *bày zá gár kà zá mbà gár ábà nd-á ngàñ*
 chief COMP leave AFF AUX ANAPH stand ASSC go-DIST 3SG
 ‘The chief_i said, “Get out of here.” He_i stood up and went back.’

Here is an example of a pronominal subject in the preceding clause and the switch reference marker in the following clause:

- (27) *à-ndà r skù mbi mì žèbér tà tkóñ*
 3SG-go D.HAB NEG ANAPH REL follow GEN 2SG
 ‘If it does not go, she should follow your [advice].’

The form *mbéñ* codes the topicalization of switch reference. The evidence that the form *mbéñ* is a topicalization marker is provided by the fact that the entity so marked is a persistent topic in discourse beyond the clause where it is the subject:

- (28) *káf yà í yà-há-ú bàt í yàt zà dzáñ kà á*
 morning call 3PL call-DIST-3SG take 3PL take AUX close AFF PREP
bìñ
 room
 ‘They called him in the morning, locked him in a room.’

tíl á dāmù
 leave PREP bush
 ‘He [the one who was doing the locking] went into the bush.’
séy mbáñ bàt nèwén tà dindém díyá bām
 SO ANAPH take salt GEN sweet start eat
 ‘Then he [the one who was locked] took the sugar and started eating.’

- (29) *tíl ngàñ á wtá vàn gòyl gòyl bày á kàtàf mbàn zàm*
 leave 3SG PREP home rain hit hit chief PREP road ANAPH eat
wàdà tók vàn tók zà
 food finish rain finish AUX
 ‘When he_i returned home, the rain hit the chief on the road. He_i finished eating and the rain also finished.’

When the switch-reference marker is used for topicalization, it may be followed by subject pronouns, like any other topicalized subject:

- (30) *séy pá í vəl-á-ñ jèni*
 so give 3PL give-OBJ-3SG ax
 ‘So, they gave him an ax.’

mbéŋ à tik-é tàlàn káyàk kà jèni
 ANAPH 3SG tilt-DIST head earth PREP ax
 ‘He inclined his head because of the ax.’

- (31) *tséy wàž-yî bá-yî zá vl-á nènén mbéŋ à zá*
 so children-PL chief-PL COMP give-OBJ 1PL.EXCL ANAPH 3SG COMP
gwáɗ á bîŋ kàcîŋ
 plenty PREP room DEM
 ‘The children of the chief said, “Give [it] to us.”
 He said, “There is plenty here in the room.”’

If the event has many participants, thus providing the possibility of ambiguous interpretations, the switch-reference marker in sequential clauses may be followed by the pronominal subject:

- (32) *bày ábà nd-á ngèn séy mbéŋ gàmíyíɗ-yíi mà nd-á-y zà*
 chief ASSC go-DIST 3SG so ANAPH chimp-PL REL go-DIST-STAT AUX
kà báɓ pây wàcîŋ
 INF eat tree DEM
 ‘The chief went back. Then the monkeys came to eat the fruit of that tree.’

2.3 The source of the switch reference function of *mbí*

The form *mbí* is an independent third-person anaphor. In its full, i.e. phonologically non-reduced, form it is used as the third-person singular subject-in-focus marker:

- (33) *mbí mà tr-á-k kà*
 ANAPH REL save-OBJ-1SG AFF
 ‘It is he who saved me!’ (*tàr* = to separate people who are fighting; save)
- (34) *tséy mbí déw kà bá-y ndá bàt màmbèŋ ábà cîŋ*
 so ANAPH sit like chief go take his mother ASSC his father
 ‘Then he became a chief, he came to take his mother and father’

The anaphor *mbí* in the object function can also have a proposition as its antecedent. In the following example, the anaphor has the form *mbí* rather than the reduced form *mbà*, as it occurs in the negative clause, which in Mina, as in a number of other Chadic languages, shares the same morpho-syntactic characteristics as the focus construction:

- (35) *ángà hìdà nd-á ngèn à sàɓ mbí sku*
 if man ‘from his birth’ 3SG know ANAPH NEG
 ‘If somebody says that since his birth he does not know that ...’

The form *mbí* also serves as an anaphor in prepositional phrases. Its antecedent, but not the one immediately preceding, can be a human noun, an event, or a proposition:

- (36) *dòk zá hà kúl kà dzàm ábà mbí skù*
 horse COMP 2SG able INF wrestle ASSC ANAPH NEG
 ‘The horse said, “You can’t fight with him.”’
- (37) *mə dāl-í dá kà mbí skù*
 REL do-STAT exist like ANAPH NEG
 ‘It is not done like that.’
- (38) *séy áb dùwàn mbí í*
 then ASSC back ANAPH 3PL
 ‘After that they ...’

The three examples above clearly point to the function of *mbí* as an independent third-person pronoun. Such a pronoun has cognates in other Chadic languages, e.g. in Mupun, a West Chadic language.

The focus function of the form *mbí* as well as its independent pronoun form, allows it to occur alone or in combination with other third-person pronouns and with relative markers. All three cases are illustrated by the following example:

- (39) *á gèr mbéŋ mə nzè mbí tátà ngàm á gèr mə mār*
 3SG want ANAPH REL be ANAPH alone because 3SG want REL control
ngùl á mbí tátà
 husband 3SG ANAPH alone
 ‘She wants to stay by herself, because she wants to control her husband herself.’

The idea of the switch-reference function, the focus function, and the independent pronoun function belonging to the same grammaticalization chain is quite plausible. The most likely path of grammaticalization is from independent third-person pronoun to contrastive focus marker to contrastive switch-reference marker. As a switch-reference marker and focus marker the third person anaphor *mbí* can also be used with the first-person pronoun. That form does not code switch reference with respect to the subject of the preceding clause, but rather contrastive focus with respect to another subject in the discourse:

- (40) *à zá ááá mbí sà nè kí yàn-á tàŋ àmmá sà bə*
 3SG COMP ah, ANAPH 1SG FUT INF move-DIST DEM but 1SG ASSC
idá
 house
 ‘She said, “I would have moved but I have a house.”’

2.4 Potential sources of the independent pronoun *mbí*

There is no obvious source for the marker *mbí*, in the sense that there is no independent lexical item having the same segmental and tonal structure. The closest lexical item is the form *mbù* whose primary meaning is ‘child’. Note, however, that the tone on *mbù* ‘child’ is low, whereas the tone on the switch reference and focus marker is

high. The evidence for *mbù* being the lexeme for child is provided by its use as either a subject or an object of a clause:

- (41) *bìyláv mǎ vl-á-k mbǎ táŋ kúl kǎ bǎt bǎ déwli*
 God REL give-OBJ-1SG child DEF can INF take ASSC force (Ful. dole)
skù
 NEG
 'It is God_i that gave me this child, he_j cannot take it away with force.'
- (42) *mbǎ à vl-á sǎŋ sǎ mǎ káp-ú*
 child 3SG say-DIST 1SG 1SG REL break-3SG
 'The child said, "It is me that broke it."'
- (43) *à-zá hidì wǎ á wǎk rà mbǎ bàhá à-zá mà-n dǎhá*
 3S.COMP MAN dem 3S go crazy HAB child also 3S.COMP mother exist
 'She said, "This man is crazy. The child also said, "I have a mother."'

The grammaticalization from 'child' to the independent pronoun is semantically plausible because in many Chadic languages the lexeme meaning 'child' often also designates any human being. The use of the term for human being as an anaphor for a [+human] noun is a common phenomenon across languages of the world:

- 1_14_1 <640 A> *I ^had a 'soldier* . [?]a I ^had a :c\orporal#
 1_14_1 <641 A> ^who "st\upid 'man got "t\ight# -
 1_14_1 <642 A> ^coming back from leave from [lǎ fr@ @] ^quite a
 _good [ko] !s\oldier {t\oo#}#.

The arguments against *mbí* deriving from *mbù* are phonological, the first one being the tone, low on *mbù* 'child', and high on the third-person independent pronoun; and the second, the vowel *i* on the pronoun and *u* on the word 'child'. Moreover, in Mina the lexeme for 'man' is the phonologically unrelated *hidì*. Thus, the possibility of a direct derivation from *mbù* to *mbí* has to be ruled out.

3. Motivation for switch-reference coding: The negative evidence

The existence of the category 'subject' in Mandarin has always been a highly controversial issue (cf. Chao 1968; Li & Thompson 1981). Judging from the existing literature, it appears that there are no formal properties that would delimit the category 'subject', and that there are no structures that crucially depend on that category. The aim of the present discussion is not to address directly the issue of the existence of subject in Mandarin; nevertheless, the coding of reference in Mandarin appears to provide an argument for those who claim that the category 'subject' does not exist in Mandarin. Recall that the initial hypothesis of the present study is that if a language codes the category 'subject' not only the role of an element as the subject must be transparent to the hearer but also the identity of the subject must be transparent. If Mandarin has the

category ‘subject’, which element is the subject in a clause must be transparent and also the identity of the subject must be transparent. Yet, neither of these conditions holds true for Mandarin:

- (44) *nèi běn shū chūbǎn le*
 that CL book publish PFV
 ‘That book, (someone) has published it’ (Li & Thompson 1981:88)

In a sequence of clauses neither coreference nor switch reference has to be coded. Compare the following fragment from natural discourse:

- (45) *wài-biānr fēi jìn-lái zhème dà gè ézi*
 out-side fly enter-com this big CL moth
 ‘There was a huge moth₁ that flew in from outside’
tā jiù cuān-shàng cuān-xià
 it then leap-up leap-down
 ‘It [family cat] immediately jumped up that high’
cuān nème-lǎo-gāo
 leap that-much-high
 ‘[it [the cat]] jumped up that high’
tā zhuō-zhù tā
 it catch-stop it
 ‘The cat caught the moth’
wár wár wár
 play play play
 ‘and [it] played and played [with it].’
wár fēi-le
 play fly-PFV
 ‘[The moth] was played [with] [and was caused by the cat] to fly away.’
 (Tao 1993:244)

It is not the case that the speakers of Mandarin cannot identify the participants in event. The translations of the preceding examples provide the necessary evidence for that. Tao 2001 suggests that the speakers of Mandarin identify the referents through analysis of other elements in discourse.

4. Conclusions

It has been shown that the coding of coreference and switch reference is a consequence of the language having grammaticalized the category ‘subject’. The implication of such grammaticalization is not only that the category ‘subject’ must be transparent to the addressee but also that the identity of the subject must be transparent. Languages may use various strategies to code coreference and switch reference. What strategy is used

depends very much on the coding means available. If a language has no agreement system and only nouns and pronouns as the coding means for reference, the pronoun is used to code coreference (if the subject has to be coded) and the noun codes switch reference. If a language has an extensive agreement system, agreement codes coreference and pronouns code switch reference. If a language does not have an agreement system but has two sets of pronouns, coreference is coded by one system of pronouns and switch reference is coded by another system. In all three types of languages, there exists the means of coding coreference through the absence of subject coding.

Note

* Although this paper treats a motivation for grammaticalization that has not figured in Heine's work, it is a tribute to Heine's contributions to the theory and methodology of grammaticalization research. I would like to thank the participants in the workshop, especially Salikoko Mufwene and Tom Givón, for helpful comments.

References

- Benis, H., Pica, P., & Rooryck, J. (Eds.). (1997). *Atomism and Binding*. Dordrecht: Foris.
- Broadwell, A. (1997). Binding theory and switch-reference. In H. Benis et al. (pp. 31–50).
- Chao, Yuen-ren (1968). *A Grammar of Spoken Chinese*. Berkeley, CA: University of California Press.
- Finer, D. (1985). The syntax of switch reference. *Linguistic Inquiry*, 16, 35–55.
- Frajzyngier, Z. (1987). Grammaticization through analysis. In S. DeLancey & R. Tomlin (Eds.), *Papers from the 2nd Pacific Linguistics Conference* (pp. 125–140).
- Frajzyngier, Z. (1997). Pronouns and agreement: Systems interaction in the coding of reference. In H. Benis et al. (pp. 115–140).
- Haiman, J. (1983). On some origins of switch-reference marking. In J. Haiman & P. Munro, 105–128.
- , & P. Munro (Eds.). (1983). *Switch-reference and Universal Grammar*. Amsterdam: John Benjamins.
- Kurcz, I., Lewicki, A., Sambor, J., Szafran, K., & Woronczak, J. (1990). *Słownik frekwencyjny polszczyzny współczesnej*. Kraków: Polska Akademia Nauk.
- Li, C. N., & S. A. Thompson (1981). *Mandarin Chinese. A functional reference grammar*. Berkeley: University of California Press.
- Matthews, P. (1997). *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- Mithun, M. (1999). *The Languages of Native America*. Cambridge: Cambridge University Press.
- Stirling, L. (1993). *Switch-reference and Discourse Representation*. Cambridge: Cambridge University Press.
- Tao, L. (1993). Zero Anaphora: Cognitive strategies in discourse processing. PhD. Dissertation, University of Colorado.
- Tao, L. (2001). Switch reference and zero anaphora: Emergent reference in discourse processing. In A. Cienki, B. J. Luka, & M. B. Smith (Eds.), *Conceptual and Discourse Factors in Linguistic Structure* (pp. 253–267). Stanford CA: CSLI.

Complex predicates based on generic auxiliaries as an areal feature in Northeast Africa^{*}

Tom Güldemann

Max-Planck Institute for Evolutionary Anthropology
and University of Leipzig

This paper explores an areal feature of unrelated languages and language families in Northeast Africa whereby a complex predicate is formed by a semantically generic auxiliary and a content sign of different lexical type. I diverge from previous studies of this phenomenon in arguing that the auxiliary must not be reduced in meaning to a speech verb. Instead of starting from a purported grammaticalization chain “*say* > AUXILIARY”, I start from the actual observed phenomena. I identify two central functions of the complex predicate which are in principle independent of speech verb semantics: first, if the content sign is itself a verb, the construction conveys a predicate that is focused or in some other way modified pragmatically; otherwise it enables a non-verbal content sign to be used as a predicate. On the basis of this approach, the areal isogloss can be defined more precisely. Finally, I give a preliminary and tentative diachronic sketch of how the feature arose in Northeast Africa and what implications this may have for areal-typological research in Africa as a whole.

1. The polyfunctionality pattern of an auxiliary

1.1 Introduction

The grammaticalization of an auxiliary verb, purportedly meaning ‘say’, to a predicative base of certain complex predicates in Northeast Africa has been repeatedly reported and discussed, for example, by Waley and Armbruster (1934), Armbruster (1960:29–32), Palmer (1974), Ferguson (1976:71–2, 75), Crass et al. (2001), and most recently and in more detail Cohen et al. (2002). In Güldemann (2001), I have surveyed reported-discourse and related constructions across Africa based on a sample of 39 genealogically widely distributed languages and could confirm the restricted areal distribution of this particular polyfunctionality pattern.

Before I start the discussion, there is one important preliminary remark in order. In contrast to all previous studies, I view the two uses of the same verb, namely in reported discourse and in predicate formation, not as two stages in a diachronic development, but rather as collateral functions which are the outcome of grammaticalizations that are in principle independent. Therefore, I do not use the term “grammaticalization chain”, but instead “polyfunctionality pattern”. The crux of the matter is the ultimate etymology of verbs used as the predicative nucleus of quotative constructions. It is not possible to lay out the quite extensive and complex argumentation regarding this problem, which can be found in Güldemann (2001). Here, it must suffice to say that, the few and historically motivated exceptions aside, the verbs concerned are etymologically not canonical speech verbs comparable to English *say*, French *dire*, German *sagen*, etc. They are rather verbal lexemes which originally encoded semantically more generic notions (e.g., manner deixis and similarity ‘like (this)’, action ‘do, make’, inchoativity ‘become’, and equation ‘be’) and which have been subject to PARALLEL grammaticalization in the domains of predicate formation AND reported discourse.¹ Hence, they will be labeled in the text and glossed in the examples more neutrally, namely as “auxiliary” (AUX) in the complex predicate to be discussed in this paper and as “quotative verb” (QV) in reported discourse.

The aim of this paper is threefold. First, I will try to show in the rest of Section 1 that the complex predicate typical for Northeast African languages comprises potentially two functional components, which are in principle independent from each other and thus have to be kept separate. Based on this observation and further exemplification of the predicate pattern across Northeast African languages, the areal isogloss will be defined more precisely and its cross-language variation be outlined in Section 2. The major conclusion there is that the common areal denominator is not the mere existence of the relevant predicate pattern, but rather a cluster of features that conspired in the emergence of a particular variant of such complex predicates. The polyfunctionality pattern outlined initially is only one aspect of the isogloss and has been focused on too much by previous researchers. The final Section 3 will be a short discussion of the historical-geographical setting in which the relevant isogloss evolved, including some possible implications for areal-typological research in Africa as a whole.

1.2 “Descriptive compounds” in Afar

I start with presenting the complex predicate with data from Afar (Cushitic, Afroasiatic stock). Here, it has been subject to analysis from various view points in previous works like Longacre (1990: 17–21), Hayward (1996: 540–4), and Cohen et al. (2002) and is called “compound verb” or “descriptive compound”. I refer the reader to these works for a more detailed description and confine myself to mentioning only the features important for the present discussion.

The complex predicates in Afar are based on two semantically generic verbs: (a) *hay* ‘do, make, put’ and (b) the morphologically highly irregular *iy* which is also salient as a quotative verb in reported-discourse constructions and thus translated

commonly as ‘say’. These two verbs allow a number of linguistic signs to be used in the mould of a predicative clause nucleus whereby *hay* establishes a transitive and *iy* an intransitive predicate. The following examples demonstrate the second auxiliary in construction with a noun and an onomatopoeic root. These convey the communicated state of affairs and precede the auxiliary.

- (1) a. *tàssa-iy*-
happiness-AUX-
‘become happy’
b. *kowkàw-iy*-
ON:chatter-AUX-
‘chatter’ (Hayward 1996:541)

As shown in (2), the auxiliary is not always adjacent and thus phonetically bound to the initial meaning-bearing element so that the traditional term “compound” is misleading.

- (2) *awki bađi tibbi edde iy-a*
boy son silence PART AUX:3M.S-IPERF
‘The small boy becomes quiet’ (Cohen et al. 2002)

In the above examples, the meaning-bearing unit itself is not inherently predicative. However, the two auxiliaries also combine with verb lexemes whereby these take on a special morphological form glossed here as C(ompound) F(orm).

- (3) *fak-ka hay-’t-e* vs. *fak-’t-e*
open-CF AUX-2S/3F.S-PERF open-2S/3F.S-PERF
‘You/she opened’
(4) *in’dix-xa in-’t-e*
say-CF AUX-2S/3F.S-PERF
‘You/she said’ (Longacre 1990:18–9)

A salient function of this subtype of complex predicate is to mark the “pivotal storyline” in narratives (Longacre 1990:18–9, based on Bliese 1981). In other discourse types, it is also used to convey interpersonal pragmatic functions like “astonishment, admiration, disapproval etc.” (Cohen et al. 2002) as well as attenuation/augmentation. Compare the following:

- (5) a. *kùdda-iy*-
run.away-AUX-
‘run away a bit, scamper away, run away quickly’
b. *usùlla-iy*-
laugh-AUX-
‘laugh a bit, laugh outright’
c. *cùlla-iy*-
enter-AUX-
‘pop in’ (Hayward 1996:542)

From the Afar data given in the literature and partially presented here, the following features can be summarized. They will be shown to be valid also for other languages to be discussed later on. The type of complex predicate at issue consists of two constituents: (a) a final, semantically generic verb (called “auxiliary” or “dummy verb”) and (b) an initial linguistic sign that encodes the state of affairs, but which is not or not fully integrated in the clause from a morphosyntactic viewpoint (called in the following “content sign”). These content signs are categorially diverse items, which can be classified as follows:

1. expressive one-word signs like onomatopoeics, ideophones, etc.
2. short quotes like exclamations, interjections, imperatives
3. loan words
4. lexical items used otherwise as nouns, adjectives, adverbs
5. lexical items used otherwise as verbs
6. lexical items not used outside a complex predicate

The auxiliary, as mentioned above, is often the same as a regular or even the default marker for the embedding of quotations. Although the existence of this collateral function is relevant for the following discussion, it is not decisive for the emergence of the complex predicate. This is already corroborated by the fact that Afar *hay* ‘do, make, put’, for which no role in reported discourse has been described, is also employed in the formation of complex predicates.

1.3 Two major functional domains of the auxiliary

There exists a wide functional variation in the use of auxiliaries in general and of auxiliaries which also function as quotative verbs in particular. Hence, it is necessary to differentiate the formal and functional details of language-specific cases in order to avoid a lumping of several, in principle distinct phenomena. Two major types of auxiliary use are relevant here and can cooccur in a language, as is the case in Afar. They are called for convenience the “focus-operator complex” and the “predicativization complex”.

In the first focus-operator complex, the auxiliary is used in an expression whose content sign is a verb and serves, as implied by the term, as a predication operator with an original function of foregrounding, focusing, attenuation, etc. This function can be relevant for the clause-internal information structure or relations across clause boundaries. Such a construction provides a paradigmatic choice in the expression of a state of affairs vis-à-vis other conjugational options and is in principle available for all verbs.

In the Bantu language Shona (Benue-Congo, Niger-Congo stock), the range of uses of the defective verb *ti*, which is derived etymologically from a manner deictic ‘like this’, represents a case where the polyfunctionality pattern consists of this and the mimesis-quotative function. This is discussed in detail by Güldemann (2002). Below, I present examples of *ti* as the default quotative verb (6), as an introducer of ideophones

(7), as the auxiliary in predicates with restrictive focus rendering ‘just, only’ (8), and as the auxiliary in a marked storyline form in narratives (9). The last two Shona constructions are functionally similar to the “descriptive compounds” of Afar exemplified in (3)–(5), which are also based on verbs.

- (6) *nda-ti uya neni*
 1s:PST-QV come:IMP COM:1s
 ‘I said, “Come with me!”’ (Hannan 1984:646)
- (7) *mu-komana aka-ti zii*
 1-boy 1:REM.PST-AUX ID:be.quiet
 ‘The boy lapsed into silence’ (Dale 1972:128)
- (8) *va-mwe va-no-ti ku-ngo-dzvuta vo-dhakwa*
 2-some 2-PRS-AUX INF-only-sip 2:INIT-get.drunk
 ‘Some people just take a sip and at once get drunk’ (Hannan 1984:646)
- (9) *shumba i-ka-ti zvino ya-va ku-da ku-va-dya*
 9.lion 9-SEQ-AUX now 9:PERF-become INF-want INF-2OBJ-eat
 ‘At this point the lion was about to eat them.’ (Dale 1972:133)

Comparable cases of such a polyfunctional auxiliary are found in Ik, Krongo, Lamang, and Nguni (see Güldemann 2001:395–404) as well as in Egyptian (Faulkner 1935; Depuydt 1989). The collateral function of a quotative verb as a predication operator has a parallel in the auxiliarization of generic ‘do’-verbs in other languages. Compare as just a few examples such auxiliaries as English *do*, German *tun*, and Hausa *yii* (Wolff 1993:505; Cohen et al. 2002). Since these verbs are not used in reported discourse, the co-function in quotations is unlikely to be a precondition for the development of the relevant type of complex predicate.

It is conceivable that such a periphrastic expression grammaticalizes further (e.g., the marked narrative to an unmarked narrative and then to a historical past). This means that such an auxiliary structure can encroach potentially on what is commonly viewed as the core inventory of inflectional verb grams called “tense-aspect-modality” (TAM).²

The second type of auxiliary use, which is exemplified for Afar with (1) and (2), is called here predicativization complex, because it concerns the conversion of content signs other than verbs into predicates. It provides a greater flexibility in the syntactic manipulation of different parts of speech. Again, the predicativization function is not tied to the quotative function, because it is widely attested with generic non-speech verbs which do not play a role in reported discourse. In addition to Afar and other Northeast African languages to be mentioned below, I cite here just two relevant cases: the predicativization function is reported for *mo* ‘do’ and *ke* ‘be(come)’ in the Papuan language Wambon (Vries & Vries 1992:13–7) and for *awun* ‘do, make’ and *xun* ‘be(come)’ in the Caucasian language Lezgian (Haspelmath 1993:178–83).

Later grammaticalization stages can be characterized by phonetic coalescence between the two parts and an increasing generalization of this predicate pattern across

the lexicon. Then, the auxiliary enters the morphological domain of transcategorial derivation and eventually the subclassification in the word category verb. In general, the predicativization complex concerns primarily the organization of the lexicon and its interaction with the morphosyntax of a language.

I would like to stress again that all three grammatical uses of a semantically generic auxiliary, that is, (a) as a quotative verb with reported discourse, (b) as a focusing predication operator, and (c) as a dummy verb for the conversion of certain lexemes into predicates, are attested separately and therefore are in principle independent from each other. I will try to show in the following that what characterizes Northeast Africa is a special language-internal combination of these functions accompanied by certain formal characteristics.

2. The linguistic nature of the Northeast African isogloss

2.1 Other affected languages and lineages

The following data will give a brief survey of the distribution of the complex-predicate type in Northeast Africa across linguistic lineages³ that are relevant and on which sufficient information is available.

Zayse is an example from Omotic (Afroasiatic stock).⁴ Hayward (1990:314) reports that the verb *y(d)*, which is most salient in reported discourse, is also encountered as the inflectional basis of “compound verbs”, in which its form is reduced towards an enclitic on the content sign. The following examples show it with verbs and non-verbs (in (10), *yd* turns up as *yya*).

- (10) *yedí-itta muutta-yya*
 come:IMP-and eat:IMP
 ‘Come and eat!’ (Hayward 1990:314)
- (11) a. *’ellé-ydi*
 fast-AUX:PERF
 ‘move/travel rapidly’
- b. *bo’é-ydi*
 bald-AUX:PERF
 ‘become bald’
- c. *’eró-ydi*
 ‘O.K.’-AUX:PERF
 ‘obey, agree’ (Hayward 1990:314)

Reinisch (1881–90,1:§126, 128–9) and Thompson (1989:306–7) describe a similar situation for the isolate language Kunama (Nilo-Saharan). The verb *u*⁵ functions as a quotative verb in reported discourse and also as the inflectional basis of an open class of so-called “abgeleitete Verben [derived verbs]”. These are formed from such types of content sign as ideophones, interjections, nominal lexemes, and loan words.

- (12) a. *óo-da*
ON-AUX:INF
'heulen [cry]'
- b. *bárē-da*
TWO-AUX:INF
'entzwei reissen [tear apart]'
- c. *katábō-da*
write-AUX:INF
'schreiben [write < Arabic]'

The isolate language Nera (Nilo-Saharan) has only scarcely been described by Reinisch (1874) and Thompson (1976), so that the situation regarding complex predicates is not entirely clear. It seems that there are two types with different auxiliaries, both of which follow the content sign in accordance with the general areal pattern.

One type of complex predicate is based on a verb *man* which is also used to signal reported discourse and therefore translated as 'say'. The only example found is with an ideophone.

- (13) *tim* *mass-o* [*mass* < *man-t*]
ID:be.quiet ?say:PST-3s
'He said *tim*, i.e. he was silent, kept mum' (Thompson 1989:306)

The information available does not allow one to get a conclusive picture for this structure. It seems that the range of content signs is limited (?only expressive items like ideophones) and the two constituents do not fuse into one word.

This is different with the second type of complex predicate based on the generic verb *ay* 'do, make' (cf. Thompson 1976:487–8, 490). The content signs comprise at least loan words as well as normal verb lexemes and the bipartite structure has coalesced to one word. The consonants *s* or *n* can intervene between the two constituents, apparently depending on the type of content sign.

- (14) a. *wal-ay-t-o*
speak-AUX-PST-3s
'He spoke'
- b. *darab-s-ay-t-o*
thirst(y)-?-AUX-PST-3s
'He was thirsty'
- c. *katab-n-ay-t-o*
write-?-AUX-PST-3s
'He wrote' [< Arabic] (Thompson 1976:487–8)

The complex predicates at issue have been reported from a number of Ethiosemitic languages (Afroasiatic stock).⁶ The present data come from Tigre based on Leslau (1945:25–6) and Raz (1983:66–7). Intransitive complex predicates are formed with the generic speech verb *bela* 'say' or, far less frequently, with the verb *wada* 'do, make';

transitive complexes mostly employ *'abala*, which is the causative counterpart of *bela*. The content signs belong to different lexical and morphosyntactic categories. Example (15) shows a short utterance reported in direct style:

- (15) *'abbe belä*
 'I.refuse' say:3s.PST
 'He refused' (Leslau 1945:26)

Complex predicates based on verb lexemes are also well attested and said to convey a "meaning [that] can be specified in terms of intensity or manner of the activity, such as: augmentative, attenuate or iterative" (Raz 1983:67). This parallels the situation in Afar.

- (16) *lafqa* vs. *lafəq 'abala* vs. *lafəq bela*
 'saw' 'saw a little' 'be sawn a little' (Raz 1983:67)

Also similar to Afar, auxiliary and content sign are not necessarily adjacent to each other as in (17) so that the philological term "compound" must not be taken literally.

- (17) *'əgəl la-šūm māk 'ət bəsōtū 'á-bala-jū*
 OBJ ART-chief slap in forehead:his CAUS-say:3s.PST-3s.OBJ
 'He hit the chief a slap in his forehead' (Sundström 1914:12)

Bedauye is chosen here as another representative of the Cushitic family (Afroasiatic stock). Roper (1929:84) reports that the verb *dī* 'say' serves as an auxiliary in conjunction with "indeclinable" words; at least some items are of an onomatopoeic nature like *tīm* 'be quiet', which is also found in other languages of the area (cf. (13) from Nera). Moreover, Reinisch (1893/4,4:§308) among others states that the class of so-called "schwache Verben [weak verbs]" characterized by a suffix conjugation is the modern reflex of a periphrastic construction based on an auxiliary verb *an*; its original meaning is said to have been 'be(come)'. As an independent verb, *an* is restricted today to signalling reported discourse. Since here it is (at least in the past) an alternative to *dī* 'say', it is also translated usually as 'say'.

- (18) *aní hadám-ani*
 1s destruction-AUX:1s.PRS
 'Ich zerstöre [I destroy]' (Reinisch 1893/4,4:17–8)

The range of content signs in this older type of complex predicate is similar to that found in other Cushitic and Ethiosemitic languages, including verbs. Unfortunately, no information is provided on the functional difference between inflecting one and the same verb according to the "strong" prefix conjugation or the "weak" suffix conjugation as exemplified in (19).

- (19) *an-dír* vs. *dír-ani*
 1s.PRS-kill kill-AUX:1s.PRS
 'I kill' (Reinisch 1893/4,4:18)

The second type of complex predicate in Bedauye is of wider importance, because it has been reconstructed for Cushitic as a family. Since Praetorius (inter alia 1894: 329–32), it is a widely accepted assumption that the inherited prefix conjugation has been replaced – in some languages more than in others – through the grammaticalization of this predicate pattern and its gradual generalization in the verbal lexicon leading to the new suffix conjugation.⁷

According to Armbruster (1960: §1792–6, 2760, 2874–9, 3603–60, 3910–3), the Nubian language Dongola (Nilo-Saharan) spoken along the Nile also has two verbs which are used simultaneously as an auxiliary in complex predicates and as a quotative verb. Intransitive inchoative verbs meaning ‘become X’ are formed from the stem *án* as in (20):

- (20) a. *míll-an-*
bad-AUX-
‘go bad, deteriorate’
b. *núgud-an-*
slave-AUX-
‘become a slave’
c. *úg-an-*
night-AUX-
‘become night’ (Armbruster 1960: §3913)

The other group of complex predicates based on the cognate of the quotative verb *é* and exemplified in (21) do not display such a restricted meaning.

- (21) a. *hamd-é-*
praise-AUX-
‘praise’
b. *ǵagād-ε-* or *ǵagād-é-*
weak-AUX-
‘melt, become pliant’ (Armbruster 1960: §2874, 2877, 3615)

While both auxiliaries seem to be bound phonetically to the content sign, *é* maintains its own stress in some cases, especially with loan words as shown in (21b).

Nougayrol (1989: 34–FN1, 141–2) reports for the Maban language Aiki (Nilo-Saharan) that a number of predicative expressions are conveyed by complexes that consist of an initial content sign and a final auxiliary taking on all verbal inflection. Data given by Edgar (1989: 27–8, 43–4, 80; 1991: 342–3) suggest that other Maban languages also possess this construction type. In Aiki, there are two such auxiliary verb stems, namely *ir*, which is cognate with the default verb for marking reported discourse, and the less frequent *εs* ‘do’. The following example shows the first auxiliary in construction with an Arabic loan.

- (22) *áwán mb-r-è*
 help 2s:1s-AUX:IPFV-DECL
 ‘Je vais t’aider’ [< Arabic ‘help’] (Nougayrol 1989:99)

Finally, the entire Saharan family can be mentioned as displaying such complex predicates from a diachronic point of view. The most extensively described language is Kanuri.⁸ Hutchison (1981a:90–1, 95–111) among others reports that the cognate of the quotative verb *n*⁹ is the inflectional basis of the largest and only productive of its three verb classes. He (1981b:225, 230) states on the history of this conjugation class:

Verb class 2 is virtually unlimited in size since any lexical item of Kanuri or of a contact language can today be inflected by the class 1 verb *ngin* [= auxiliary *n*] and function as a verb, e.g. in response to a modern need. Class 2 is thus layered, since its members come from a great many sources historically.

... given: (1) the independent synchronic meanings of *ngin* i.e. ‘say, think’. (2) the synchronic independent use of *ngin* to describe ideophonic processes, (3) the phonological similarities between ideophones and class 2 verb roots, I propose that the original members of verb class 2 were ideophones, which through constant usage became fused to the verb *ngin* from their position as preceding juxtaposed objects of *ngin*.

Interestingly, there still exists the synchronic alternation between a normal ideophone construction based on the free form of the auxiliary *n* and the verbal inflection of the ideophonic root by means of its grammaticalized suffixal counterpart.

- (23) a. *kàrэгà-nzè bádák (bádák) shìn*
 heart-3s.POSS ID:heartbeat AUX:3s.NPST
 ‘His heart is going badak badak’
 b. ... (*bàdàk*) *bádák-cìn*
 beat.of.heart-3s.NPST
 ‘His heart is beating (loudly)’
 [*cìn* is the bound, inflectional counterpart of the verb *shìn*] (Hutchison 1981b:229)

2.2 Defining the isogloss

Given that this particular type of complex predicate has been repeatedly associated with Northeast Africa, it is important to recognize that partially similar linguistic phenomena are found also in other regions of this continent (and the world, for that matter). Taking a brief look at other African cases is useful for defining more precisely the areal isogloss at issue.

One case has already been discussed briefly in Section 1.3 with Shona, which is representative for many other Bantu languages. The auxiliary *tí* is simultaneously the default verb in reported discourse, the support verb when ideophones are to be used

predicatively, and an auxiliary in constructions with inherently predicative verb lexemes. There exist, however, considerable differences to the Northeast African cases described in the previous sections. The Shona auxiliary (a) does not predicativize content signs other than ideophones, (b) occurs before this constituent, and (c) retains a status as an independent phonological word.

Birom, a Plateau language (Benue-Congo, Niger-Congo stock), comes a bit closer to the Northeast African phenomenon. In (24), I exemplify a type of compound verb consisting of a content sign and a stem *ye* which is independently attested only as a quotative verb. The linkage between the two elements is apparently tighter. Nevertheless, the content signs are virtually restricted to ideophones and similar items and they follow the auxiliary.

- (24) a. *ye-tʃlót*
 AUX-ID:slide
 ‘glisser’
 b. *ye-tyɔp*
 AUX-ID:fall.in.water
 ‘plonger’
 c. *ye-cì.syáw*
 AUX-ON:atchoum
 ‘éternuer’ (Bouquiaux 1970: 225)

Fur (Nilo-Saharan), a language from the geographical area in question, also displays a related phenomenon, but would traditionally not be included in the core group of languages outlined in 2.1. Compare the following statement by Beaton (1968: 123):

A very noticeable trait in modern Fur is the use of Fur versions of Arabic words compounded with the factitive verb *’a* – to do, to make ... Such forms are numberless and are frequently used by the more sophisticated, even when there is a perfectly good Fur verb at the speaker’s disposal.

The Fur predicate based on the auxiliary *’a* seems to differ from comparable types in the larger area at least in the following respects: (a) it is constructed with a rather limited range of content signs, namely loans, and (b) the verb does not play a role in the marking of reported discourse, which sets it off from most of the other languages cited here.

It becomes clear from the above information that the Northeast African isogloss, when conceived off as an EXCLUSIVE characteristic of this area, must be described as a cluster of features cooccurring in the grammar of a language. These features are as follows:

1. a semantically generic verb serves as the inflectional basis of complex predicates
2. the respective verb is usually also used in reported discourse
3. the range of content signs converted into predicates/verbs is fairly wide
4. the auxiliary occurs after the content sign
5. the complex predicates tend to merge to a one-word sign

2.3 The variation of the isogloss across the area and a historical scenario

I have presented data on a complex-predicate strategy in genealogically widely distributed languages of Northeast Africa and contrasted it with a few partly similar phenomena in other languages of the continent. I will now show briefly how this apparently shared feature varies across the relevant group of languages on several parameters. At least the following features have been found important in this respect and are presented in Table 1:

1. Choice of auxiliary:
 - a. use of more than one auxiliary
 - b. also use of an auxiliary without a grammatical role in reported discourse
 - c. auxiliary correlates with transitivity of complex predicates
2. Range of content signs (see Section 1.2)
3. Degree of coalescence in complex predicates
4. Importance of complex predicates for inventory of predicatively expressed concepts

It must be noted that this comparison can only provide a preliminary picture. The data on some languages do not become entirely clear and/or are not presented completely in the sources used here. Therefore, some table values, especially those for the last two scalar parameters in the line 3 and 4, could only be filled in on the basis of impressionistic judgements on the available information. Nevertheless, the data are sufficient to give an idea about certain differences of the linguistic phenomenon across the area under consideration and to throw light on the historical scenario of how it emerged as a whole from a diachronic perspective. Such a historical scenario is laid

Table 1. Variation of complex predicates in Northeast African languages

Language (Family)	Zayse (Omotic)	Kunama (Isolate)	Nera (Isolate)	Tigre (Semitic)	Bedauye (Cushitic)	Dongola (Nubian)	Aiki (Maban)	Kanuri (Saharan)
1a More than one AUX	No	No	Yes	Yes	Yes	Yes	Yes	No
1b Non-quotative AUX			Yes	Yes	No	No	Yes	
1c Transitivity correlation			No	Yes	No	Yes	No	
2 Short quotes	Yes	Yes	?	Yes	?	Yes	?	Yes
Ideophones etc.	?	Yes	Yes	Yes	Yes	Yes	(Yes)	Yes
Loan words	?	Yes	Yes	?	?	Yes	Yes	Yes
Non-verb lexemes	Yes	Yes	?	Yes	Yes	Yes	Yes	Yes
Verbs	Yes	?	Yes	Yes	Yes	(Yes)	No	No
Opaque items	?	?	?	Yes	?	?	Yes	Yes
3 Coalescence	Inter	High	Inter	Low	High	Inter	Low	High
4 Predicative inventory	?	High	?	?	High	Inter	Inter	High

Note: Blank = not applicable, ? = insufficient information, (...) = less important, Inter = intermediate

out in more detail in Güldemann (2001, Section 4.4.1.3); here, I will only present a brief sketch of the assumed major developments.

STEP 1 in this scenario can be characterized by a situation where a semantically generic verb serves as a support verb providing a construction for the predicative use of pragmatically expressive signs like ideophones, exclamations, etc. As far as this is applicable in a given language, this presumably includes right from the beginning also loan words, because these lexical items, too, are marked not only in terms of morphosyntax, but also regarding their pragmatic import (see, e.g., the above quote by Beaton on Fur). At this point, the language has acquired an additional clause-formation type which is productive to the extent to which it employs pragmatically expressive meaning-bearing units as predicates. This situation is found in Fur and may also hold for the constructions with *man* in Nera and possibly *di* in Bedauye.

It is possible, but not necessary that the verb displays additional uses as a foregrounding predication operator and/or in quotations (cf. the ‘do’-verbs of Afar, Nera, Tigre, and Aiki, which are not used in reported discourse). It is, however, relevant for the general scenario in Northeast Africa that such collateral functions can potentially feed and reinforce the complex-predicate construction. For example, it is attested crosslinguistically that one-word quotes accompanied by a quotative construction can substitute for verb lexemes in the expression of certain concepts of locution. Compare an example from Mupun (Chadic, Afroasiatic stock), where (*sat*) *nə i* ‘say “Yes”’ is semantically equivalent to a normal verb ‘agree’.¹⁰

- (25) *wur sat nə i* or *wu nə i*
 3M.S say Q ‘Yes’ 3M.S Q ‘Yes’
 ‘He said yes, he agreed’ (Frajzyngier 1996: 125)

In languages taking STEP 2, this predicate pattern is increasingly transferred to other non-expressive types of content signs like ordinary adverbs, adjectives, nouns, etc. From this point on, the strategy has acquired the exclusively structural aspect of predicativization of lexemes other than verbs. The situation in Aiki seems to be on the threshold of such a stage where the verb can be called “dummy predicative”. In languages like Afar, Zayse, Tigre, etc., where the grammaticalizing verb combines also with verb lexemes and fulfills certain predication-operator functions, it is labeled for convenience “multipurpose auxiliary”.

In STEP 3, which may in principle set in simultaneously with the earlier steps, the language shows a formal adaptation of the complex predicates to normally inflected verbal predicates thereby acquiring in the long run a morphological opposition to other “canonically” conjugated verbs. In the relevant verb-final languages of Northeast Africa, this pronounced process is characterized by coalescence in the complex predicate whereby the grammaticalizing auxiliary progresses along a cline [free word > enclitic > suffix]. Given that there is a universal preference for suffixation (see, e.g., Hawkins & Cutler 1988; Hawkins & Gilligan 1988), it is not far fetched to assume that

the complex predicates' constituent order [content sign – auxiliary] has fostered the coalescence process.

STEP 4 can be identified as the systemic GENERALIZATION of this new inflection pattern across the verbal lexicon, inter alia because it is regular and phonetically transparent. The earlier syntactic pattern has now come to define a productive inflection class of verbs in opposition to a closed class of verbs that follow a conservative, morphologically irregular conjugation type. This is a typical case of “host-class formation” in grammaticalization, that is, the emergence of a lexical sub-category sharing the same inflectional morphology (see, e.g., Himmelmann (1992:17–8, 21–2) under his older term *carrier class-formation*). The function of the earlier auxiliary will be called here “conjugation type marker” and can be identified in the respective verb classes of Saharan, Cushitic, and – possibly in a less advanced stage – also in Kunama, Nera, and Dongola.

It becomes clear in some languages that another phenomenon relevant in the geographical area can contribute to a considerable extent to this last development, namely the use of the respective auxiliary as a predication operator. I refer to the possibility that the functional opposition between the unmarked simple predicate and the marked complex predicate can be given up over time whereby the latter structure is maintained and becomes functionally unmarked. These forms will increase the size of the new verb class.

A major reason for this direction of generalization has already been indicated above and also focused on as early as Reinisch (1893/4, 4:§308) who states regarding Bedauye:

Hiernach wäre wol zu vermuten, dass sämmtliche verba nach belieben stark oder schwach flektirt werden könnten. Einst dürfte das auch wirklich der fall gewesen sein, weil noch gegenwärtig eine erkleckliche anzahl von verben sowol stark wie schwach flektirt wird. Tatsächlich aber hat die überwiegende merzal der verba die starke flexion bereits eingebüsst und wird nur mer schwach flektirt. ... Der grund hiervon ist wie in den germanischen sprachen gewiss nur in der einfachern bildungsweise der schwachen verba zu suchen. [So it may be assumed that all verbs can be freely inflected either as strong or as weak verbs. This might well have been the case in the past since a considerable number of verbs are still inflected today both as strong and as weak verbs. In fact, however, the great majority of verbs have lost the strong conjugation type and are only inflected as weak verbs. ... The reason for this is surely to be found, as in the Germanic languages, in the simpler formation type of the weak verbs]

While outlining the scenario, I have already assigned the Northeast African languages treated above to a position on this historical cline. As a summary, I repeat this in Figure 1.

The auxiliary's function of syntactic conversion or transcategorical derivation of heterogeneous content signs, labelled in Section 1.3 the predication complex, is central to the above scenario, because it is relevant in all its different stages. It is im-

Predicative for expressive signs	Dummy predicative or multipurpose auxiliary	Conjugation- type marker	
?	Nera <i>man</i>	Afar, ?Bedauye (+ other Cushitic)*	Kunama Bedauye <i>an</i> (= Cushitic)
Fur	Tigre (+ other Ethiosemitic)*	Nera <i>ay</i>	Kanuri (= Saharan)
	Zayse (?+ other Omotic)*	Dongola	
	Aiki (?+ other Maban)		

Note: * verb is also predication operator, hence multipurpose auxiliary

Figure 1. Degree of grammaticalization of complex predicates in Northeast Africa

portant to note at this point that this function unites all relevant Northeast African languages. This is not the case with the other two functions: the quotative use is missing with Fur *'a* as well as with certain auxiliaries used parallel to a quotative verb (e.g., in Afar, Nera, Tigre, and Aiki); the so-called predication-operator complex fails to apply to most languages outside the Ethiopian Plateau as it is securely attested only in Cushitic, Ethiosemitic, and parts of Omotic. In view of this situation, but also considering that these collateral uses can certainly reinforce the function in the predicativization complex, it remains to be determined in the future to what extent this pool of conspiring functions of the relevant verb is responsible for bringing the developmental chain to completion, namely to the stage of a conjugation type marker.

Regarding the verb's function in reported discourse, however, one qualification can already be made. Although I have included it as a feature of the Northeast African isogloss, its salience here is a minor detail from a crosslinguistic perspective. There are several areas in the world with similar complex predicates and some where their historical development can be described by a scenario comparable to that outlined above. One such case is northwestern Australia (see, e.g., Heath (1976), Alpher (1994: 164–8), Schultze-Berndt (2000: 532–43), and McGregor (2001, 2002)). In this area, verbs employed as the default element in the embedding of reported-discourse ARE involved as auxiliaries in the complex-predicate formation; however, their role for the strategy as such is far less important. The major reason for this is that the auxiliaries in the Australian languages are mostly not just one or two in number, but come instead from a greater, although closed set of “semantically generic verbs [which] participate in an overt system of event categorisation” (Schultze-Berndt 2000: 552). This means that any reported-discourse verb there is just one among other lexemes.

3. Historical-geographical aspects of the isogloss

The functional motivation of the above scenario should not be divorced from another important aspect of the Northeast African phenomenon, namely the concrete historical-geographical circumstances of the emergence of the isogloss in this area.

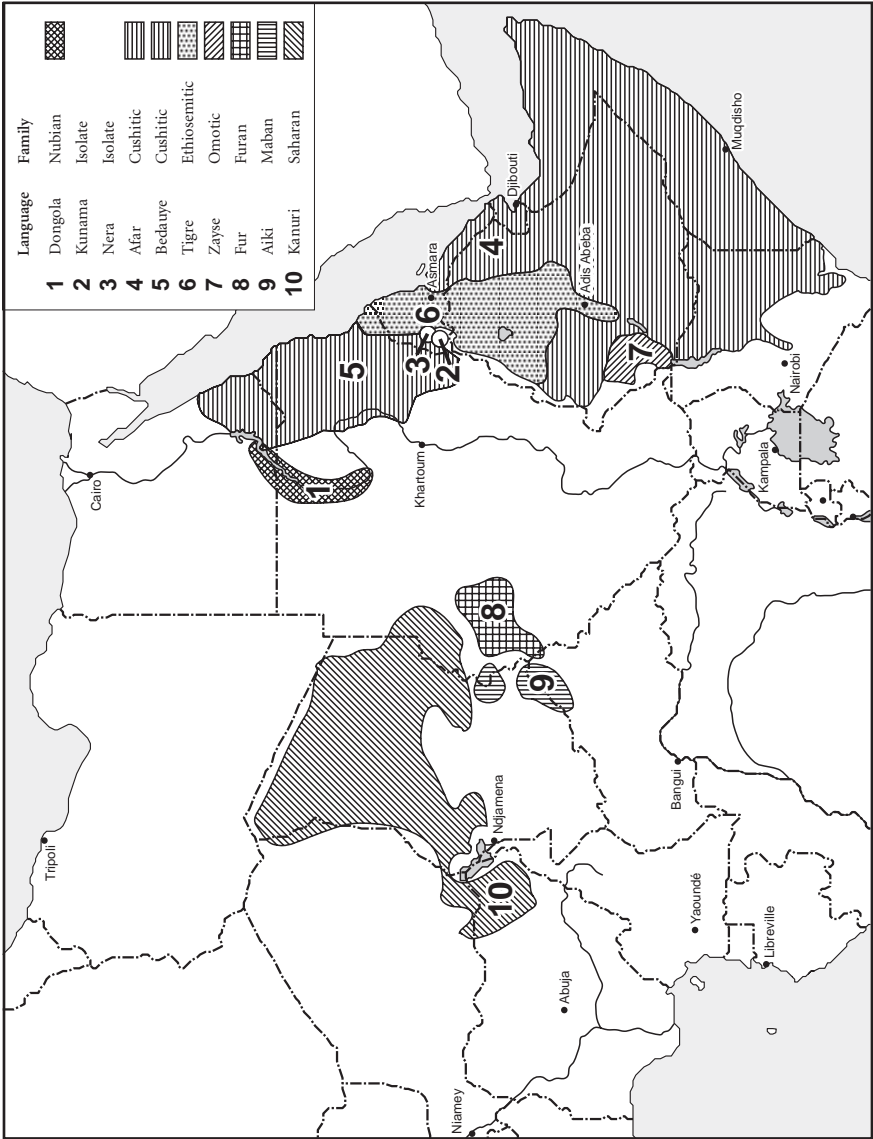
This will be discussed briefly in the following on the basis of its synchronic distribution given in Map 1.

The picture resulting from mapping only the individual languages treated in the paper could give the impression that the complex-predicate type has a rather scattered occurrence across a large area. However, this feature is repeatedly relevant for entire families like Cushitic, Ethiosemitic, Saharan, and Maban or at least for areal subgroups of families like Nubian and Omotic. Moreover, some languages chosen in this paper as representatives for a given lineage are located in the geographical peripheries of the area under consideration; the choice of another language of the family with the same characteristics would have provided a closer geographical link to other relevant and genealogically unrelated languages. This holds for Kanuri within Saharan and Aiki within Maban. Taken these considerations into account, it turns out that the synchronic distribution of the feature builds up two fairly large and compact zones: (a) the Sahel and Sahara regions around and east of Lake Chad with Saharan, Maban, and incipiently Furan languages and (b) a region comprising the Horn of Africa stretching north up to the Nile and the Nubian Desert with Nile Nubian, Cushitic, Ethiosemitic, and Omotic languages as well as the isolates Kunama and Nera.

Before discussing this general areal pattern, a note of caution is in order. There already exist gaps in the data for individual languages and lineages treated above. However, the situation is much more serious for other languages and language groups which are situated in or are at least adjacent to the relevant area with the effect that these could not be included in the survey at all. This concerns quite a few lineages which are distributed from eastern Chad, throughout Sudan, to western Ethiopia and Eritrea and are commonly assigned to Nilo-Saharan. With the availability of more extensive and reliable data on these languages, it is quite possible that the isogloss will turn out to affect even more language groups in Northeast Africa and the hypotheses to follow will have to be adapted.

A first conclusion from this geographical picture is that the isogloss is not confined to the Ethiopian Plateau; the highest genealogical diversity of the relevant languages is in fact found northeast of it (see below). For this reason, it is not a particularly suitable criterion for defining this linguistic area – pace Ferguson (1976) who lists it as an Ethiopian feature without referring to its far wider distribution. This observation somewhat echoes the more general reservations raised by Tosco (2000) against the integrity of this area. The data of this paper suggest, however, another feature in this domain possibly turning out to be shared exclusively by Ethiopian languages, namely the function of the auxiliary as a focusing predication operator.

The two geographical areas just outlined are separated by a large wedge situated first of all in modern-day Sudan. Apparently, this zone has been created by the relatively recent spread of Arabic. This begs the question about the areal picture before this expansion. It is not far fetched to assume that the languages that became submerged in this region by Arabic were areally and at least partly also genealogically related to languages in the East and West which possess the feature at issue. So there is good reason for the hypothesis that the languages and lineages affected by the isogloss formed



Map 2. The distribution of complex predicates in Northeast Africa

in the past a fairly coherent belt stretching from the Horn of Africa across most of Ethiopia, Eritrea, northern Sudan, and northern Chad up to the Lake-Chad Basin. The western and eastern peripheries of this large area are quite homogeneous in terms of genealogical classification in that they are constituted by Saharan and Cushitic, respectively. As opposed to this, the areal center around and east of the Nile is heterogeneous in displaying five lineages which are genealogically distant from each other: Nubian, Kunama, Nera, Cushitic, and Ethiosemitic.

There is ample evidence apart from purely geographical considerations that the feature is historically deeply entrenched in the area. First, the language-specific strategies often display a high degree of grammaticalization. This can be discerned from the fact that (a) the earlier complex predicates have become one-word signs and have been generalized across the verbal lexicon in lineages like Saharan and Cushitic and (b) the respective auxiliaries are mostly opaque etymologically, display a highly irregular conjugation involving stem suppletion, and have very little phonetic substance. Second, the complex-predicate strategy has entered at least in Cushitic a new cycle in that another lexeme is employed both as a reported-discourse verb and as an auxiliary in a renovated and productive formation pattern of complex predicates.¹¹

The clear historical layering of the isogloss in Northeast Africa can also account for a potential problem of my general scenario. I have claimed above that the relevant auxiliaries are originally not canonical 'say'-verbs. However, this is contradicted at least by Ethiosemitic and seemingly also by modern Cushitic. For the first family, the majority of language-specific verbs are clearly cognate with a root **bhl* which Leslau (1991:89) can trace back as a speech verb 'say' to Afroasiatic in general. For the second family, there are good chances to reconstruct a stem **d'h* (Black 1974:302; Sasse 1979:61) with an original meaning 'say' and this is recruited in some modern languages for the formation of complex predicates. A solution to this apparent problem is the assumption that the synchronically attested use of these speech verbs has been influenced by the existence of the relevant polyfunctionality pattern either in the previous language state, as can be assumed for Cushitic, or in a substrate, as can be assumed for Ethiosemitic which has been influenced strongly by Cushitic (see inter alia Hetzron (1975:113) regarding the feature at issue). In other words, it is hypothesized that the synchronic range of usage of Ethiosemitic **bhl* and Cushitic **d'h* is not the result of canonical grammaticalization from 'say' to a multipurpose auxiliary. It rather reflects the calquing of an established polyfunctionality pattern of an older, more grammaticalized non-speech verb onto a younger speech verb, because the latter was linked to the former by its use in reported discourse. This is a non-canonical process of context extension which does not contradict an idealized path of grammaticalization that is exclusively steered by functional principles (see Heath (1998) and Güldemann (2001, Sections 4.4.1.3 and 4.5) for more discussion of this type of process and its relevance for the Northeast African isogloss).

The last remarks indicate that an attempt to trace the complex-predicate strategy as far as possible to its origin in a smaller area or group of languages and to reconstruct at least partially its history across time, space, and different linguistic lineages

is a worthwhile undertaking for future research. This is so because several (groups of) languages or geographical areas where the feature is found today are unlikely candidates for its origin. In some languages, its existence is best explained as mediated historically through structural interference from contact languages where the feature is old. This holds with all probability for Ethiosemitic; more data are needed to determine whether, for example, Nile Nubian and the affected Omotic languages also belong here. Similarly, the distribution of the feature in the western and southeastern areas of the isogloss can be motivated by an outward spread of Saharan and Cushitic, respectively. These are both families where the feature is deeply entrenched historically and can be traced back to an early language state. My hunch at the present is that the feature started out from a region in the center of the larger area close to the Nile; Cushitic and Saharan seem to have played a central role in its further geographical expansion. Given that a number of languages possibly involved are not sufficiently described and the area in question can not yet be surveyed exhaustively for the feature, it is clear that my observations here are at best preliminary and far from giving any conclusive answer to this complex question.

This line of research will be intimately related to a final consideration. The identification of this Northeast African isogloss is likely to be significant for research on areal typology in Africa as a whole, in particular for the future task of cutting up the continent into larger areas of linguistic convergence and diffusion of individual features. One result of the present investigation is that the identified isogloss does not particularly strengthen the idea that the Ethiopian Plateau is a longstanding linguistic area that can be clearly delimited geographically from its neighboring areas. Since the Ethiopian Plateau is one of the fairly few regions in Africa where an areal-typological approach has been entertained previously, the present observation is not a trivial one; it indicates how much there is to catch up in the way of this kind of research on this continent.

Moreover, taking the particular distribution and great time depth of the feature into account, the question arises whether it is a reflex of a larger linguistic macro-area in Northeast Africa which only includes the Ethiopian Plateau as a subpart. In the necessary search for other linguistic features coinciding geographically with this isogloss, there comes at least one other clear and well-known candidate to mind. The distribution area of complex predicates is largely identical with the most important and compact African concentration of languages and lineages with verb-final or generally head-final constituent order (see Heine (1976) under his "type D").¹² Since word order has been shown to be more sensitive to language contact than other linguistic features (see, e.g., Nichols 1992), this geographical parallel could certainly be accommodated in the areal-typological approach pursued here. In fact, the feature complex of constituent order alone has already been entertained by Heine (1975: 41–2) for proposing a convergence area "Chad-Ethiopia"; this would be consolidated by the distribution of complex predicates.

More research on this topic is needed, however, because other features that have been claimed to be of areal significance in this part of Africa do not obviously pattern

in the same way. This holds, for example, for Schadeberg's (1987) survey of two assumed phonological isoglosses, namely (a) the lack of a voiceless-voiced distinction in plosives or at least a tendency towards its neutralization and (b) the existence of five places of articulation with two distinctive segment types in the dental-postalveolar region. According to Schadeberg, these features exclude the Horn of Africa, but include languages in the Rift valley as far south as Uganda and Kenya.

Thus, the question of whether the similar distribution of head-final languages on the one hand and of languages with complex predicates on the other hand reflects indeed an old linguistic macro-area on the African continent has yet to be answered conclusively.

Notes

* The research on the present subject was made possible by a generous scholarship of the "Volkswagen-Stiftung" for which I am grateful. My thanks go also to the participants of the conference for an interesting discussion about this topic as well as to Stefan Elders, Orin Gensler, Maarten Mous, Eva Schultze-Berndt, Martine Vanhove, and Ekkehard Wolff for valuable comments on earlier drafts of this paper. The abbreviations in example glosses are as follows: ART article, AUX auxiliary, CAUS causative, CF compound form, DECL declarative, F feminine, ID ideophone, IMP imperative, INF infinitive, INIT initiative, IPERF imperfect, IPFV imperfective, M masculine, NPST non-past, OBJ object, ON onomatopoeic, P plural, PART particle, PERF perfect, POL polite, POSS possessor, PRS present, PST past, Q quotative, QV quotative verb, REM remote, S singular, SEQ sequential. Bare Arabic numbers indicate noun class indices, otherwise person categories.

1. Such cases have been called "polygrammaticalization" in previous studies (cf. Craig 1991).
2. Since the label TAM is often used in a functionally very loose way, I will continue to employ the more inclusive term PREDICATION-OPERATOR functions.
3. The term "lineage" is used here for a genealogically defined language group irrespective of its age (cf. Nichols 1992:25). Regarding language classification, I do not start from Greenberg's (1963) assumption of just four genealogical supergroups in Africa. Instead, I recognize only those units which have been established or are likely to become established in the foreseeable future by the historical-comparative method. For the present discussion, this is relevant for languages subsumed under Nilo-Saharan: they are not claimed here to be all related genealogically.
4. See Azeb Amha (2001) for similar data on Wolaitta.
5. The verb has a very irregular conjugation involving stem suppletion. Several inflected forms regularly display a vowel *u* as the only phonetic substance of the verb stem, which motivates the present citation form. Other sources give a different citation form based on the suppletive stem *da*, which is found in the imperative and infinitive.
6. Apart from grammatical descriptions of individual languages, additional examples can be found in Crass et al. (2001), who discuss Amharic and Zay, and in the survey by Cohen et al. (2002).
7. Banti (2001), however, challenges Praetorius' reconstruction and tries to show that the suffix conjugation can be explained without the help of this grammaticalization scenario.

8. See, inter alia, Cyffer (1991) and Crass et al. (2001) for data on other Saharan languages.
9. The normal citation form for this verb in Kanuri philology is the 1st-person non-past *ngin*.
10. Cf. also Benveniste's (1971) discussion of so-called "delocutive verbs" in Indo-European like Latin *salutare* from *Salus!*, German *bejahren* from *Ja!*, etc.
11. The historical-geographical pattern of the isogloss in general and the situation in Cushitic in particular do not support Banti's general rejection of Praetorius' idea on reconstructing the relevant suffix conjugation, but rather corroborate this traditional view.
12. There are only four isolated cases of verb-final languages outside this Northeast African region, namely the families Dogon and Ijoid in western Africa, the isolate Sandawe in eastern Africa, and the Khoe-Kwadi group in southern Africa.

References

- Alpher, B. (1994). Yir-Yoront ideophones. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound Symbolism* (pp. 161–177). Cambridge: Cambridge University Press.
- Armbruster, C. H. (1960). *Dongolese Nubian: A grammar*. Cambridge: Cambridge University Press.
- Azeb Amha (2001). Ideophones and compound verbs in Wolaitta. In F. K. E. Voeltz & C. Kilian-Hatz (Eds.), 49–62.
- Banti, G. (2001). New perspectives on the Cushitic verbal system. Paper presented at the 32nd Annual Conference on African Linguistics, Berkeley, March 23–25, 2001.
- Beaton, A. C. (1968). *A Grammar of the Fur Language*. Khartoum: Sudan Research Unit, Faculty of Arts, University of Khartoum [Linguistic Monographs 1].
- Benveniste, E. (1971). Delocutive verbs. In E. Benveniste (Ed.), *Problems in General Linguistics* (pp. 239–246). Coral Gables: University of Miami Press [Miami Linguistics Series 8].
- Black, P. D. (1974). Lowland East Cushitic: Subgrouping and reconstruction. PhD. Dissertation, Yale University.
- Bliese, L. F. (1981). *A Generative Grammar of Afar*. Arlington, TX: University of Texas [Summer Institute of Linguistics Publications in Linguistics 65].
- Bouquiaux, L. (1970). *La langue birom (Nigeria septentrional): Phonologie, morphologie, syntaxe*. Paris: Les Belles Lettres [Bibliothèque de la Faculté de Philosophie et Lettres de l'Université de Liège 185].
- Cohen, D., M.-C. Simeone-Senelle, & M. Vanhove (2002). The grammaticalization of 'say' and 'do': An areal phenomenon in East Africa. In T. Güldemann & M. von Roncador (Eds.), pp. 227–251.
- Craig, C. S. (1991). Ways to go in Rama: A case study in polygrammaticalization. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*, Vol. 2 (pp. 455–492). Amsterdam: John Benjamins [Typological Studies in Language 19:2].
- Crass, J., Dehnhard, B., Meyer, R., & Wetter, A. (2001). Von 'sagen' zum Verbbildungsmorphem: Die Grammatikalisierung des Verbs 'sagen' einmal anders. *Afrikanistische Arbeitspapiere*, 65, 129–141.
- Cyffer, N. (1991). The Zaghawa verb structure and its relation to other Saharan languages. In M. L. Bender (Ed.), *Proceedings of the 4. Nilo-Saharan Conference: Bayreuth, Aug.30–Sep.2, 1989* (pp. 79–90). Hamburg: Helmut Buske [Nilo-Saharan Linguistic Analyses and Documentation 7].

- Dale, D. (1972). *Shona Companion*. Gwelo, Zimbabwe: Mambo Press.
- Depuydt, L. (1989). The contingent tenses of Egyptian. *Orientalia, Nova Series*, 58, 1–27. Roma: Pontificum Institutum Biblicum.
- Edgar, J. (1989). *A Masalit Grammar with Notes on Other Languages of Darfur and Wadai* Berlin: Dietrich Reimer [Sprache und Oralität in Afrika 3].
- . (1991). *Maba-group Lexicon* Berlin: Dietrich Reimer [Sprache und Oralität in Afrika 13].
- Faulkner, R. O. (1935). The verb 'i 'to say' and its developments. *Journal of Egyptian Archaeology*, 21, 177–190.
- Ferguson, C. A. (1976). The Ethiopian language area. In M. L. Bender et al. (Eds.), *Language in Ethiopia* (pp. 63–76). London: Oxford University Press.
- Frajzyngier, Z. (1996). *Grammaticalization of the Complex Sentence: A case study in Chadic* Amsterdam: John Benjamins [Studies in Language Companion Series 32].
- Greenberg, J. H. (1963). *The Languages of Africa*. Bloomington, IN: University of Indiana [Research Center in Anthropology, Folklore, and Linguistics Publications 25].
- Güldemann, T. (2001). Quotative Constructions in African Languages: A synchronic and diachronic survey. 'Habilitation' thesis. Leipzig: Institut für Afrikanistik, Universität Leipzig.
- . (2002). When 'say' is not say: The functional versatility of the Bantu quotative marker *ti* with special reference to Shona. In T. Güldemann & M. von Roncador (Eds.), pp. 253–287.
- & M. von Roncador (Eds.). (2002). *Reported Discourse: A meeting ground for different linguistic domains*. Amsterdam: John Benjamins [Typological Studies in Language 52].
- Hannan, M. (1984). *Standard Shona Dictionary*. Harare/ Bulawayo: College Press.
- Haspelmath, M. (1993). *A Grammar of Lezgian*. Berlin: Mouton de Gruyter [Mouton Grammar Library 9].
- Hawkins, J. A. & A. Cutler (1988.) Psycholinguistic factors in morphological asymmetry. In J. A. Hawkins (Ed.), *Explaining Language Universals* (pp. 280–317). Oxford: Basil Blackwell.
- & Gilligan, G. (1988). Prefixing and suffixing universals in relation to basic word order. *Lingua*, 74, 219–259.
- Hayward, R. J. (1990). Notes on the Zayse language. In R. J. Hayward (Ed.), *Omotic Language Studies* (pp. 210–355). London: School of Oriental and African Studies.
- . (1996). Compounding in Qafar. *Bulletin of the School of Oriental and African Studies*, 59 (3), 525–545.
- Heath, J. (1976). Simple and compound verbs: Conjugation by auxiliaries in Australian verbal systems – North-east Arnhem Land. In R. M. W. Dixon (Ed.), *Grammatical Categories in Australian Languages* (pp. 735–740). Canberra: Australian Institute of Aboriginal Studies [Linguistic Series 22].
- . (1998). Hermit crabs: Formal renewal of morphology by phonologically mediated affix substitution. *Language*, 74, 728–759.
- Heine, B. (1975). Language typology and convergence areas in Africa. *Linguistics*, 144, 27–47.
- . (1976). *A Typology of African Languages Based on the Order of Meaningful Elements* Berlin: Dietrich Reimer [Kölner Beiträge zur Afrikanistik 4].
- Hetzron, R. (1975). Genetic classification and Ethiopian Semitic. In J. Bynon & T. Bynon (Eds.), *Hamito-Semitic* (pp. 103–127). The Hague: Mouton.
- Himmelman, N. P. (1992). *Grammaticalization and Grammar*. Köln: Institut für Sprachwissenschaft, Universität zu Köln [Arbeitspapiere, Neue Folge 16].
- Hutchison, J. P. (1981a). *The Kanuri Language – A Reference Grammar*. Madison, WI: African Studies Program, University of Wisconsin.

- — —. (1981b). Kanuri word formation and the structure of the lexicon. In T. C. Schadeberg & M. L. Bender (Eds.), *Nilo-Saharan – Proceedings of the 1st Nilo-Saharan Linguistics Colloquium, Leiden, September 8–10, 1980* (pp. 217–237). Dordrecht: Foris.
- Leslau, W. (1945). The verb in Tigré (North-Ethiopic). *Journal of the American Oriental Society*, 65 (1), 1–26.
- — —. (1991). *Comparative Dictionary of Ge'ez (Classical Ethiopic): Ge'ez-English/English-Ge'ez with an index of the Semitic roots*. Wiesbaden: Otto Harrassowitz.
- Longacre, R. E. (1990). *Storyline Concerns and Word Order Typology in East and West Africa*. Los Angeles CA: Department of Linguistics, University of California [Studies in African Linguistics, Supplement 10].
- McGregor, W. (2001). Ideophones as the source of verbs in Northern Australian languages. In F. K. E. & C. Kilian-Hatz (Eds.), pp. 205–221.
- — —. (2002). *Verb Classification in Australian Languages*. Berlin: Mouton de Gruyter [Empirical Approaches to Language Typology 25].
- Nougayrol, P. (1989). *La langue des Aïki dits rounga (Tchad, République Centrafricaine): Esquisse descriptive et lexicale*. Paris: Geuthner.
- Nichols, J. (1992). *Linguistic Diversity in Space and Time*. Chicago, IL: University of Chicago Press.
- Palmer, F. R. (1974). Some remarks on the grammar and phonology of the 'compound verbs' in Cushitic and Ethiopian Semitic. In *IV Congresso Internazionale di Studi Etiopici (Roma 10–15 aprile 1972), Tomo II (Sezione Linguistica)* (pp. 71–77). Roma: Accademia Nazionale dei Lincei [Problemi Attuali di Scienza e di Cultura, Quaderno 191].
- Praetorius, F. (1894). Über die hamitischen Sprachen Ostafrika's. *Beiträge zur Assyriologie und vergleichenden Sprachwissenschaft*, 2, 312–341.
- Raz, S. (1983). *Tigre Grammar and Texts* Malibu: Undena [Afroasiatic Dialects 4].
- Reinisch, L. (1874). *Die Barea-Sprache: Grammatik, Text und Wörterbuch; nach den handschriftlichen Materialien von Werner Munzinger Pascha*. Wien: Wilhelm Braumüller.
- — —. (1881–90). Die Kunama-Sprache in Nordost-Afrika I–IV. *Sitzungsberichte der philosophisch-historischen Classe der kaiserlichen Akademie der Wissenschaften Wien*, 98 (1), 87–174/ 119,5/ 122,5/ 123,1.
- — —. (1893/4). Die Bed'auye-Sprache in Nordost-Afrika I–IV. *Sitzungsberichte der philosophisch-historischen Classe der kaiserlichen Akademie der Wissenschaften Wien*, 128 (3): 1–74/ 128, (7): 1–80/ 130, (7): 1–80/ 131, (3): 1–44.
- Roper, E. M. (1929). *Tu Bed'awi: An elementary handbook for the use of Sudan government officials*. Hertford: Stephen Austin and Sons.
- Sasse, H-J. (1979). *The Consonant Phonemes of Proto-East-Cushitic (PEC): A first approximation*. Malibu: Undena [Afroasiatic Linguistics 7:1].
- Schadeberg, T. C. (1987). Zwei areale Sprachmerkmale im Ostsudan unter besonderer Berücksichtigung des Nera-bena. In H. B. Mukarovsky (Ed.), *Leo Reinisch: Werk und Erbe* (pp. 213–229). Wien: Österreichische Akademie der Wissenschaften [Sitzungsberichte der Philosophisch-Historischen Klasse 492].
- Schultze-Berndt, E. (2000). *Simple and Complex Verbs in Jaminjung: A study of event categorisation in an Australian language*. Wageningen: Ponsen and Looijen [MPI Series in Psycholinguistics 14].
- Sundström, R. (1914). Some Tigre texts with transliteration and translation). *Le Monde Oriental*, 8 (1), 1–15.

- Thompson, E. D. (1976). Nera. In M. L. Bender (Ed.), *The Non-Semitic Languages of Ethiopia* (pp. 484–494). East Lansing MI: African Studies Center, Michigan State University [Occasional Papers Series, Monographs 5].
- . (1989). Kunama verb phrase. In M. L. Bender (Ed.), *Topics in Nilo-Saharan Linguistics* (pp. 305–346). Hamburg: Helmut Buske [Nilo-Saharan Linguistic Analyses and Documentation 3].
- Tosco, M. (2000). Is there an ‘Ethiopian language area’? *Anthropological Linguistics*, 42 (3), 329–365.
- Voeltz, F. K. E., & C. Kilian-Hatz (Eds.). (2001). *Ideophones*. Amsterdam: John Benjamins [Typological Studies in Language 44].
- Vries, L. J. de, & R. de Vries-Wiersma (1992). *The Morphology of Wambon of the Irian Jaya Upper-Digul Area with an Introduction to its Phonology*. Leiden: KITLV Press [Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 151].
- Waley, A., & C. H. Armbruster (1934). The verb ‘to say’ as an auxiliary in Africa and China. *Bulletin of the School of Oriental and African Studies*, 7 (3), 573–576.
- Wolff, H. E. (1993). Referenzgrammatik des Hausa: Zur Begleitung des Fremdsprachenunterrichts und zur Einführung in das Selbststudium. Münster/ Hamburg: LIT [Hamburger Beiträge zur Afrikanistik 2].

The OHO Constraint

Richard J. Hayward

University of London

A research interest in investigating the prosodic systems of a range of Ethio-Eritrean languages, has drawn my attention to a constraint of such wide occurrence that I am inclined to think that some interesting linguistic generalization may possibly lie at the heart of it. Nevertheless, the purpose of the present paper is not so much to speculate on an explanation for the constraint, but rather to highlight the phenomenon itself, and to define as carefully as possible the domains where it does and where it does not operate. It will then be appropriate to seek for any evidence of it outside the languages of The Horn.

The initial observations to be made relate to the distribution of higher pitch in word-sized nominal items¹ in a range of Ethio-Eritrean languages that from the viewpoint of prosody belong to widely differing types.

1. Higher pitch within words

In two earlier studies I have attempted to establish that certain Ometo languages, specifically Zayse (Hayward 1990) and Gamo (Hayward 1994), have prosodic systems that can best be regarded as of the tonal accent type.² It is not relevant or necessary to recapitulate those arguments here, nor to enter into any of the details of the analysis of those languages in terms of accentual and tonal components in their phonologies. What is germane to the present purpose is to appreciate that in the representative surface pitch contours of Zayse and Gamo nominal words, as shown in (1), higher pitch (HP) – what I shall later identify as the exponent of prosodic prominence – features once only. Actually, for the sake of Zayse, where we do find a relatively small number of items that show no higher-pitched part, we shall modify this observation to say that within a nominal word HP features once at most.³ In the following and in all subsequent examples every mora bearing HP will be marked with an acute accent and every lower pitched mora with a grave accent.⁴

- (1) a. GAMO
 gìró mole *’òðsánčà* labourer
 káálò to follow *míššàzàtà* the trees

b. ZAYSE

<i>gàidé</i>	cattle	<i>'itúm'à</i>	type of food
<i>'óók'k'àrò</i>	frog	<i>kààrànkó</i>	bat

An analysis of the East Cushitic language Qafar (Hayward 1991) claimed that this language too should be considered as having a tonal accent system. Once again HP for words uttered in isolation has to be seen as a once-per-word phenomenon. The same holds true for the Irob variety of Saho,⁵ a not-too-distant relative of Qafar. Forms from both languages are presented in (2)

(2) a. QAFAR

<i>gilé</i>	type of knife/sword	<i>màkààbánnà</i>	wisdom, sagacity
<i>kútà</i>	dog	<i>kòntàytó</i>	treesp.
<i>hàntùútà</i>	rat	<i>mòynób</i>	bull

b. IROBSAHO

<i>ínkà^s</i>	lice	<i>kùràákùr</i>	puppies
<i>ʃimbòòbá</i>	flower	<i>ʃimbòòbàytó</i>	a single flower
<i>kùrkùr</i>	puppy	<i>ʃàwààníytà</i>	a single locust

A recent investigation of Nara (Hayward 2000)⁶ shows that although this Nilo-Saharan language is a true 'tone language', HP within nominals is distributed in a somewhat similar way to what has been demonstrated for the tonal accent languages considered above. Thus, although Nara is a language in which tonal features are not at all compatible with an analysis involving an accentual apparatus, the fact that tone is not being employed in a totally paradigmatic way is to be suspected by the fact that there is an obvious distributional gap in the inventory of tone patterns found on nominal words; specifically, there is no HLH pattern. Given the relatively heavy functional load of tone in this language one might have expected such a pattern to occur. Typical examples of short words are given in (3).⁷

(3) NARA

<i>hàdíí</i>	skin	<i>wítàà</i>	five
<i>sùrñ</i>	grass	<i>sùrñ</i>	poison
<i>èdòò</i>	prayer-mat	<i>sààn</i>	three

As the examples suggest, High and Low tones are best assigned to morae, rather than syllables, and in addition to vowels sonorant consonants in post-vocalic (coda) position can bear tone. Moreover, the two morae of a long vowel may bear independent tones. The examples in (3) are all relatively short; but Nara words are commonly quite long, consisting of many morae. Words in which High (or Low) tones occur on a succession of contiguous morae are very frequent in the language. However, within the general assumptions for mapping tone to tone-bearing units as articulated in Autosegmental Phonology and, in particular, the meta-constraint known as the Obligatory Contour Principle (OCP) (Leben 1973), a sequence of like tones occurring on an uninterrupted sequence of tone-bearing units should usually be interpreted as resulting

2. Higher pitch within phrases

In the earlier work referred to on pitch in NPs in Gamo, Zayse, Qafar and Saho it was demonstrated that syntactic constituents comprising two or, occasionally, more words exhibit just one HP. Within the tonal accent analyses that were proposed for these languages, a single High tone was posited within what was termed a phonological phrase domain, and rules were formulated that allowed that tone to associate only with one word.¹⁰ The observation that led to such an account is that any word following the word bearing High tone is pronounced on a fairly low uniform pitch – regardless of what accentual properties it might possess lexically or what pitch features it might have when uttered in isolation.¹¹ Examples of NPs in the various languages are shown in (6)–(9) below.

(6) GAMO

- a. *bóóššà bòðrà*
a white ox cf. *bóóšši* ‘white’; *bóóra* ‘ox’
- b. *sékkì kùndidà mìššài*
that tree that fell cf. *sékkì* ‘that’; *kùndidà* ‘that fell’; *mìššài* ‘tree (Nom.)’

(7) ZAYSE

- a. *mààhé pìšò*
a leopard’s tail cf. *mààhé* ‘leopard’; *pìšò* ‘tail’
- b. *’èrállò kùllirì*
some guinea-fowls cf. *’èrállò* ‘some’; *kùllirì* ‘guinea-fowl (Plu.)’

(8) QAFAR

- a. *sìdihá sàgà*
three cows cf. *sìdihá* ‘three (Attrib.)’; *sàgà* ‘cow’
- b. *wòó ’ùndà ’àri*
that little house cf. *wòó* ‘that’; *’ùndà* ‘little’; *’àri* ‘house’

(9) IROB SAHO

- a. *kùlús sàgà*
a fat cow cf. *kùlús* ‘fat’; *sàgà* ‘cow’
- b. *kùlùs ságòg*
fat cows cf. *kùlús* ‘fat’; *ságòg* ‘cows’
- c. *táy dèq hìyàwtih ’àrè*
this tall man’s house cf. *táy* ‘this’; *dèq* ‘tall’; *hìyàwtih* ‘man’; *’àrè* ‘house’

In one of the two possessive construction types occurring in Nara NPs we find the order possessor – possessee. In such phrases the overall tone pattern is sometimes what would be expected in terms of combining the tone patterns of the component words of the phrase; but in other cases this does not occur. In all such exceptions, what we observe is that a non-initial (i.e., head) nominal appears without an expected High tone. NPs in which High tone drops on a head always consist of words having indi-

vidual lexical tone patterns that if combined within a phrase would create a contour containing two non-contiguous High tones, e.g.

(10) NARA

- | | | | |
|----|----------|----------------------|--|
| a. | H + LH | <i>gírbá hùbùr</i> | |
| | | colour of a mongoose | cf. <i>gírbá</i> 'mongoose'; <i>hùbùr</i> 'colour' |
| b. | H + LHL | <i>kúú bàttèg</i> | |
| | | the man's melon | cf. <i>kúú</i> 'man'; <i>bàttèg</i> 'melon' |
| c. | LH + LH | <i>àbbàá tòòkkù</i> | |
| | | father's wife | cf. <i>àbbàá</i> 'father'; <i>tòòkkù</i> 'wife' |
| d. | HL + LHL | <i>nààmbà àsàr</i> | |
| | | footprint of a calf | cf. <i>nààmbà</i> 'calf'; <i>àsàr</i> 'footprint' |
| e. | LHL + LH | <i>àbsòónà šìi</i> | |
| | | claw of vulture | cf. <i>àbsòónà</i> 'vulture'; <i>šìi</i> 'claw' |
| f. | HL + H | <i>ánìn dùmjàà</i> | |
| | | the woman's lamb | cf. <i>ánìn</i> 'woman'; <i>dùmjàà</i> 'lamb' |

If the combination of tone patterns on the nominals involved would not create such a contour, however, the High tone does not drop on the head, nor does any other tonal change occur, e.g.

(11) NARA

- | | | | |
|----|--------|---------------------|--|
| a. | HL + L | <i>wíínì nòò</i> | |
| | | eye of fly | cf. <i>wíínì</i> 'fly'; <i>nòò</i> 'eye' |
| b. | L + LH | <i>ngòò wòl'</i> | |
| | | our house | cf. <i>ngòò</i> 'our'; <i>wòl'</i> 'house' |
| c. | L + L | <i>gòò nòò</i> | |
| | | eye of frog | cf. <i>gòò</i> 'frog'; <i>nòò</i> 'eye' |
| d. | LH + H | <i>kàmbèré káló</i> | |
| | | camel's food | cf. <i>kàmbèré</i> 'camel'; <i>káló</i> 'food' |
| e. | H + H | <i>téé dùmjàá</i> | |
| | | his lamb | cf. <i>téé</i> 'his'; <i>dùmjàá</i> 'lamb' |

It would be out of the question to analyse the variation shown in (10) as 'grammatical', as, for example, the morphosyntactic expression of the genitive relationship. Rather, it is suggested that the phenomena seen here are all driven by a phonological constraint disallowing two (non-contiguous) High tones.

Indeed, I would claim that what we have seen in NPs in the tonal accent languages considered earlier and here in Nara are all manifestations of the same thing, which from this point on I shall refer to as the 'One-High-Only' (OHO) Constraint. Before proceeding to think more about this constraint, it will be necessary to establish that there is a correlation between languages where we find this constraint operating and head-final syntax.

3. The OHO Constraint and NP word order

Not all Ethio-Eritrean languages have thoroughgoing head-final syntax. For example, in most Southern Lowland East Cushitic languages nominal modifiers (adjectives, determiners, genitive NPs, relative clauses, etc.) follow the head noun in the NP, i.e., they belong to Heine's D2 Type (Heine 1976). Interestingly, the constraint we have been considering seems not to hold in such languages, so that we frequently find more than one HP per phrase. This is not to say that there may not be tonal sandhi phenomena in evidence; nor does it preclude tonal morphology being involved in expressing relationships – such as the genitive – within NPs. Indeed, there may be a number of factors present that militate against a purely compositional expression of pitch/tonal within the NP. But these can and should be distinguished sharply from the OHO Constraint. In examples (12)–(14) we see a variety of post-head modifiers in three D2 Type languages, which show the occurrences of pitch contours containing two (non-contiguous) occurrences of HP. In accordance with the view being advanced here that the OHO Constraint should operate whatever the morphotonological composition of the phrase is, an accompanying tonal analysis of the constituent words is not provided, though an understanding of these can be obtained from the sources indicated.

(12) SOMALI (Saeed 1999)

- a. *ʕààno lo'áád*
cow's milk (ibid. p. 64)
- b. *nínkùù sàddèh'áád*
the third man (ibid. p. 72)
- c. *nìmánkà bùùgággà kèèná*
the men who bring the books (ibid. p. 213)
- d. *šúqùl àdág*
hard work (ibid. p. 177)

(13) HARAR OROMO (Owens 1985)

- a. *dùbàrtíí dūrétííín*
the rich girl (ibid. p. 101)
- b. *bìnèyán tòrbán sàñ*
those seven animals (ibid. p. 91)
- c. *málláàn jòdòllée*
the children's cheeks (ibid. p. 103)
- d. *nàmníí bìnènsá àjjèssè*
the person who killed the wild animal (ibid. p. 132)

(14) ARBORE (Hayward 1984)

- a. *sááltà gùd'dà*
a fat woman (ibid. p. 201)

- b. *kèrólò séézzèlò*
these three dogs (ibid. p. 191)
- c. *hìndéérú náág*
the boy's flute (ibid. p. 155)
- d. *móhàrdò yééččè*
the man who came (ibid. p. 317)

As we have seen, the OHO Constraint operates in single words as well as in phrases. We might just expect therefore that non-observance of the constraint in an expanded NP with post-head modifiers might have a reflection in unexpanded NPs, i.e., in single nominals uttered in isolation. We certainly could see support for this expectation in (15) and (16), though, as the translations suggest, all of these words contain possessive or deictic determiner suffixes/clitics, which might suggest that they are really still phrasal.¹² However, in D'irayta, a Konsoid variety that shares the general Oromoid D2 word order typology, we find that in addition to polymorphemic nominals similar to those exemplified for Somali and Arbore in (15) and (16), we also find a number of monomorphemic nominals that violate the OHO Constraint, cf. (17)

- (15) SOMALI (Saeed 1999)
 - a. *gúrīgāygà*¹³
my house (ibid. p. 115)
 - b. *géèlàyágà*
our (exclusive) camels (ibid. p. 115)
 - c. *géèdkíí*
the tree (ibid. p. 174)
- (16) ARBORE (Hayward 1984)
 - a. *lúkkùtássèt*
her hens (ibid. p. 187)
 - b. *húndùttáw*
my backyard (ibid. p. 188)
 - c. *húzzùk'ló*
this star (ibid. p. 197)
- (17) D'IRAYTA
 - a. *tápàyyá* 'rat
 - b. *dákàllá* 'stone'
 - c. *káwnèlélá* 'yellow'

It might be suggested of course that the post-head order for modifiers and the failure to observe the OHO Constraint is simply a coincidence – and certainly there can be no way of knowing whether Somali, Oromo, Arbore and D'irayta would have observed the OHO Constraint if they had had pre-head modifiers. However, in support of the claim

that there is indeed a correlation involved, we note that certain of the languages that observe the OHO Constraint with pre-head modifying elements may also have post-head modifying constituents in the NP. Qafar has a type of relative clause¹⁴ terminating in *-iyya* which occupies this position. Significantly, the OHO Constraint does not hold in an NP structured in this way.

- (18) QAFAR (Parker & Hayward 1985:239)

yì ʕàmmiy gàddá lih yànih iyyà ..

my uncle who is rich cf. *yì gadda-lih ʕammi ..* ‘my uncle who is rich’

In addition to the relative verb, Irob Saho also allows adjectives and numerals in post-head modifying position. These occur with *iyya*, a form clearly cognate with Qafar *iyya* just considered. In none of these is the OHO Constraint in evidence, e.g.

- (19) IROB SAHO

a. *ʕìyàwtì yèmèètéh yìnè iyyà*

a man who came cf. *yèmèètéh yìnè ʕìyàwtì* ‘a man who came’

b. *ʕìyàwtì dèd iyyà*

a tall man cf. *dèd ʕìyàwtì* ‘a tall man’

c. *àdóh yángùltì iyyà*

three hyaenas Literally: ‘three, which are hyaenas’ cf. *àdòhá yàngùlà* ‘three hyaenas’

It might appear that in both phrases in (19c) the numeral precedes the head. However, in NPs containing numerals and terminating with *iyya*, it is the numeral that is really the head of the phrase. That this is so can easily be seen when the numeral is one that ends in a vowel, for then it can be seen to inflect for case, which in this language is always a property of a head, e.g.,

- (20) IROB SAHO

a. *làrààtánn-à lùbák iyyá yìgdifè*

twenty-Abs. lion which-are he-killed

(He) killed twenty lions

b. *làrààtánn-í lùbák iyyá yèmèètè*

twenty-Nom. lion which-are it came

Twenty lions came

Although attention has been directed to the very unique way in which Nara achieves conformity to the OHO Constraint, NPs in this language are of special significance at this point in the discussion in that they also allow an alternative post-head order for certain types of modifiers. Genitive NPs are one such type of modifier, and numerals are another.¹⁵ Deictic determiners and adjectives seem only to follow the head. Consider the following examples, where it will be observed that none of the tonal changes that take place in pre-head position (see (10)) are in evidence, e.g.

(21) NARA

- a. *nóótà yígù*
these eyes cf. *nóótà* 'eyes', *yígù* 'these'
- b. *wòl wóó gàà*
my house cf. *wòl* 'house', *wóó gàà* 'mine'
cf. pre-head order: *wóó wòl* 'my house'
- c. *tîl' ngírùkù*
a long python cf. *tîl'* 'python', *ngírùkù* 'long'
- d. *šîí kèssèl' gàà*
claw of a leopard cf. *šîí* 'claw', *kèssèl' gàà* 'of leopard'
cf. pre-head order: *kèssèl' šîí* 'claw of a leopard'
- e. *wíná šòónà*
four flies cf. *wíná* 'flies', *šòónà* 'four'
cf. pre-head order: *šòón' wìnà* 'four flies'

It is concluded in this section that there are good grounds for maintaining that a definite correlation exists between pre-head order of modifiers in NPs and the OHO Constraint.

4. Is the OHO Constraint simply an areal phenomenon?

It would seem very pertinent to ask whether the constraint we have been describing is confined to the Ethio-Eritrean linguistic area, where there is a heavy concentration of languages with strictly head-final syntactic systems, or whether it has any cross-linguistic (perhaps universal) manifestations outside this area in languages of similar syntactic type. The empirical evidence that could answer such a question would of course only be obtained through a very extensive research programme. It did however prove possible for me to conduct a brief pilot investigation of one non-Ethiopian language of the requisite type; namely Japanese. Japanese exhibits thorough-going head-final syntax; it also has a well-investigated tonal accent system; thus it furnishes an excellent test case.¹⁶

The following account of the facts of tonal accent prosody in Standard (Tokyo) Japanese has a fairly wide currency; beginning from, for example, McCawley 1968.¹⁷

When a word is uttered in isolation the manifestation of what I have been referring to in this paper as HP occurs on all TBU's except the first (which unless it is accented is always low-pitched).¹⁸ A word may either be accented on the first mora of one of its syllables, or it may be unaccented. If a word contains an accent, there is a drop in pitch following the accented TBU. If, however, a word is unaccented, HP continues throughout the word; e.g.

(22) JAPANESE

	accentual status ¹⁹		manifestation of HP
a.	<i>mAkura</i>	pillow	<i>mákùrà</i>
b.	<i>kagAmi</i>	mirror	<i>kàgámì</i>
c.	<i>atamA</i>	head	<i>àtámá</i>
d.	<i>sakana</i>	fish	<i>sàkáná</i>

The situation just described leads to an analytical indeterminacy in distinguishing un-accented words from words accented finally, as in the examples (22) c. & d. However, when what McCawley terms 'postpositions'²⁰ follow such words, the indeterminacy is resolved; e.g.

(23) JAPANESE

Phonological phrase (N + postposition *kara* 'from')

- a. *mákùrà kàrà*
- b. *kàgámì kàrà*
- c. *àtámá kàrà*
- d. *sàkáná kàrà*

The examples we have considered thus far clearly conform to the OHO Constraint. Moreover, as McCawley points out, there some 'postpositions' in Japanese that (unlike *kara*) appear themselves to be accented. The examples in (24) show that these behave distinctly with unaccented nouns and with nouns accented finally.

(24) JAPANESE

Phonological phrase (N + postposition *made* 'to')

- | | | | |
|----|--------------------|---|--------------------|
| a. | <i>mákùrà màdè</i> | ← | <i>mAkura mAde</i> |
| b. | <i>kàgámì màdè</i> | ← | <i>kagAmi mAde</i> |
| c. | <i>àtámá màdè</i> | ← | <i>atama made</i> |
| d. | <i>sàkáná màdè</i> | ← | <i>sakana made</i> |

In order to arrive at the surface pitch contours of a., b., and c. in (24), McCawley argues that "there must be a rule that eliminates an accent that is preceded by another accent in the same phrase" (ibid. p. 115). This is, of course, very reminiscent of what we have been discussing in the Ethio-Eritrean tonal accent languages. But more importantly, we must note that the result of this rule of elimination is a conformity to the OHO Constraint within this type of phrase.

Now it is proper to ask whether there are any manifestations of the OHO Constraint in expanded NPs in Japanese similar to what was seen in the Ethio-Eritrean languages. Here again we encounter positive results. Thus, in NPs containing pre-head modifiers such as adjectives, genitive NPs, numeral quantifiers, or relative clauses, we find at least an optional conformity to the OHO Constraint; e.g.

(25) JAPANESE

- a. *àkâi kàgàmi* (~ *kàgámi*) 'red mirror'
cf. *akAi* 'red'; *kagAmi* 'mirror'
- b. *kàsíkói àtàmà* (~ *àtámá*) 'clever brain'
cf. *kasikO-i* 'clever'; *atamA* 'head'
- c. *kàrè-nò tàràgò* (~ *tàrágò*) 'his egg'
cf. *kAre no* 'him- possessive'; *tamAgo* 'egg'
- d. *kòwàrètà kàgàmi* (~ *kàgámi*) 'the mirror that broke'
cf. *kowAreta* 'that broke'; *kagAmi*
- e. *nísàtù-nò hòñ (hòñ)* 'two books'
cf. *nI-satu-no* 'two-num.classfr-of'; *hOñ* 'book'

Moreover, there are other structures in Japanese in which, as in the 'postpositional phrases' considered earlier, there is clear evidence of the suppression of HP features in head words, such that a conformity to the OHO Constraint could be claimed. For example, a nominal complement + copula construction or a direct object or locative complement + verb construction clearly do show this, though only in the former case does the process seem to be obligatory. In example (26c), where the nominal complement consists of an unaccented noun, we observe the same behaviour as in the postpositional phrases, i.e., that the accent on the copula is not eliminated because there is no preceding accent in the phrase. E.g.

(26) JAPANESE

- a. (*kòré wá*) *àtámá dèsità* (This) was a head
cf. *atamA* 'head'; *dEs-ita* 'e-past'
- b. (*kòré wá*) *kàgámi dèsità* (This) was a mirror
cf. *kagAmi* 'mirror'; *dEs-ita* 'be-past'
- c. (*kòré wá*) *sàkáná dèsità* (This) was a fish
cf. *sakana* 'fish'; *dEs-ita* 'be-past'
- d. *ìé ò mîtà* (~ *míta*) I saw a house
cf. *iE o* 'house-object'; *mI-ta* 'see-past'
- e. *gòhàn ò tàbètà ~ gòhà- ò tàbètà* (I) ate cooked rice
cf. *gOhañ o* 'cooked rice-object'; *tAbe-ta* 'eat-past'
- f. *ìé-è kàèrù* (~ *ìé è kàèrù*) 'to return home'
cf. *iE-e* 'house-to'; *kAeru* 'to return'

Although it was not mentioned earlier, following the presentation of this last set of Japanese examples it is now appropriate to observe that focussed complement + verb phrases (i.e., core VPs) behave in a virtually identical way in languages of the Ethio-Eritrean region. The occurrence of this has been noted with regard to Zayse (Hayward 1990: 239–40), Qafar (Parker & Hayward 1985: 221–3), and Nara (Hayward 2000).

(27) ZAYSE

- a. *hàméttèt wòrgè* 'I want to go'; cf. *hàméttèt* 'it-is-I-going', *wòrgè* 'want'
- b. *ʔésà gwíidín* 'I hit him'; cf. *ʔésà* 'him', *gwíidín* 'hit-past'

(28) QAFAR

- a. *yòó tùblè* 'She saw me'; cf. *yòó* 'me', *tùblè(h)* 'she saw'
- b. *kímàl gèndè* 'We went yesterday'; cf. *kímàl* 'yesterday', *gèndè(h)* 'we went'

(29) NARA

- a. *tîl sàg* 'Kill (the) python!'; cf. *tîl* 'python', *sàg* 'kill!'
- b. *téb* [*kùdùl síttò*]_{VP} 'He / she killed (the) hyaena'; cf. *kùdùl* 'hyaena', *síttò* 'he / she killed'
- c. *àggí* [*sàà láàgò*]_{VP} 'We drank (the) milk'; cf. *sàà* 'milk', *láàgò* 'we drank'

Although this investigation of the possible occurrence of the OHO Constraint outside of the Ethio-Eritrean language area has been restricted to a consideration of one language only, namely Japanese, the positive results obtained encourage one to extend the research with an enlarged linguistic data base.

5. Conclusion

It has been the purpose of this paper to draw attention to a pervasive constraint affecting the distribution of high pitch in words and phrases in languages of the Ethio-Eritrean area. Although what I have here dubbed the OHO Constraint manifests itself in languages of several distinct prosodic types, it appears to do so only if the language exhibits strict head-final syntax, so that it is consistently absent from NPs in languages in which modifying elements follow (rather than precede) their heads. A brief examination of the distribution of high pitch in Japanese, a well-investigated language with thorough-going head-final syntax suggests that the constraint may not be confined to the language area where it was first noted. This invites a careful examination in the future of a much wider range of languages having the requisite syntax. Should the constraint turn out to be a quite general one, it will naturally lead on to the important question of how to account for it in a theoretically satisfying way.

Notes

1. The term 'nominal' is based on morphological considerations and is used to refer not just to nouns and pronouns but also, very commonly, to adjectives and numerals; though the precise definition of the set is a language-specific matter.
2. Cf. also Azeb Amha 1996 concerning prosody in the major Ometo variety Wolaitta.

3. In the appropriate contexts both Gamo and Zayse words exhibit a process that spreads a high tone rightwards onto a following tone-bearing unit. There are also a number of Gamo nominals in which HP appears on consecutive syllables/morae lexically. These facts in no way invalidate the statement made here. The point receives general justification in the treatment of Nara below.
4. In the segmental transcription of examples in this study the following symbols differ from IPA usage: \dot{s} = IPA [S]; $s\ddot{Y}$ = IPA [$t^{\circ}s\ddot{O}$]; $s\ddot{Y}'$ = IPA [$t^{\circ}s'$]; (= IPA [$t^{\circ}S\ddot{O}$]; (' = IPA [$t^{\circ}S'$]; d' = IPA [()]; y = IPA [j]; and ' = IPA [/]. Double consonant or vowel letters represent geminate consonants or long vowels respectively. Some of the languages treated here have tonal accent systems, and an appropriate analysis of these needs only to indicate one (syntagmatic) pitch event for any given domain. For the purposes of the present paper, however, pitch has been indicated on every tone bearing unit. This has been done in order to achieve greater comparability between such languages and languages that exhibit more clearly (paradigmatic) tonal properties.
5. The point could probably be extended to all Saho varieties, though I do not have the necessary field data to hand to establish it. A partial description of the tonal accent system of Irob Saho appears in Hayward 1991.
6. Nara, sometimes incorrectly called 'Nera', has four dialects; the material discussed here represents the Higir dialect.
7. The Low tone in Nara is marked with a grave accent.
8. A cursory examination of Nara words containing sequences of like tones on contiguous morae makes it clear that some diacritic device, such as pre-association of certain tones, would be necessary to handle tone-to-TBU association adequately. The point does not need to be taken up here.
9. For example: Rendille (Pillinger 1989), Gimira (Breeze 1990; Wedekind 1995), Shinassha (Lamberti 1993a), Yemsa (Lamberti 1993b), D'irayta (Hayward 1998), Arbore (Hayward 1984), Kunama (Connell, Hayward, & John Abraha Ashkaba 2000)
10. In three of these languages the word with which the High tone associates is the first (left-most); in the case of Irob Saho, however, the High tone may sometimes appear in the second word; for the details, see Hayward 1991.
11. I.e., as a one-word phonological phrase.
12. Indeed in Owens' treatment of Harar Oromo possessive and deictic determiners are given separate word status.
13. The diacritic > indicates falling pitch, analysed by Saeed (and others) as H + L. Saeed indicates this tonal sequence by means of a grave accent. To avoid confusion with the representation of L tone in other languages considered here, I have replaced \ddot{Y} by >.
14. Parker and Hayward (1985) termed this type of relative 'non-restrictive'. Bliese (1981:24) notes that such forms occur obligatorily when an NP occurs with a relative clause preceded by some other modifying element.
15. It will be observed that when the genitive follows the head, it is followed by *gàà*. Similarly, there are distinct pre-head and post-head forms for some numerals. No attempt is made here to explain these differences, and I do not believe that the correlation being argued for is in any way invalidated by such syntactic and morphological differences.
16. I wish to express here my gratitude to Taeko Maeda, a former doctoral research student at SOAS, for her very valuable and perceptive cooperation in an investigation which led to the observations reported here.

17. Many of the examples in this section are taken from McCawley's article, though, as I shall point out, my interpretation of accent differs from his.
18. That the TBU in Japanese is the mora, rather than the syllable, is argued for by Maeda (2001).
19. For clarity, accentuated vowels are indicated by writing them in capital letters.
20. While some of the items that are referred to as 'kagami postpositions' would be recognised as such cross-linguistically (i.e., they are post-nominal adpositions), there are also a number of items (e.g., the topic marker *wa*, the subject marker *ga*, a word *sika* translating as 'only', etc.) that do not fit into such a syntactic category. Since, however, both McCawley's and our concern is a prosodic constituent, namely the phonological phrase, no serious objection can be sustained on grounds of the heterogeneity of this set from a syntactic point of view.

References

- Andrzejewski, B. W. (1964). *The Declensions of Somali Nouns*. London: School of Oriental and African Studies.
- . (1970). The role of tone in the Borana Dialect of Galla. In *Proceedings of the III International Conference of Ethiopian Studies* (pp. 88–89). Addis Ababa: Institute of Ethiopian Studies.
- Azeb Amha (1996). Tone-accent and prosodic domains in Wolaitta. *Studies in African Linguistics*, 25 (2), 111–138.
- Banti, G. (1988). Two Cushitic systems: Somali and Oromo nouns. In H. van der Hulst & N. Smith (Eds.), *Autosegmental Studies on Pitch Accent* (pp. 11–49). Dordrecht: Foris.
- Bliese, Loren (1981). *A Generative Grammar of Afar*. Dallas, TX: Sumer Institute of Linguistics.
- Connell, B., Hayward, R. J., & J. A. Ashkaba (2000). Observations on Kunama tone (Barka dialect). *Studies in African Linguistics*, 29 (1), 1–41.
- Breeze, M. (1990). A sketch of the phonology and grammar of Gimira (Benchnon). In R. J. Hayward (Ed.), *Omotic Language Studies* (pp. 1–66). London: School of Oriental and African Studies.
- Hayward, R. J. (1984). *The Arbore Language: A first investigation (including a vocabulary)*. Hamburg: Helmut Buske Verlag [Kuschitische Sprachstudien 2].
- . (1990). Notes on the Zayse language. In R. J. Hayward (Ed.), *Omotic Language Studies* (pp. 210–355). London: School of Oriental and African Studies.
- . (1991). Tone and accent in the Qafar noun. *York Papers in Linguistics*, 15, 117–137.
- . (1994). A preliminary analysis of the behaviour of pitch in Gamo. In B. Zewde et al. (Eds.), *The Proceedings of the 11th International Conference of Ethiopian Studies, Addis Ababa, April 1991*, Vol. 1 (pp. 481–494). Addis Ababa: Addis Ababa University Press.
- . (1998). Vowel reduction, tone and nominal declension in D'irayta. In M. Janse (Ed.), *Productivity and Creativity: Studies in general and descriptive linguistics in honour of E. M. Uhlenbeck* (pp. 503–519). Berlin: Mouton de Gruyter [Trends in Linguistics, Studies and Monographs 116].
- . (2000). Observations on tone in the Higur dialect of Nara. In R. Vossen, A. Mietzner, & A. Meissner (Eds.), *'Mehr als nur Worte.' Afrikanistische Beiträge zum 25. Geburtstag von Franz Rottland* (pp. 247–267). Köln: Rüdiger Köppe Verlag.
- . (Ed.) (1990). *Omotic Language Studies*. London: School of Oriental and African Studies.

- , & Gemetchu Megerssa (1996). Nominal tone in Western Oromo. In R. J. Hayward I. M. Lewis (Eds.), *Voice and Power: The culture of language in North-East Africa. Essays in honour of B. W. Andrzejewski* (pp. 21–36). London: School of Oriental and African Studies, University of London [African Languages and Cultures Supplement 3].
- Heine, B. (1976). *A Typology of African Languages Based on the Order of Meaningful Elements*. Berlin: Reimer [Kölner Beiträge zur Afrikanistik 4].
- Hyman, L. M. (1981). L'accento tonale in somalo. In G. R. Cardona & F. Agostini (Eds.), *Fonologia e lessico. Studi Somali 1* (pp. 109–139). Roma: Ministero per gli Affari Esteri e Comitato tecnico linguistico per l'Università nazionale Somala.
- Lamberti, M. (1993a). *Die Shinassha-Sprache: Materialien zum Boro* [Studia Linguarum Africae Orientalis, 4]. Heidelberg: Universitätsverlag C. Winter.
- . (1993b). *Materialien zum Yemsa*. Heidelberg: Universitätsverlag C. Winter [Studia Linguarum Africae Orientalis 5].
- Leben, W. (1973). *Suprasegmental Phonology*, PhD Dissertation, MIT.
- Moreno, M. M. (1938). *Grammatica teorico-practica della lingua galla*. Milano: A. Mondadori.
- Maeda, T. (2001). *The Mora and the Syllable in KiMvita (Mombasa Swahili) and Japanese*, PhD Dissertation, School of Oriental and African Studies, University of London.
- McCawley, J. (1968). *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.
- Owens, J. (1980). Observations on tone in the Booran dialect of Orome (Galla). *African Language Studies*, 17, 141–196.
- . (1985). *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Helmut Buske Verlag [Kuschitische Sprachstudien 4].
- Parker, E. M., & R. J. Hayward (1985). *An Afar – English – French Dictionary (with grammatical notes in English)*. London: School of Oriental and African Studies.
- Pillinger, O. S. (1989). *Accent, Tone and Prosodic Structure in Rendille*, PhD Dissertation, School of Oriental and African Studies, University of London.
- Saeed, J. (1999). *Somali*. Amsterdam: John Benjamins.

The word in Luganda

Larry M. Hyman and Francis X. Katamba

University of California, Berkeley / University of Lancaster

In this paper we take a close look at the phonological, morphological, syntactic, and semantic properties of the word in Luganda. We show that criteria for word status and word delimitation conflict, whether taken from different components of the grammar, e.g. syntactic word vs. morphological word, or even when limited to one component, e.g. the word-like domains phonologically defined by tonal vs. vowel length criteria do not coincide. In the course of our study we also distinguish lexical words, which can be simplex (root + affixes) or compounded, from postlexical words, which may consist of lexical words + clitics, as well as phrasal words which can be whole sentences.

1. Introduction

This paper investigates the word in Luganda, a member of the Narrow Bantu subgroup of the Bantu sub-branch of the Benue-Congo branch of the Niger-Kordofanian language family which is spoken in southern Uganda.¹ The major questions posed by the paper are the following:

- a. Can the “word” be defined?
- b. If not, why not?
- c. If yes, is the “word” a universal?

As a Narrow Bantu language, Luganda is typical with its rich, largely agglutinative morphology. In addition, its extraordinarily complex segmental and tonal phonology makes frequent reference to the notion of the word. Luganda provides fertile ground for addressing the above questions. The need to disentangle different aspects or kinds of words as shown in (1) is generally recognised (Dixon & Aikhenvald 2002).

- (1)
 - a. the semantic word
 - b. the syntactic word
 - c. the morphological word
 - d. the phonological word
 - e. other (orthographic word, cognitive word etc.)

Semantic and morphological criteria for the word are problematic for a number of reasons.

First, the amount of semantic material that can be incorporated into a word varies enormously, cross-linguistically. What is said using one word in one language may require a sentence containing many words in another. Compare the English and Luganda utterances in (2).

- (2) a. *te tú- lí kí bá- gùl- ir- a*
 not we FUTURE it them buy APPLICATIVE INFL.SUFFIX
 b. 'we will not buy it for them'

The Luganda utterance in (2a) consists of a single word containing eight morphemes.² Its English translation in (2b) by most accounts would be analyzed as seven (monomorphemic) words.

Second, a morphological definition of the word as a stem plus affixes runs into the problem of words that consist of a proclitic + enclitic (i.e. no stem).

- (3) a. *byaa=mú* 'of in there'
 b. *waa=ki* 'of what?'

We return to these words of this type in (11) and (12) below.

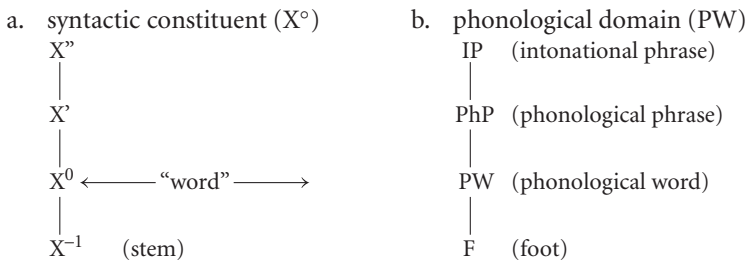
Third, syntactic phrases may enter into a paradigm with words (morphological objects), e.g. comparatives of some monosyllabic adjectives appear to block the periphrastic comparative: *bigger, smaller* vs. (*)*more big, (*more small* (see Poser 1992, who distinguishes morphological vs. word formation rules).

Finally, there is the notorious problem of "phrase words" (Bloomfield 1933) such as *Johnny-come-lately* in English, which constitute a single word (e.g. noun), but which have complex multi-word syntactic structure. As seen in (4):

- (4) "Phrase words" (Bloomfield 1933) such as *Johnny-come-lately*
- a. "orig. U.S.: (a) a newcomer; (b) = Johnny Raw; (c) fig. and attrib." (OED Online, 2nd Edition, 1989)
 - b. 'But it's Johnny Comelately, aint it, you?' said a young mizzen topman. (1839 C. F. BRIGGS Adv. H. Franco I. 249)
 - c. He may be an old barbarian, but he's entitled to more consideration than these Johnny-come-lately's who cruise along the coast after trade. (1924. R. DALY' Outpost xiv. 139)
 - d. The Midlands are.all Johnny-come-latelys who coined money out of the war. (1952 E. COXHEAD Play Toward iii. 88)
 - e. Postwar planning in these United States was no Johnny-come-lately. (1946 M. SHULMAN Zebra Derby iii. 22)
 - f. The excessive power and renown of many Johnny-come-lately anti-Communists. (1953 Amer. Scholar XXIII. 17)

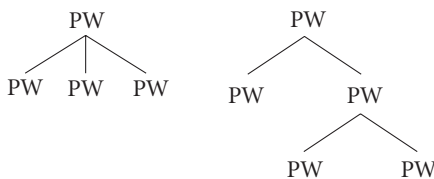
Scholars are not always sure how to treat phrase words, e.g. whether to write them as one word or hyphenate them, whether to indicate their plural with vs. without an apostrophe etc. As we shall see below, Luganda is full of such entities, which pose even greater problems than in English. Although a consistent definition of the word remains elusive, its place in the linguistic hierarchy is, at first blush, reasonably clear. It occurs in an intermediate position within a hierarchy of syntagmemes. Many modern linguists would place it at the position in the hierarchy shown in (5).

- (5) The word occurs in an intermediate position within a hierarchy of syntagmemes



According to Prosodic Domain Theory (Selkirk 1984; Nespor & Vogel 1986), which subjects the phonological hierarchy in (5b) to the Strict Layer Hypothesis (SLH), feet should be grouped into phonological words, which are grouped into phonological phrases, which are then grouped into intonational phrases (and finally into an utterance). This approach encounters problems, however, since, for a variety of reasons, elements do not always occupy the same position in the hierarchy. First, nesting, schematized in (6), is not uncommon, e.g. a PW can potentially consist of more than one PW, with one of the two structures indicated.

- (6) Nesting



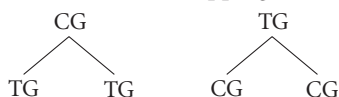
Second, although also outlawed by the SLH, recursion, schematized in (7), does occur, due to the phenomenon of rank-shifting, which was illustrated with the Johnny-come-lately example in (4) and to which we return below with Luganda examples such as *mugenzi tázzê* 'delinquent debtor' (lit. 'traveller has not returned') in (17).

- (7) Recursion/rank-shifting



Lastly, intersection of domains relevant for hierarchies may be found such that for the purposes of certain phonological processes the relevant domains may overlap or intersect. Thus, in Luganda, the clitic group (CG) and tone group (TG) intersect (Hyman, Katamba & Walusimbi 1987; Hyman 1988), as shown in (8), and cannot therefore be neatly placed in a single phonological hierarchy:

(8) Intersection/overlapping



We return to this below in Section 4.

For the above – and perhaps other reasons – the word in linguistics cannot be characterised in a simple manner. But that is not to say that any attempt to elucidate its position is bound to be fruitless. A possible useful strategy is to seek generalizations on how the various word-like entities function and interact in a total grammar (cf. Zwicky 1990; Dixon & Aikhenvald 2002). This involves, first, determining the criteria that define Xo and PW in a wide array of languages, and second, comparing word-like entities of a given language to the canonical (or prototypical) word defined by the morphology.

Luganda provides a rich laboratory for studying such questions. In the following sections, we first provide a brief background to Luganda morphology (§2), followed by a discussion of the two phonological properties that have been criterial for establishing word status in the language: quantity (§3) and tone (§4). We return to the questions raised in this section in a brief conclusion in § 5.

2. Background: Luganda morphology

We begin by sketching the structure of the canonical morphological word in Luganda, which is very clearly agglutinating. The Luganda canonical morphological word (CMW) is summarized in (9a).

- (9) a. CMW → prefix(es) + stem
 b. Noun → (augment) (noun class prefix) + stem e.g. *o-mu-limi* ‘farmer’
 Adj V- CV-/N- *mu-limi* ‘he’s a farmer’
 c. Verb → (prestem) + stem
 Stem → root + extensions + FV (= inflectional final vowel, usually -a)
 e.g. [*a- bá- tà- lì- [lì-m- ir agan- a]*]_{stem}
 AUG SUBJ NEG FUT ROOT APPL RECIP FV
 ‘They who will not cultivate for each other’

As seen, the CMW consists of one or more prefixes followed by a stem. The typical noun and verb structures are indicated and exemplified in (9b, c). While these struc-

tures are typical, they are not the only possibilities. Thus, as seen in (10), words can appear without prefixes:

- (10) a. noun class 1a Ø- (plural via noun class 2a proclitic *ba=*)
katonda ‘God’ cf. *mu-limi* ‘farmer’ (class 1)
ba=katonda ‘gods’ *ba-limi* ‘farmers’ (class 2)
lujuuju ‘drunkard’ *lu-sózi* ‘hill’ (class 11)
ba=lujuuju ‘drunkards’ *n-sózi* ‘hills’ (class 10)
doodô ‘spinach’ (no. pl.)
Walúsimbi (proper name) *ba=Walúsimbi* ‘the Walusimbis’
- b. imperative verbs (2nd person singular affirmative without object)
siba ‘tie!’ (2sg.) cf. *n-sibâ* ‘tie me!’
/sib-a/ *tô-sibâ* ‘don’t tie!’ */te-ô-/* (NEG+2sg.)
 FV *mu-sibê* ‘tie!’ (2pl.)

The singular class 1a nouns in (10a) lack a prefix and show that words can be monomorphemic. Second person singular affirmative imperative verbs such as *siba* ‘tie!’ in (10b) consist of a prefixless stem.

In addition, (11) and (12) show that words exist in Luganda which do not have an obvious root or stem structure:

- (11) prefix(es) + suffix (?), e.g. demonstrative */-o/* ‘that/those’ (near hearer)
- | | | |
|----------------------------|----------------------------|-----------------------------|
| cl. 1 <i>oyo</i> /o-i-o/ | cl. 6 <i>ago</i> /a-ga-o/ | cl. 11 <i>olwo</i> /o-lu-o/ |
| cl. 2 <i>abo</i> /a-ba-o/ | cl. 7 <i>ekyo</i> /e-ki-o/ | cl. 12 <i>ako</i> /a-ka-o/ |
| cl. 3 <i>ogwo</i> /o-gu-o/ | cl. 8 <i>ebyo</i> /e-bi-o/ | cl. 13 <i>otwo</i> /o-tu-o/ |
| cl. 4 <i>egyo</i> /e-gi-o/ | cl. 9 <i>eyo</i> /e-i-o/ | cl. 14 <i>obwo</i> /o-bu-o/ |
| cl. 5 <i>elyo</i> /e-li-o/ | cl. 10 <i>ezo</i> /e-zi-o/ | cl. 15 <i>okwo</i> /o-ku-o/ |
- (12) proclitic + enclitic
- | | |
|---------------------------|--|
| <i>mu=kí</i> ‘in what?’ | <i>kyaa=kí</i> ‘it’s for what?’ (‘what’s it for?’) |
| <i>ku=kí</i> ‘on what?’ | <i>byaa=mú</i> ‘those (belonging) inside’ |
| <i>na=kí</i> ‘with what?’ | <i>byaa=kô</i> ‘those (belonging) on there’ |

The demonstratives in (11) appear to consist of two prefixes, e.g. class 2 *a-ba-*, followed by a suffix *-o* ‘this’. The combinations in (12) consist of a combination of proclitic + enclitic (see §3). Assuming that the forms in (11) and (12) have word status, not all morphological words are “canonical,” thus making it difficult to provide a clear, single morphological definition of the word in Luganda.³

Any grammatical definition of the word is further complicated by at least three factors. First, the existence of cliticised forms that, on the one hand, are not autonomous words, but, on the other, enjoy greater freedom than affixes. These clitics show a considerable degree of grammatical diversity, e.g. attaching to hosts from all morphological categories. Examples of nominal proclitics enclitics are shown in (13a) and (13b), respectively (cf. §3 for verbal clitics):

- (13) a. *ba*= *Walusimbi* 'the Walusimbis'
 ku= *ki-tabo* 'on the book'
 mu= *n-nyumba* 'in the house'
 na= *mu-limi* 'with a farmer'
 byaa= *mu-limi* 'those (cl. 8) of the farmer'
- b. *ki-dee* = *kyo* 'your (sg.) bell'
 ki-dee = *kye* 'his/her bell'
 ki-dee = *ki* 'which bell?'

Words containing the so-called ‘-a of relationship’ (a.k.a. connective, associative, genitive), which we analyze as a proclitic, are especially problematic and cause problems of word division in the standard orthography which are due to their unclear word status. Ashton et al (1954:104) recommend treating the forms in (14) as single orthographic words.

- [illegible]

The justification for this is that we have lexicalised here nouns built up on “the possessive noun form but without an expressed antecedent”. Elsewhere, as in (15), Ashton recommends writing the proclitic as a separate word, although phonologically and grammatically it is no different:

- (15) a. *wa mukazi* ‘of the woman’ (*waa= mukàzi*)⁴
 b. *ba Mulondo* ‘of Mulondo’ (*waa= Mulondo*)

It seems the distinction is meant to show that lexicalisation has taken place in (14), but not in (15).

Similar problems potentially arise in lexicalized and productively created compounds. Compounds are a special case of words that are built up using phrase structure rules to combine pre-existing words into syntactically complex words. In this respect they resemble syntactic phrases. Examples of compounds are given in (16).

- (16) a. *mulwa – kujjula* ‘woman slow to serve food’ (lit. delayer + to serve food)
 katwé – kàsa ‘stupid person’ (lit. small head + empty)
 nkyá – mùzi ‘type of bark-cloth tree’ (lit. morning + small root)
 b. *mulyá – màtooke* ‘plantain-eater’ (lit. eater + plantains)
 munywá – mwènge ‘beer-drinker’ (lit. drinker + beer)
 mumwá – mítwê ‘head-shaver’ (lit. shaver + head)

When compounding is encountered, a question that always arises is how does one distinguish between compounds and syntactic phrases? We will turn to this task

shortly. But first we will consider a third category of problematic complex words, namely phrase words, illustrated in (17).

- (17) a. *mugenzi – tázzê* ‘delinquent debtor’ (lit. traveler hasn’t returned)
 mwáámì – akóóyè ‘easy chair’ (lit. chief has tired)
 kyáálà – kimpáddè ‘thief’ (lit. fingernail has given me)
- b. *Nsí – yàlèèta* (proper name) (lit. country brought)
 Túlíná – ómùbéezi (proper name) (lit. we have a helper)
 Sílivá – kùno (proper name) (lit. I will not leave here)
- c. *ndábírwáá =kô* ‘mirror’ (lit. I am seen from it)
 kyaa= kulábira =kô ‘example’ (lit. that to see from)
- d. *Wáálábyèè =kí* (proper name) (lit. what have you seen yet?)
 gwe - bátákígámbyé (proper name) (lit. one that they haven’t said it to)

As seen, these forms appear to be words based on full sentences. As seen, they can have a wide range of internal syntactic structure, e.g. two word combinations of subject-verb or verb-object in (17a, b). Other comparable forms involve clitics, as in (17c, d). Finally, as seen in the glosses, many of these are proper names, some of them derived from proverbs.

Previous literature on Luganda has assumed that phonology provides a number of litmus tests for wordhood:

- (18) a. long vowels are allegedly shortened at the end of a word “the final syllable of a word spoken in isolation is always short. Within the sentence too, final syllables of words are usually short, and this fact has been of great value in assessing word division.” (Tucker 1962: 155)
- b. a word allegedly may not contain more than one HL tonal sequence
- c. “(at most) a single High-Low sequence occurs in every major lexical item in the language.” (Heny 1974: 1)

Both vowel length (18a) and tone (18b) have been said to be indicators of word status. In the next section we evaluate the extent to which the claim that the word in Luganda can be defined phonologically is justified.

3. Quantity as a criterion for word status

Over the last 40 years, the literature on Luganda phonology has remarked on a process by which long vowels are shortened in final position. See the quote from Tucker in (18a). In this context the assumption is that “final position” means “at the end of a word”. But this conceals a crucial ambiguity. All researchers are aware that such “words”, postulated to account for final vowel length, may consist of a “full” word (or “host”) and one or more proclitics or enclitics. For instance, commenting on *asomyé* ‘he has read’, Stevick (1969:4) states: “the last syllable of the isolated word is short. Be-

fore an enclitic, however, it receives the expected two moras.” We will explore this with the data in (19).

(19) Final Vowel Shortening (FVS)

- a. *kulábwà* ‘to be seen’ /*ku-láb-u-a/* → *kulábwàà* → *kulábwà*
- b. *kulábwà Walúsìmbi* ‘to be seen by Walusimbi’ gliding+CL FVS
- c. *kulábwàà =kô* ‘to be seen a little’

The word-final vowel [a] is phonetically short in (19a) where it is in absolute final position, and in (19b) where it is followed by an object NP, which is not an enclitic. By contrast, its phonological length is preserved in (19c) where it is followed by the locative enclitic =*kô*.

As (20) shows, proclitics also fail to undergo FVS the length of their final vowel being protected by the host:⁵

- (20) a. -*aa*=⁶ ‘genitive linker’: *kikópó kyáá= Wálúsìmbi* ‘cup of Walusimbi’ /*ki-aa=*/
luggí lwáá= Wálúsìmbi ‘door of Walusimbi’ /*lu-aa=*/
 b. -*ee*= ‘subject cleft marker’: *kikópò kyèè= kyáágwà* ‘it’s a cup that fell’ /*ki-ee=*/
luggí hwèè= lwáágwà ‘it’s a door that fell’ /*lu-ee=*/

It is also possible to have a daisy chain of proclitics preceding the host or enclitics following it. In either event, each clitic protects the vowel length of the clitic to its left which, as a consequence, escapes FVS:

- (21) a. two proclitics: *kikópó kyáá= wáá=* ‘the cup of the one of
Wálúsìmbi Walusimbi’
 c. three enclitics: *yákítééséé =múù =kòð =kí* ‘what did he put a little of in?’
 d. one of each *kyaa= musíkáá =wange* ‘the one of my heir’

It is tempting to hypothesise that FVS is a rule that applies at the end of a clitic group (CG) as depicted in (22).

- (22) $VV \rightarrow V / __]_{CG}$

Such temptation should be resisted. Rule (22) makes the right prediction in many but not all cases because it overlooks some of the subtleties of the process.

An adequate account must take on board the distinction between syntactic clitic and phonological clitic (cf. Klavans 1985). The two types of clitics are not identical.

(23) Syntactic clitic ≠ phonological clitic (cf. Klavans 1985)

- a. syntactic procliticization of object cleft/relative clause marker (≠ phonological proclitic; vs. (13b))
 - (i) *kikópò Walúsìmbi kye yalábà* ‘it’s a cup that Walusimbi saw’ /*ki-e/*
 - (ii) *embwà Walúsìmbi gye yalábà* ‘the dog that Walusimbi saw’ /*gi-e/*
- b. syntactic encliticization of emphatic pronouns (≠ phonological enclitic)
 - (i) *yalyá kô* ‘he ate IT’ (class 12)
 - cf. *yalyáá =kô* ‘he ate a little’

- In (23a), we observe that the object cleft and relative clause marker /-e/ is syntactically procliticized to the verb, hence translatable as “it’s a cup Walusimbi that he saw”. Despite this, /ki-e/ is realized as [kye], not as [kyee], which it would be, if it were a phonological proclitic. In the case of the syntactic enclitics in (23b), which must immediately follow the verb, the final length of the latter is not preserved, as it would be if *kô*, *yê* etc. were phonological enclitics. The same is true of other short particles such as the topic marker in (23c).

In verbs, the length of a final bimoraic (heteromorphemic) syllable which arises when the first of two vowels is glided or deleted, the second being lengthened in compensation, is preserved before an enclitic:

- As another source of length, there is a minimality requirement that words belonging to lexical categories must satisfy. Such words must contain at least two moras. So, if a verb has a monosyllabic stem, the one stem syllable must be bimoraic. As seen in (25a), this length is preserved before an enclitic:

- The sole exception to this, seen in (25b), is the copula *-li*, which can be analyzed as either non-lexical or as not being a stem.

The third source of final vowel length comes from the realization of contour tones (Ashton et al. 1954:424, 452; Tucker 1962:157; Cole 1967:67–68, 88; Stevick 1969:6; Hyman 1982:13). Setting aside an utterance-level downstepping phenomenon, Luganda has two surface tones, H(igh) and L(ow), which, conditions being met, can

combine to form a HL falling contour tone. (The language does not allow LH rising tones.) Within a word this HL contour (i.e. falling tone) must be realized on two separate moras. In word-final position, if a vowel bears a H tone that H tone is always realised as a HL (falling) contour tone.) Length from final HL contour length is preserved if HL is realized on the surface as in (26a), otherwise length from this source is not preserved as in (26b).

- (26) a. *te-bá-bálâ* 'they don't count'
te-bá-bálâà=kô 'they don't do much counting'
te-báá-gúlê 'they will not buy'
te-báá-gúlêè=kí 'what will they not buy?' (echo Q)
- b. *a-balâ* 'he who counts'
a-balá=kô 'he who counts a bit'
a-náá-bálâ 'he who will count'
a-náá-bálá=kí 'he who will count what?' (echo Q)

Turning to nouns in clitic groups we observe that FVS does not always apply quite in the same way as it applies to verbs. First, final bimoraic syllable (hetero- or monomorphemic) length in a noun is not preserved before an enclitic:

- (27) a. *ki-wábyò* 'sickle' *ki-wábyò=kí* 'which sickle?'
ki-wábyò=kyè 'his sickle'
- b. 'his sickle' 'deed' *kí-kòlwa=kí* 'which deed?'
/kol-u-a/ (do-passive-FV) *kí-kòlwa=kyè* 'his deed'
- c. *ku-lábwà* 'being seen' *ku-lábwà=kí* 'which being seen?'
 cf. *ku-lábwàà=kí* 'to be seen by what?'

By contrast, the vowel length of a monosyllabic stem is always maintained in order to ensure that the violation of the bimoraicity constraint on the structure of words belonging to lexical categories is averted.

- (28) a. *ki-de* 'bell' *?!-de(e)/* *ki-dee=kí* 'which bell?'
ki-déé=kyè 'his bell'
- b. *n-te* 'cow' *?!-te(e)/* *n-tee=kí* 'which cow?'
n-téé=yè 'his cow'

Finally, the length of the final vowel of a noun triggered by a final HL contour is preserved obligatorily if HL is realized and optionally if the HL contour is not realised.

- (29) a. *ki-síkí* 'log' *ki-sikî=kí* 'which log'
ki-síkí(í)=kyè 'his log'
- b. *mu-sotâ* 'snake' *mu-sotâà=kí* 'which snake'
mu-sotá(á)=gwè 'his snake'

The difference between FVS in nouns and verbal CGs is summarised in (30).

(30) Comparison of FVS in nominal and verbal CGs

Source of length	Verb=encl	Noun=encl
Underlying	long	short
Monosyllabic stem	long?	long
HL contour (realized)	long	long
HL contour (unrealized)	short	variable

We suggest that the internal FVS properties of Verb=encl vs. Noun=encl should be interpreted as follows. Speakers generalize the realization of the final vowel of a noun in isolation, where it is short, to situations where the noun occurs in a clitic group. As a result they obligatorily shorten the final vowel of a noun in a CG, e.g. *ki-wábyò* = *kí* ‘which sickle?’ or *mu-gwáágwá* = *kí* ‘which fool?’ Likewise, the length of a final vowel bearing a contour tone is optionally realised in a nominal CG, e.g. *ki-sikíí* = *kyè* ‘his log’ by analogy to the preservation of vowel length based on the contour of the isolation form *ki-sikí*.

However analogy is not invoked in the same way in the case of verbs in CGs and so speakers do not generalize what happens to final vowel length in bare verbs to final vowels in CGs. The reason for this might be the pressure exerted by verbal paradigms which are more pervasive and “tight” than nominal paradigms. The isolation forms of verbs are less complete in themselves than nominal forms occurring on their own.



Nominalisation provides an excellent context in which to observe the asymmetry between nouns and verbs with respect to FVS. FVS will apply as it does to nouns if a verb is nominalized (including in a phrase word) as (31) shows:

- (31) a. *kí-kòlwa* ‘deed’
kí-kòlwa = *kí* ‘which deed?’
 < /*kí-kól-u-a*/ ‘it is done’
 cf. *kí-kòlwaa* = *kí* ‘it is done by what?’
- b. *muzáddè - t-á-lyà* ‘(proper name)’
 (lit. a parent doesn’t eat)
muzáddè - t-á-lyà = *kí* ‘which M-T.?’
muzáddè - t-á-lyà = *wè* ‘his M-T.’ (*...*lyàà* = *wè*)
- c. *mugenzi - tázzè* ‘delinquent debtor’
 (lit. traveler hasn’t returned)
mugenzi - tázzèè = *kí* ‘which delinquent debtor?’
mugenzi - tázzé(é) = *wè* ‘his delinquent debtor’
- d. *nnámpá - wè - n-gwá* ‘a neutral person’
 (lit. there is nowhere that I fall)
nnámpá - wè - n-gwá = *wè* ‘his N.’ (?...*n-gwáá* = *wè*)
 ~ *nnámpá - wé - n-gwá* = *wè* (= with H tone plateauing)

Finally, as the final element of a compound or phrase word, monosyllabic stems undergo FVS optionally if the phrase word is a common noun and obligatorily if it is a proper name, as seen in (32).

- (32) a. *mutunda - bide* ‘bell-seller’ *mutunda - bide(e)*=*ki* ‘which...’
cf. *bi-dee*=*ki* ‘which bells’
mutunda - bidé(é)=*wè* ‘his...’
cf. *bi-dée*=*býè* ‘his bells’
- b. *akisá – ènte* (proper name) *akisá - ènte*=*ki* ‘which A-E.?’
(lit. he who hides cows) cf. *n-tee*=*ki* ‘which cow’
akisá - ènté=*wè* ‘his A-E.’
cf. *n-tég*=*wè* ‘his cow’

A possible interpretation of the variation in the realisation of vowel length is that speakers are unclear in (32a) whether the enclitic cliticises to preceding PW (N₃), as in (33a), or to the whole syntactic word X^o (N₁), as in (33b).

- (33) a. 
mutunda - bidee
- b. 
mutunda - bide

We have shown in this section the final vowel length *per se* is not a reliable indicator of word status. In the next section we turn to another indicator, that has been proposed, namely tone, and see whether it fares any better.

4. Tone as a criterion for the word

We have already referred to Heny (1974) who proposes the tonal criterion of one pitch drop per lexical item to characterise the tone word (TW). In (34) a fuller set of tonal criteria for word determining word status is provided:

- (34) a. at most one HL pitch drop per TW
b. mapping of phrasal %LH% boundary tones to toneless TWs
c. function of TW in H tone plateauing (HTP)
d. one overall “tonal configuration” (e.g. in verb reduplications – see (53))

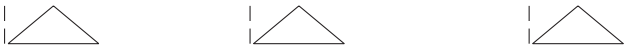
In the following paragraphs we elaborate on these criteria.

Morphological words (MWs) can have one drop from H to L, as in (35a), or none (in which case they are lexically toneless), as in (35b).

- (35) a. *ki-bê* jackal *ki-sásilo* rubbish *ki-jũko* spoon
ki-kópò cup *ki-yulĩfũ* torn (cl.7) *ki-bíínà* society
ki-siki log *ki-wójjólò* butterfly *ki-wúúgúlũ* owl
kilòkwà weed *ki-bónèlezo* punishment *ki-sàànikizo* cover, lid
ki-kòlwa deed *ki-begábèga* shoulder *ki-sáákáátè* reed
b. *ki-de* bell *ki-tabo* book *ki-tooke* plantain
ki-bya bowl *ki-muli* flower *ki-seenge* room
ki-lagiro command *ki-biiliti* match(-box) *ki-sanilizo* comb
ki-papajjo branch *kin-njaalo* bean *ki-sumuluzo* key

MWs are marked for the tones with which they exit the lexical (word-level) phonology: (á) = H, (à) = L, (a) = toneless. Toneless lexical items may acquire tone postlexically.

Multiword forms are marked with the tones they carry after the application of H tone plateauing (HTP) e.g. in (41) below. Toneless moras acquire a H or L by operation of rules at the phrase level. A major rule applying at the phrasal ensures that at the left edge of a phrase, a toneless word is realized L-H_n (%L on the first mora and H% on remaining moras), as seen in (36).

- (36) a. *kì-dé* bell *kì-tábá* book *kì-tóóké* plantain
kì-byá bowl *kì-múli* flower *kì-sééngé* room
b. *kì-lágíró* command *kì-bíilíté* match(-box) *kì-sánílító* comb
kì-pápájjó branch *kì-jánjááló* bean *kì-súmúlúzó* key
c. %L H% %L H% %L H%

kì-pápájjó branch *kì-jánjááló* bean *kì-súmúlúzó* key

By this criterion, the augment (a.k.a. initial vowel), e-, in (37) is a prefix for when it appears in a lexically toneless word, a low tone falls on the initial syllable and any remaining syllables receive high tone.

- (37) a. *è-kí-dé* bell *è-kí-tábó* book *è-kí-tóóké* plantain
è-kí-kbyá bowl *è-kí-múli* flower *è-kí-sééngé* room
b. *è-kí-lágíró* command *è-kí-bíilítí* match *è-kí-sánílízó* comb
è-kí-pápájjó branch *èkí-jánjááló* bean *è-kí-súmúlúzó* key

By the same criterion, the proposed elements in (38) are proclitics since we do not get the pattern of L tone on the first syllable and H on the rest where the word belonging to a lexical category is lexically toneless.⁷

- (38) a. class 2a *ba*=
bà=kàtóná gods *bà=Mùkásá* the Mukasas
b. locatives *ku*= (class 17) and *mu*= (class 18)
kù=kì-dé on the bell *mù=kì-dé* in the bell
kù=kì-tábó on the book *mù=kì-tábó* in the book
kù=kì-sánízíó on the comb *mù=kì-sánízíó* in the comb

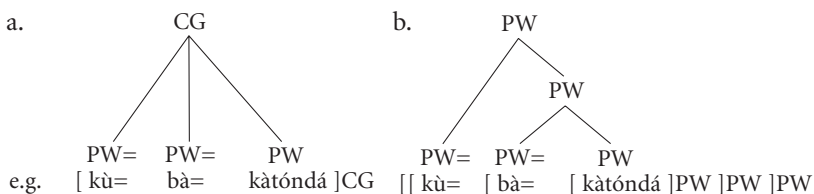
- c. *na* ‘with’, ‘and’ (comitative, instrumental, associative)
nà=kì-dé with a bell *nà=Kàtòndà* with God
nà=kitábó with a book *nà=Mùkásá* with Mukasa
nà=kì-sánízíó with a comb
- d. genitive linker *-aa*; here: *byaa*= ‘those of (class 8)’
byàà=mù-ntú those of a person *byàà=Kàtòndó* those of God
byàà=mù-límí those of a farmer *byàà=Mùkásá* those of Mukasa
byàà=mù-lámútí those of a judge

An initial string of proclitics will all remain L (plus one more L on the first syllable of the noun), as seen in (39), indicating again that we are not dealing with a single lexically toneless MW.

- (39) a. *kù=bà=katóndá* on the gods
nà=bà=kàtòndá with the gods
mù=bà=kàtòndá in the gods
byàà=bà=kàtòndá those of the gods
- b. *nà=kù=bà=kàtòndá* and on the gods
nà=byàà=bà=kàtòndá with those of the gods
byàa=kù=bà=kàtòndá those on the gods
byàa=wàà=bà=kàtòndá those of the one of the gods
- c. *nà=byàa=kù=bà=kàtòndá* with those on the gods
nà=byàa=wàà=bà=kàtòndá with those of the one of the gods

Two reasonable representations can be proposed to account for one or more proclitics combining in a CG with toneless MWs:

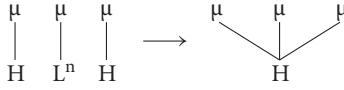
- (40) Two reasonable proposed structures to account for proclitic(s) + toneless MWs



We opt for (40b) which reflects the cyclic nature of the attachment of proclitics in (39).

We shall now address the question of whether “at most one HL pitch drop” is an adequate criterion for identifying TWs. The answer will be in the negative. This is because of the extensive use of the rule of High Tone Plateauing (HTP) depicted in (41), whose effect is exemplified in (42).

(41) High Tone Plateauing (HTP)



(42)

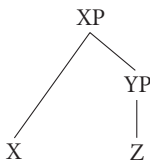
a. Noun + possessive

<i>ki-bê</i>	+ <i>kyaa=Walúsimbi</i>	→ <i>ki-bé</i>	<i>kyáá=Wálúsimbi</i>	Walusimbi's jackal
<i>ki-kópò</i>	+ <i>kyaa=Walúsimbi</i>	→ <i>ki-kópó</i>	<i>kyáá=Wálúsimbi</i>	Walusimbi's cup
<i>ki-sikí</i>	+ <i>kyaa=Walúsimbi</i>	→ <i>ki-sikí</i>	<i>kyáá=Wálúsimbi</i>	Walusimbi's log
<i>ki-lókwà</i>	+ <i>kyaa=Walúsimbi</i>	→ <i>ki-lókwá</i>	<i>kyáá=Wálúsimbi</i>	Walusimbi's weed
<i>ki-kòlwa</i>	+ <i>kyaa=Walúsimbi</i>	→ <i>kí-kólwá</i>	<i>kyáá=Wálúsimbi</i>	Walusimbi's deed

b. (Affirmative) verb + following "word"

<i>twáá-làbà</i>	+ <i>Walúsimbi</i>	→ <i>twáá-lábá</i>	<i>Wálúsimbi</i>	we saw Walusimbe
<i>twáá-làbwà</i>	+ <i>Walúsimbi</i>	→ <i>twáá-lébwá</i>	<i>Wálúsimbi</i>	we were seen by Walusimbi
<i>twáá-gèndà</i>	+ <i>tútùtu</i>	→ <i>twáá-géndá</i>	<i>tútùtu</i>	we went slowly
<i>twáá-gèndà</i>	+ <i>lulí</i>	→ <i>twáá-géndá</i>	<i>lulí</i>	we went day before yesterday
<i>twáá-làbà</i>	+ <i>ki-kópò</i>	→ <i>twáá-lábá</i>	<i>kí-kópò</i>	we saw a big cup
<i>kinénè</i>				
H L L	H L	HL	H	H <u>L</u> H L

What the data in (42) show is that there is "at most one HL pitch drop" per tone group (TG). As seen in (42a), a N + gen=N constitutes a single TG. Similarly, in (42b), in most affirmative tenses, a verb will form a TG with the PW that follows it. We propose to define the TG as shown in (43).

(43) Definition of the Tone Group (TG): X + Z (see Hyman & Katamba 1990/91, 1993, a&b)

- where: (i) X ≠ negative, imperative or infinitive of verb
(ii) Z = PW
(iii) Z does not begin with an augment

The role of syntactic information in defining TGs is crucial. As well as requiring the presence of appropriate tonal properties for a string to count as a tone group, it must also meet the syntactic characteristics specified in (43). The postlexical application of HTP is restricted to a TG. Some constraints on its application will now be considered as a way of illustrating TGs.

As seen in the following examples, HTP may also apply to compounds, whether lexicalized, as in (44a), or productively created, as in (44b).

- (44) a. *mwásà + jjútè? mwásá - jjútè* ‘hard chair’ (lit. boil-breaker)
 H L H L H-----H L
 b. *mu-témà + bi-síkí mu-témá - bísíkí* ‘log-chopper’ (chopper + logs)
 H L HL H-----HL

We now show that HTP is a domain-juncture rule (Selkirk 1984, Nespor & Vogel 1986). Plateauing is not permitted where a toneless word is intercalated between words with the requisite H tones as seen in (45).

- (45) a. *twáá-làbà ki-tabo kyaa= Walúsìmbi* ‘we saw Walusimbi’s book’
 [H L L] [] [H L]
 b. *twáá-làbà mugenzi - tázzê* ‘we saw the delinquent debtor’
 [H L L] [] [H HL]
 c. *mukúbà - balimi waa= Walúsìmbi* ‘Walusimbi’s farmer-beater’
 [H L] [] [H L]

Our interpretation is that the intervening toneless forms in (45) are PWs, which are visible to HTP. Now compare comparable situations that arise when the intervening form is a clitic. As seen in (46), HTP is obligatory when the toneless words are proclitics which are not visible to HTP.

- (46) HTP is, however, obligatory when the toneless “words” are proclitics (which are not visible to HTP -- which seems to argue for the structure in (32b))
- a. *twáá-làbà + byaa= Walúsìmbi ? twáá-lábá byáá=Walúsìmbi*
 H L L H L H-----H L
 ‘we saw Walusimbi’s’
- b. *twáá-gèndà + na= Walúsìmbi ? twáá-géndá ná= Wálúsìmbi*
 H L L H L H-----H L
 ‘we went with Walusimbi’
- c. *twáá-géndá + ná= byáá= kú= bá=Wálúsìmbi*
 H-----H L
 ‘we went with those on the Walusimbis’

Our conclusion is that such clitics are not PWs, rather join their host to define a complex PW. This fact, plus the mapping of left-boundary %L onto each proclitic in (39), seems to argue for the branching structure in (40b), though with the proclitics not identified as PWs.

Returning to the syntactic conditions, we find that HTP will not apply where X does not c-command Z. Thus, for example, because a subject and the verb that follows it are not in a c-command relationship, a subject-verb sequence (including phrase-words as in (47b)) will never constitute a TG. Hence they are not a viable domain for HTP:

- (49) a. one TG, one CG: *tú-lyáá=kô* ‘we eat a little’ (cf. *tú-lyà* ‘we eat’)
H-----HL (“Z” = enclitic) H L
b. two TGs, two CGs: *te-tú-ly-à mu-púùnga* ‘we don’t eat rice’
H L HL
(two TGs because verb = negative)

In (49a), the affirmative verb *tú-lyàà* + following enclitic *=kô* constitutes a single TG with the enclitic, and HTP applies. The two elements also form a single CG, within which FVS thus cannot apply. The vowel of the verb thus remains long. In (49b), FVS applies, since the post-verbal element is not an enclitic, and HTP fails to apply, because the verb is negative. We thus have two TGs and two CGs.

These two situations contrast with the two in (50), where there is a mismatch, or non-isomorphism, between TGs and CGs:

- (50) a. one TG, two CGs: *tú-lyá mú-púunga* ‘we eat rice’
H-----HL (two CGs because ‘rice’ ≠ enclitic)
b. two TGs, one CG: *te-tú-ly-àà=kô* ‘we don’t eat any’
H L HL
(two TGs because verb = negative)

In (50a) we have one TG, since the verb is affirmative, and HTP applies. However, since the verb is followed by a noun, rather than by an enclitic, FVS applies, indicating that the sequence consists of two CGs. Just the reverse situation obtains in (50b), where there are two TGs, because the verb is negative, but one CG, because the verb is followed by the enclitic =*kô*. Such intersection is typical in Luganda, e.g. a CG can function as a single tone word (TW), one layer down from the TG, as can two PWs. This is exemplified in (51).

- (51) a. *ki-tabo + kyângè* → *ki-tabó = kyângè* ‘my book’ cf. *kyângè* ‘mine’ (class 7)
ki-lagiro + kyângè → *ki-lagíró = kyângè* ‘my command’
ki-sumuluzo + kyângè → *ki-sumúlüzò = kyângè* ‘my key’
 b. *η-ηumya + mu-twê* → *ηηumyá – mùtwê* ‘hard, solid person’
 (-*gumya* ‘make solid’ + head)
ka-mwa + ka-bî → *kamwá – kàbì* (personal name) (lit. small mouth + bad)
o-mu-ntu + mu-lamú → *omúntù – mùlà mù* ‘person of noble character’
 (-*lamú* ‘healthy’)
g-gulu + d-dénè → *ggulú – ddènè* ‘elephant’ (lit. big foot + big)

While a full discussion would take us very far afield, the essential characteristic of a complex TW is that the H of a second stem or clitic is mapped onto the first stem. In the input in (51a), for example, we see that the 1sg possessive enclitic has a H on the first mora. In class 7, it would be pronounced [kyààngè] ‘mine’ when there is a null nominal head. However, the output shows that the H of the enclitic is reassigned to the second mora of the preceding noun stem (followed by all Ls). In the compounds

in (51b), we have noun+noun and noun+adjective inputs, i.e. two words, as can be seen from the FVS of N-Numya and ka-mwa in the first two examples. In this case, the H from the second word is reassigned to the second mora of the stem of the first word.⁸

The Table in (52) summarizes our findings, showing how these different “words” and “groups” intersect.

(52)

	TW	PW	TG	CG
proclitic + noun	1	1	1	1
affirm.verb	2	1	1	1
(i) noun=poss.pronoun	1	2	1	1
(ii) stem reduplication				
neg.verb= <i>kô</i> (~ <i>mû</i>)	2	1	2	1
irreg. noun compounds (51b)	1	2	1	2
adjective= <i>kô</i> (~ <i>mû</i>)	2	2	1 ~ 2	1
noun= <i>kí</i> ‘which’	2	2	2	1
(i) affirm. verb + noun	2	2	1	2
(ii) noun + poss.noun				
(iii) noun compounds (44)				
other “word” + “word”	2	2	2	2

Intersecting “Words” and “Groups” in Luganda

Reduplication is the final arena of the intersection of phonological domains that we will consider. It is also an area where often the application of morphological processes is circumscribed by prosodic parameters. Reduplication is highly productive in Luganda. Here we restrict ourselves to verbal reduplication, which is total.⁹ For noun-and adjective reduplication, see Hyman & Katamba (1990).

Verbal reduplication has a variety of uses including signalling an action done frequently, or ‘here and there’ or without real commitment. Here it is the phonological properties of reduplication which have a bearing on word recognition criteria that will be treated. We will begin by considering the realisation of word-final underlying length. As seen in (53), untypically of length from this source, in a verb the underlying length of the base (first part) of a reduplicated form is not preserved. By contrast, the expected length survives in the reduplicant if it is followed by an enclitic.

- (53) a. *ku-sasulwa-sasulwaa* =*kô* ‘to be paid a bit’ < *ku-sasulwa* ‘to be paid’
 b. *ku-wúlirwa-wulirwaa* =*kô* ‘to be heard a bit’ < *ku-wúlirwa* ‘to be heard’

Where the final length is induced by a contour tone that is realised on the surfaces as in the reduplicant in (54a), length is saved before an enclitic. But where the final HL contour fails to surface, the length associated with it also perishes with it as in (54b).

- (54) a. *mu-wulile-wulilèè* =kô '(you pl.) hear a bit!' < *mu-wulilè* '(you pl.) hear!'
mu-labilile-labililèè =kô '(you) look after a bit!' < *mu-labililè* '(you) look after!'
 b. *a-wúlílá-wúlílá* =kô 'he who hears a bit' < *a-wúlílâ* 'he who hears'
a-lábílílá-lábílílá =kô 'he who looks after a bit' < *a-lábílílâ* 'he who looks after'

As seen in (55), monosyllabic length is preserved in both parts of the reduplicated verb. This seems to be motivated by the requirement for words belonging lexical category to have stems that are at least bimoraic:

- (55) a. *ku-mwaa-mwaa* =kô 'to shave a bit' < *ku-mwa* 'to shave'
 b. *ku-lyáá-lyàà* =kô 'to eat a bit' < *ku-lyâ* 'to eat'

Bimoraic CVCV stems require an iambic base and hence become CVCVV in reduplication. Thus, the FV of the base is lengthened in (56a) to assure the iambic structure. Where there is underlying length as in (56b), it is also preserved, but not because of the input, here /bal-u-a/, but because of the iambic condition on bisyllabic bases.

- (56) a. *ku-balaa-bala* =kô 'to count a bit' < *ku-bala* 'to count'
 b. *ku-balwaa-balwa* =kô 'to be counted a bit' < *ku-balwa* 'to be counted'

Finally, let us consider the forms in (57), which consist of non-tonic compounds followed by a possessive enclitic:

- (57) a. *mu-gemera + wala* 'gun' *mu-gemera - walá* =gwàngè 'my gun'
 (lit. preventer (from) far) ~ *mu-gemèrà - wàlà* =gwàngè
 b. *mu-tunda + bitabo* 'book-seller' *mutunda - bitabó* =wàngè 'my book-seller'
 (lit. seller + book) ~ *mutúndà - bitàbò* =wàngè
 c. *n-tabaaaza + bakadde* 'beer' *ntabaaza - bakááddè* =yàngè 'my beer'
 (lit. I make old people go to war) ~ *ntabáázà - bàkàddè* =yàngè

As in the forms in (51), the possessive 'my' shifts its H to the second mora of the stem in the preceding compound. But which stem? There are two in each case: the stem of the immediately preceding word, or the stem of the first word of the compound. While the first form is preferred in each case, i.e. where the H goes on the second mora of the preceding stem, speakers also frequently accept to place the H on the second mora of the first stem, as indicated.

Our interpretation of this variation is shown in (58).

- (58) a.
$$\begin{array}{c} N_1 \\ \swarrow \quad \searrow \\ N_2 \quad N_3 = \text{gwàngè} \\ \text{mugemera} - \text{walá} \end{array}$$
 b.
$$\begin{array}{c} N_1 = \text{gwàngè} \\ \swarrow \quad \searrow \\ N_2 \quad N_3 \\ \text{mugemèrà} - \text{wàlà} \end{array}$$

In our view, speakers are unclear in (57a) whether the enclitic cliticizes to the preceding PW (N_3), as in (58a), or to the Xo (N_1) as a whole, as in (58b). The difference is expressed via the landing sight of the H of the possessive enclitic.

5. Conclusion

Let us recall the workshop questions that were posed at the beginning of §1.

- a. Can the word be defined?
- b. If not, why not?
- c. If yes, is the word a universal?

When the Luganda facts are considered in the context of these questions, the answers that emerge are not simple. It is obvious that the word cannot be uniquely defined in any of the senses commonly recognised by linguists using a consistent set of criteria. Why should that be so? In our view, this is because the different criteria conflict in at least two ways:

First, there are conflicts between the different components of the word (morphology, syntax, phonology).

Second, there are conflicts even within the same component, e.g. with regard to phonological criteria vowel length conflicts with tone.; tone also conflicts with itself.

Although intuitively it makes sense to recognise words as key building blocks of language, finding consistent and reliable ways of characterising words in one language, let alone cross-linguistically is Sisyphean task. The linguist's desire to categorise and compartmentalize and to label entities neatly is frustrated because words are amorphous entities in the sense that probably there is no one point where all relevant information about a word is packaged together. While the result at any one time is a partial view of a changing situation, one needn't fret. The excitement of such a study as the one in which we have been engaged in Luganda for a number of years has taught us much about how the different concerns of a language interact and ultimately conflict.

In producing this study, we realized how central the word has been in our previous work on the morphology-phonology and syntax-phonology interfaces in Luganda. While we cannot define, we can delimit, and we can also advance hypotheses as to why these complexities and contradictions exist. We suspect that rather than being different from other aspects of language, the problems we have set out could be applied to most any aspect of language: the sentence, the syllable, etc. Even the morpheme and the phoneme, which are presented as if easily defined, are not exempt from definitional and analytical problems. *Raison de plus* to keep at it.

Notes

1. In Guthrie's widely used classification system of Bantu languages, it is classified as Zone E.15, modified by Tervuren to J.15.
2. In these and other examples, an acute accent (*á*) marks high (H) tone, a grave accent (*à*) marks low (L) tone, and a circumflex (*â*) a HL falling tone. Vowels lacking a tone mark are toneless and receive a H or L tone according to the phrasal tonology. Long vowels are transcribed as double throughout this study. The symbol (=) separates proclitics and enclitics from their host. Hyphens, when present, mark morpheme boundaries, though not all internal morphology is marked in the examples.
3. It is likely that the clitic combinations in (12) are formed postlexically, i.e. at the syntactic level. Since they escape the lexical morphology, it could be said that they are not morphological words. We would find it difficult to draw the same conclusion concerning the demonstratives in (11), however.
4. The fact that the vowel of the proclitic is always long unless it is followed by a word whose first consonant is a geminate is not reflected in the standard orthography. Cf. §3.
5. Note that the cliticisation does not induce lengthening; it only helps preserves it where is already has some reason for being present present. Observe the short vowels of *ba=* and *ku=* in (13a), for example.
6. Although not relevant for our present purpose, for evidence that some vocalic morphemes should be analyzed as VV and others as V, see Hyman & Katamba (1990).
7. We have two other tonal criteria which yield the same result: (i) tone retraction (*ki-sikí* vs. *e-ki-síkí* 'log', with retraction, vs. *na= kisikí* 'with a log', without retraction); (ii) Meeussen's Rule (H-H ? H-L), e.g. *a-láb-à* 'he sees' vs. *tú-làb-à* 'we see'; cf. *né= Kígúndu* 'with Kígúndu', where *né=* is the [+augment] form of *na=*.
8. In the case of *omúntù - mùlà mù*, pronounced [*òmúúntù - mùlà mù*], the H is assigned to second mora of [*muú*], then spreads to both moras to avoid a LH rising tone.
9. It is hence difficult to determine which is the base and which is the reduplicant. We will assume a base-reduplicant structure.

References

- Ashton, E. O., Mulira, E. M. K., Ndawula, G. G. M., & Tucker, A. N. (1954). *A Luganda Grammar*. London: Longmans, Green and Co.
- Bloomfield, L. (1933). *Language*. New York NY: Holt.
- Cole, D. T. (1967). *Some Features of Ganda Linguistic Structure*. Johannesburg: Witwatersrand University Press.
- Dixon, R. M. W., & A. Y. Aikhenvald (2002). Word: A typological framework. In R. M. W. Dixon & A. Y. Aikhenvald (Eds.), *Word: A cross-linguistic typology*. Cambridge: Cambridge University Press.
- Heny, F. (1974). Ganda tone as 'pitch accent'. Ms. University of Massachusetts, Amherst.
- Hyman, L. M. (1982). Globality and the accentual analysis of Luganda tone. *Journal of Linguistic Research*, 1, 1–40.

- . (1988). Direct vs. indirect syntactic conditioning of phonological rules. *Proceedings of E.S.C.O.L. 1987* (pp. 147–163). Ohio State University.
- & F. X. Katamba (1990). Final vowel shortening in Luganda. *Studies in African Linguistics*, 21, 1–59.
- . (1990–91). The augment in Luganda tonology. *Journal of African Languages and Linguistics*, 12, 1–45.
- . (1993a). A new approach to tone in Luganda. *Language*, 69, 34–67.
- . (1993b). The augment in Luganda: Syntax or pragmatics? In S. Mchombo (Ed.), *Theoretical Aspects of Bantu Grammar* (pp. 209–256). Stanford CA: CSLI.
- & L. Walusimbi (1987). Luganda and the strict layer hypothesis. *Phonology (Yearbook)*, 4, 87–108.
- Klavans, J. L. (1985). The independence of syntax and phonology in cliticization. *Language*, 61, 95–120.
- Nespor, M., & I. Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.
- Poser, W. J. (1992). Blocking of phrasal constructions by lexical items. In I. A. Sag & A. Szabolsci (Eds.), *Lexical matters* (pp. 111–130). Stanford CA: CSLI.
- Selkirk, E. O. (1984). *Phonology and Syntax: The relation between sound and structure*. Cambridge MA: The MIT Press.
- Tucker, A. N. (1962). The syllable in Luganda: A prosodic account. *Journal of African Languages*, 1, 122–166.
- Stevick, E. W. (1969). Pitch and duration in Ganda. *Journal of African Languages*, 8, 1–28.
- Zwicky, A. M. (1990). Syntactic words and morphological words, simple and composite. *Yearbook of Morphology*, 3, 201–216.

Case in Africa

On categorial misbehavior

Christa König

Johann Wolfgang Goethe Universität, Frankfurt

In the present paper I will propose a fairly far reaching explanation for the categorial misbehavior of case markings found in Ik.¹ While it will seem that the kind of data here considered is somewhat unusual it will be shown, based on a general consideration of case, that it is to be expected, considering the development of the structures concerned.

1. Introduction

Case as a grammaticalized category is rather a rare phenomenon in Africa. It occurs only in two of the four language families: Nilo-Saharan and Afroasiatic, although within Niger Congo at least some West-Bantu languages (Blanchon 1988) are claimed to have a case distinction expressed by tone. Khoisan has never been mentioned in this connection. In East Africa there is an abundance of languages with case inflexions. Case may, therefore, be an areal phenomenon. Genetically, the subfamilies Cushitic, Semitic, Chadic and Nilotic are mainly concerned.

World-wide, case systems are distinguished with regard to the behavior of the three core participants, which are subject of an intransitive sentence (S), subject of a transitive sentence (A), and object of an transitive sentence (O). If S and A are treated the same and simultaneously different from O, Dixon (1994) speaks of an accusative system. If S and O are treated the same and simultaneously different from A, Dixon speaks of an ergative system. Within Africa the accusative system has the broadest occurrence, especially in East Africa again.

Ergativity is fairly rare in Africa. Only for the West Nilotic languages Pări and Shilluk an ergative case marker is well described (Andersen 1988; Miller & Gilley 2001). Traces of ergativity, that is not as a grammaticalized case system but rather with regard to word order, cross reference, verbal plural, are described again for West Nilotic languages such as Anywa (Reh 1996), Chadic languages such as Mandara (Frajzyngier 1984), Tangale and Hausa (Böhm 1983), and Loma (Rude 1983), a Mande language.

Interestingly, the only documented African instance of a flexional split-ergative system, Pāri, is typologically exceptional: Following Mallinson & Blake (1981:123) in ergative systems a constituent order where O precedes A is not allowed. But this is what can be observed in Pāri, as it has the rare OVA order.

The East African language region is known for another peculiarity which is world-wide nearly unique: Within a proper accusative language the nominative is morphologically the unmarked case, is used in citation while the accusative is the marked case. In East Africa there are a number of languages which on the one hand show an accusative pattern with regard to S, A and O, as S and A are treated the same and simultaneously different from O. On the other hand there is the accusative, traditionally called absolutive in the literature, morphologically the most unmarked case, used in citation and the nominative is morphologically the marked case. This system is called marked nominative (Dixon 1994). Languages with a marked nominative system are for instance Sidamo, Ometo, Oromo, Somali, Rendille, Dasenech, Kemantney, Gidole, Beja, Arbore of Cushitic, Kalenjin, Maa, Teso, Turkana, of Nilotic and Diding'a, Murle of Surmic.

A further case type called active system (Klimov 1974) is rarely mentioned in the African literature: Namely the Soharam languages Beria, and Tedadaza, some Berber languages, and Loma, a Mande language. According to Rude (1983), Loma has a split-active system with regard to cross reference. In an active system, S is not marked coherently, with some verbs it is marked in the same way as A, and with others in the same way as O (see König in print).

1.1 The notion categorial misbehavior

Prototypically certain grammatical features do occur in languages only with certain word classes. Nouns can be pluralized, can be modified for instance by numerals or adjectives, can take determiners like demonstratives, can take genitive attributes, can be inflected for case, etc. Verbs, on the other hand, are associated with tense, aspect and modality, can be negated, etc.

But there are languages where a given word class shows features which typically are associated with another word class. Such cases are referred to by Gerrit Dimmen-daal (p.c.) as categorial misbehavior. In this paper I want to deal with one instance of categorial misbehavior that does not seem to have received much attention in the linguistic literature.

The language to be considered is Ik, a language with an elaborated case and a split-accusative system. Ik distinguishes seven cases; the nouns are marked by suffixes. Each case marker occurs in two different forms: One is called the final form and the other the non-final form. The final form is basically used at the end of sentences or phrases and the non-final form elsewhere. The forms are given in Table 1.

A further typological characteristic of Ik is the presence of voiceless vowels. They are indicated by raised vowel symbols in the examples. Ik is a VSO-language. And, Ik

Table 1. The case inflexions of Ik

Case	Abbreviation	Final	Non-final
Nominative	NOM	*V-[a]	*V-[a]
Accusative	ACC	<i>n</i>	<i>[-a]</i>
Dative	DAT	<i>-ke</i>	<i>-e</i>
Genitive	GEN	<i>-e (-i)</i>	<i>-e</i>
Ablative	ABL	<i>-o (-u)</i>	<i>-o</i>
Copulative	COP	<i>-ko</i>	<i>-o</i>
Oblique	OBL	\emptyset	\emptyset

is a tone language with two tone levels, the high tone is marked by an acute accent, the low tone remains unmarked.

Normally case is a category which is associated with nouns or pronouns. As will be shown below, case in Ik is not restricted to these word classes. Function words like conjunctions, postpositions, prepositions, adverbs and even verbs are inflected for case as well.

2. Data

2.1 Case on conjunctions

In Ik there are five elements documented which all serve as conjunctions and which are all case inflected. These are listed under Table 2. For the most part they are used as complementizers.

In the present I will limit myself mainly to a discussion of one of them, namely *kɔɔbáa*.

In examples (1) and (2) the element *kɔɔbáa* functions as a conjunction translated by 'what', in examples (3) and (4) as a conjunction 'why'. In (5), *tóimen* functions as a conjunction with the meaning 'that' and in (6) *na* occurs as a conjunction meaning 'where'.

Looking at these examples some observations are called for: The conjunction *kɔɔbáa* occurs in (1) in the nominative case but in (2) in the accusative case (*kɔɔbádí-a*). In (3), *kɔɔbáa* occurs again in the nominative, but at the end of the subordinate clause there is an additional dummy pronoun *d^e*. In (4) *kɔɔbáa* occurs in the ac-

Table 2. Ik Conjunctions

Conjunction	Meaning
<i>tómeda(na)</i>	'where'
<i>mená</i>	'what'
<i>kɔɔbáa</i>	'what'
<i>na</i>	'where'
<i>tóimen</i>	'that'

cusative, marked again with the additional dummy pronoun at the end of the subordinate clause. In (5) *tóimen* occurs in the dative case (*tóimen -k^e*) and in (6) *na* in the ablative case (*n-ée*), marked again with the dummy pronoun.

In (1) to (6) the conjunctions are evidently case inflected, the following cases being used: Nominative, accusative, dative and ablative.

- (1) *ńtá ye-í-í kɔɔbá-a itiyá-id^a*
 NEG know-1.SG-NEG what-NOM do-2.SG-a
 ‘I don’t know what you do.’
- (2) *ńtá ye-atⁱ kɔɔbádí-a itiy-at^a*
 NEG know-3.PL-NEG what-ACC do-3.PL-a
 ‘They don’t know what they do.’
- (3) *ńtá ye-í-í kɔɔbá-a im-á ǵɔd-á d^e*
 NEG know-I-NEG what-NOM child-NOM cry-a DP
 ‘I don’t know why the child cries.’
- (4) *ńtá ye-í kɔɔbádí-a im-á ǵɔd-á d^e*
 NEG know-NEG what-ACC child-NOM cry-a DP
 ‘He don’t know why the child cries.’
- (5) *itét-í-a ńa tóimení-k^e ńǵ-a nyéǵa bi-k^a*
 notice-1.SG-a ENC.SG that-DAT eat-a hunger-NOM you-ACC
 ‘I noticed that you felt hungry (Lit: hunger ate you).’
- (6) *moo ńts-a it-ét-í óropoi-é n-ée no ats-á*
 NEG he-NOM reach-VEN-NEG Oropoi-DAT where-ABL ENC.PAST come-a
d^e
 DP
 ‘He could not return to Oropoi where he came from.’

These facts raise a number of questions: (a) Why are conjunctions inflected for case? (b) Why does the subordinate clause need a dummy pronoun in some cases, but not in others? (c) How can one explain that in the nearly identical examples (1) and (2) *kɔɔbáa* does not occur in the same case, in (1) in the nominative and in (2) in the accusative?

One explanation for the aberrant behavior of conjunctions may be found in the source of the conjunctions. In all cases here documented, conjunctions can be shown to be derived from full nouns. The relevant sources are indicated for each case in Table 7. Not for all case inflected conjunctions are the nouns still used actively. *tóimen* has the original meaning ‘problem’ and *na* the original meaning ‘place’. For others, like *tómeda*, the original meaning is unknown; still, a nominal source is very likely. Others again, like *mena* and *kɔɔbáa* are still used as nouns, both meaning ‘thing’.

I will illustrate the grammaticalization process of the case inflected conjunctions by means of *kɔɔbáa*. At stage 0 of the grammaticalization, *kɔɔbáa* functions as a noun in an object position. *kɔɔbáa* is marked by the case which is obligatory for nouns used in this slot.

Now I have to describe a special behavior of the case system in Ik. Ik shows no homogeneous accusative system. It rather has a split system in the sense that the core participants, such as subject and object, either show an accusative pattern or no distinction at all: The encoding of the core participants is characterized by all kinds of irregularities. It is sensitive to factors like word order, syntactic construction, person of the subject, etc. In the clause type discussed here the accusative encodes only objects when the subject does refer to the third person. All other objects are encoded in the nominative case. In (König 2002) I have called this phenomenon ‘case anomaly’.

I will illustrate this case anomaly by examples (7) and (8). In (7) the object, girl, occurs in the accusative case with a subject referring to the third person. This is what we would expect from a proper accusative system. In (8) the object, children, is instead encoded in the nominative case with a subject referring to the first person. In (8) both core participants subject and object are encoded identical in the nominative case. The accusative system is neutralized in this environment.

- (7) *en-ugót-á ím-a nyárama-k^a*
 see-AND-a child-NOM girl-ACC
 ‘The child sees the girl.’
- (8) *en-és-isín-a njín^a wík^a*
 see-IRR-1.PL.IC-a we.IC-NOM children-NOM
 ‘We (incl.) will see the children.’

This irregular object behavior illustrated in the above examples (7) and (8) is identical with the behavior of the conjunction *kóróbáa*: In examples (1) and (3) *kóróbáa* occurs in the nominative case with a subject referring to the first person, in examples (2) and (4) *kóróbáa* occurs in the accusative case with a subject referring to the third person. Therefore in this respect *kóróbáa* behaves like any other object in Ik. The dative case of *tóimen* in (5) is also motivated by the main verb. *ité-es* requires an object in the dative case when having the reading ‘notice’, and it requires the accusative to encode the locative participant in the reading ‘to reach’.

For the grammaticalization process the following scenario can be established: At stage I the noun in object position is extended by a relative clause. Example (9) may illustrate this. The subordinate clause shows all the features of any other relative clause in Ik. The dummy pronoun also fits into the picture: In relative clauses a dummy pronoun is obligatory if the head of the relative clause is a peripheral participant, otherwise no dummy pronoun is used.

- (9) *ńtá ye-í-í kóróbá-a na ím-á gód-á d^e*
 NEG know-I-NEG what-NOM REL.SG child-NOM cry-a DP
 ‘I don’t know why the child cries.’

This accounts for the fact that sometimes the subordinate clause shows a final dummy pronoun and sometimes not. *kóróbáa* meaning ‘what’ would always be a core participant of the following erstwhile relative clause, as in (1), and therefore no dummy pronoun is needed at the end of the erstwhile relative clause. *kóróbáa* meaning ‘why’

would be always a peripheral participant of the erstwhile relative clause as in (3). The same is true in the other examples with and without dummy pronoun, respectively.

At stage II the context remains the same as at stage I, but the relative pronoun is deleted and its head remains the only marker which introduces the subordinate clause. This marker is the case inflected conjunction. Otherwise in Ik it is ungrammatical to drop the relative pronoun in relative clauses. Therefore the status of the remaining clause can no longer be interpreted as a relative clause. It also cannot be interpreted as a headless relative clause because this construction is not headless. The head remains in the form of the new conjunction.

At stage III the new conjunction widens its function: It is no longer restricted to introduce following subordinate clauses, it may now also introduce preceding clauses. Hereby a further feature of the nominal source is lost. The conjunction is frozen in one invariable form, which is the nominative form. This can be seen by examples (10) to (14): In these examples the subordinate clause precedes the main clause and the conjunction appears always in the nominative case irrespective of the case the conjunction had when used in the following subordinate clause: In example (11), *kɔɔbáa* occurs in the nominative case whereas in the corresponding example (4) with a different ordering of main and subordinate clause *kɔɔbáa* occurs in the accusative case. The same holds true for the corresponding examples (14) and (6): While in (14) *na* occurs in the nominative case, it occurs in the ablative case in (7).

- (10) *kɔɔbá-a im-á ɔɔd-^a ńtá iye-ń*
 what-NOM child-NOM cry-a NEG know-1.SG-NEG
 ‘Why the child cries I don’t know.’
- (11) *kɔɔbá-a im-á ɔɔd-^a ńtá iye-ń*
 what-NOM child-NOM cry-a NEG know-NEG
 ‘Why the child cries he doesn’t know.’
- (12) *kɔɔbá-a ńtíya-ńd-^a ńtá iye-ń-ń*
 what-NOM do-2.SG-a NEG know-1.SG-NEG
 ‘What you are doing I don’t know.’
- (13) *kɔɔbá-^a ńtíya-ńd-^a ńtá iye-ń*
 what-NOM do-2.SG-a NEG know-NEG
 ‘What you are doing he doesn’t know.’
- (14) *n-á ats-á d^e mo ńt-ńt-ń nabó*
 where-NOM come-a DP NEG reach-VEN-NEG again
 ‘Where he came from he could not return to again.’

These observations suggest that we are dealing with a grammaticalization process as summarized in Table 3.

Table 3 represents the grammaticalization path from noun to conjunction, as it can be reconstructed for Ik. The bold parts reflect the part of the construction which has been grammaticalized. The non-bold parts reflect the context in which the grammaticalization took place. The source structure is a case inflected noun used as an

Table 3. Grammaticalization from noun to conjunction in Ik

Source structure			
0	MC:S + OBJECT = N-CASE		⇒ MC
Target structure			
I	MC + N-CASE + REL = Conjunction		⇒ MC + sc-Complement
II	N-CASE = Conjunction		⇒ MC + sc-Complement
IIIa	N-NOM = Conjunction	+ MC used freely	⇒ sc + MC-Complement
IIIb	N-CASE = Conjunction		⇒ MC + sc-Adjunct

Table 4. Some features of grammaticalization of noun to conjunction in Ik

Property	0	> I	> II	> IIIa
	Noun	Noun + REL	Case inflected conjunction	invariable conjunction
May be pluralized	+	–	–	–
May be modified	+	–	–	–
May be determined	+	–	–	–
May be head of relative clause	+	+	–	–
Is case inflected	+	+	+	–

object in a main clause, represented in Table 3 as stage 0. *kóróbá-a* is still used as a noun meaning ‘thing’. At stage I of the target structure the object slot of the main clause is filled with an object with a relative clause functioning as a subordinate clause. At stage II the relative marker is deleted and only its head remains as a marker of the subordinate clause. At stage IIIa the new conjunction gains the ability to be used freely in its new function. It is no longer restricted to introduce a subordinate clause following the main clause; instead it is used in a structure where the subordinate clause precedes the main clause. Hereby the conjunction loses the ability to be case inflected. It is now frozen in an invariable form which always corresponds to the nominative form of the former noun.

Alternatively, the conjunction may expand its new function in a different way, illustrated by stage IIIb: Here the expansion relates to the kind of subordinate clauses it may introduce. It is no longer restricted to complement clauses; rather it may introduce adjunct clauses as well. The latter usage is documented for *na* in example (6). In this example, *na* cannot refer to any participant required by the main verb. The locative slot of the main verb is already filled by *óropoi-e*. Instead *na* shows the case required by the verb of the subordinate clause. This subordinate clause therefore introduces no longer a complement clause but rather an adjunct clause. In table 4 the features which the conjunctions show at each stage of their grammaticalization from noun to conjunction are summarized.

In Table 4, the grammaticalization process from noun to conjunction is presented from a different perspective. Whereas in Table 3 the grammaticalization was primar-

ily described with regard to syntactic context, in Table 4 the grammaticalization is described with regard to the features the grammaticalized item shows itself, such as nominal or conjunctive features.

As mentioned earlier, the source item is a noun and as a noun it shares all features a noun typically has in Ik. I will illustrate this again with *kɔɔbáa*, as *kɔɔbáa* is still used productively as a noun meaning ‘thing’. As a noun *kɔɔbáa* can be used in its plural-form *kóúóbá-a*, it may be modified by a count word or determined by a demonstrative or it may be the head of a relative clause. Example (15) may illustrate this: In (15), *kɔɔbáa* is used in its plural-form, it is modified by the count word *lébétsé* ‘two’ and determined by the demonstrative pronoun *ni*. It also is the head of a relative clause. Therefore all features listed in Table 4 in the left column have “+”.

- (15) *ńtá ye-í-í kóúóbá-a ni lébétsé ni íťya-íd-a*
 NEG know-1.SG-NEG thing.PL-NOM DEM.PL two.OBL REL.PL do-2.SG-a
 ‘I don’t know these two things which you are doing.’

As a conjunction the item loses more and more nominal features. At stage I *kɔɔbáa* loses most of its nominal features: It can no longer be pluralized, it can no longer be modified or determined. The only nominal features which remain are the ability to be head of a relative clause and be inflected for case. This stage can be illustrated by example (9). Therefore the first three have “+”.

A more careful look at example (15) shows that *kɔɔbáa* is used as a noun in the same context as the conjunction *kɔɔbáa* in the earlier mentioned example (1). Nevertheless, *kɔɔbáa* does not serve as a conjunction in (15). This can be seen as additional evidence of the fact that the grammaticalization process already started at stage I, when the conjunction still consists of two word units.

At stage II, which can be illustrated by examples (1)–(2), the relative pronoun is deleted, that is, *kɔɔbáa* alone serves as a conjunction. At this stage *kɔɔbáa* loses also the ability to take a relative clause introduced by a relative pronoun. The only nominal feature which remains is the case inflexion. Therefore in Table 4 all features get “–”, except the last one.

At stage IIIa, which is illustrated by examples (10) to (13), *kɔɔbáa* serves as a conjunction in an invariable form, therefore the last nominal feature, the ability to be case inflected, is also lost. Therefore in Table 4 all features have “–”.

There are other languages where conjunctions are grammaticalized in a similar way as in Ik, Ewe and Nama, for instance. The main difference between these languages and Ik lies in the fact that case inflected conjunctions appear to be unusual in the languages of the world.² The reason for this lies partly in the fact that not all languages have a grammaticalized case system; but even more important is the presence of the case anomaly which is responsible for case inflexion of the kind found in Ik.

2.2 Case markings on verbs

The example discussed in 2.1 is not an isolated case as can be shown with the following case, which is another instance of categorial misbehavior of case inflexion in Ik: Case is used on verbs and became part of verbal inflexion. The dative suffix *-ik^e* is used on verbs as a subjunctive.³

The dative suffix *-ik^e* is used at least in twelve different functions. Most of these are associated with nouns. In one function the dative case suffix is used with verbs to function as a subjunctive, that is, a verbal marker used to encode subordinate clauses. (16) to (19) may illustrate the grammaticalization chain from dative to subjunctive: In (16) the dative expresses the purpose role with the noun *kíja* ‘country’; in (17) the dative expresses also the purpose role with a verbal noun *ítín-es^a* ‘to cook’; in (18) again the purpose role is expressed now by a whole clause meaning ‘so that they carry water’. In this case the dative is suffixed to a fully inflected verb form. The latter is suggested by the fact that the verb *tír-es^a* ‘to hold’ carries the bound personal suffix for the third person plural *-at-*. Therefore *tír-es^a* shows the same behavior as any other inflected verb form in Ik. A closer look at (18) shows that it is not the dative alone which covers the new function subjunctive. The dative occurs in the suffix forms *-k^e* or *-e*; in the subjunctive the suffix is *-i-ke* or *-ie*. The subjunctive has an additional vowel *-i-*. This vowel *-i-* goes back to an optative marker. The optative marker alone is part of the verbal inflexion. It is used in expressions of commands or wishes including the imperative. In (19) the optative is used to express the purpose role with a full-fledged verb *wet-es^a* ‘to drink’.

- (16) *béd-í-a gó-on^a nci kíja-k^e*
 want-1SG-a go-INF.NOM I.OBL contry-DAT
 ‘I want to walk for my country (at a walking-competition).’
- (17) *ma ná im-a tɔbɔŋw-á bíy-e ítín-esí-k^e*
 give-a ENC boy-NOM food-ACC you-DAT cook-INF-DAT
 ‘The boy gave you the food for cooking.’
- (18) *na suban-át-ik^e tɔkɔɔd-íní logo-íkw-a ni kwátsi-k^a*
 when ready-3.PL-SBJ hold-NAR.3.PL container-PL-NOM REL.PL small-PL
tír-at-ík^e cué-k^a
 hold-3.PL-SBJ water-ACC
 ‘When they were ready for the journey they held small containers for carrying water. (Lit.: they held containers which are small so that they carry water.)’
- (19) *maa-ée cué níti-k^e wet-át-i*
 give-IMP2.SG water.OBL they-DAT drink-3PL-OPT
 ‘Give them water to drink! (Lit.: Give them water so that they may drink.)’

As (16) to (19) illustrate, the purpose role can either be expressed by the dative or by the optative, or by a combination of both, which is called the subjunctive form. The dative encodes the purpose role prototypically with nouns or items used in a nominal

Table 5. The sources of the subjunctive

Optative	+	Dative	>	Subjunctive
-i-	+	-k ^e (-e)	>	-ik ^e (-ie)

way, the optative with full-fledged verbs in commands and the subjunctive with full fledged verbs which cover the same slot as dative participants expressing the purpose role. Table 5 illustrates the grammaticalization of the subjunctive: The optative *-i-* and the dative *-k^e* are merged to the subjunctive *-ik^e*.

Once grammaticalized, the subjunctive has widened its range of uses: It has become an obligatory marker of certain subordinate clauses, such as all clauses introduced by the conjunction *na* ‘when’, which take the verb in the subjunctive. Semantically subjunctive clauses are no longer restricted to encoding the purpose role; they may as well encode temporal or conditional clauses, as in (20). Syntactically the subjunctive is no longer restricted to the following subordinate clause, it may also be used with a preceding clause, as (20) may show.

- (20) *ná* *ats-an-é* *aw-é* *jej-id-o* *bi-a* *sába-k^e*
 when come-IPS-SBJ home-DAT remain-2.SG-NAR you-NOM river-DAT
 ‘When coming home you [will] remain at the river. (Text 4/60)’

Table 6 gives an overview of the stages in the grammaticalization from dative to subjunctive. As mentioned earlier, stage 0 consists of a noun in the dative case expressing a purpose role. Instead of the noun, a verbal noun is used to express the purpose role at stage I. At stage II, the purpose role is no longer expressed by a noun phrase but rather by an embedded clause. Here the dative together with the optative is suffixed to a full-fledged verb. The whole embedded clause again expresses the purpose role. At stage III, the subjunctive widens its new function. It is no longer restricted to purpose clauses but rather can be used as a general marker for clause embedding. The only feature that changes from stage II to stage III is a widening of the new function.

To sum up, at the stages 0 to II the semantics remains the same: A purpose role is expressed; only at stage III the semantics is widened. The means changes from dative

Table 6. Grammaticalization of the subjunctive

Stage	Morphosyntax	Shape	Word status	Semantics	Syntactic status
0	Noun-DAT	-k ^e /-e	Noun	Purpose	NP
I	Verb-INF-DAT	-esa-k ^e /-ona- k ^e -esa -e/-ona-e	Verbal noun	Purpose	NP
II	Verb-(PPRON)-OPT-DAT (OPT-DAT= SBJ)	(at)-i-k ^e /(at)-i-e	Finite verb	Purpose	Embedded clause
III	Verb-(PPRON)-OPT-DAT (OPT-DAT=SBJ)	(at)-i-k ^e /(at)-i-e	Finite verb	General	Embedded clause

to subjunctive, which is a merger of optative and dative. The element which carries the marker varies from noun phrase to embedded clause.

A similar grammaticalization process can be found in Kanuri (a Nilo-Saharan language spoken at lake Chad) where also a dative suffix *-ro* has been grammaticalized to a subordination marker used with verbs (see Heine 1990: 131, 140–142).

2.3 An overview

But the examples discussed in Sections 2.1 and 2.2 are not the only cases of misbehavior found in Ik. Table 7 gives an overview of other items showing categorial misbehavior involving case inflected adverbs, prepositions, and postpositions. In all cases the source item is a noun.

Table 7. Some items showing categorial misbehavior in Ik

Source		Target		Case inflexions
Dative	<i>-k^e /-e</i>	Subjunctive	<i>-ik^e -ie</i>	
Copulative	<i>-k^o /-o</i>	Narrative	<i>-uo</i>	
NOUN		CASE INFLECTED CONJUNCTION		
	<i>tómeda (na)</i> ?		'where'	NOM, ACC
	<i>mená</i> 'thing'		'what'	NOM, ACC
	<i>kɔɔbáa</i> 'thing'		'what'	NOM, ACC
	<i>na</i> 'place'		'where'	NOM, ACC, DAT, ABL
	<i>tóimen</i> 'problem'		'that'	NOM, ACC, DAT
NOUN		CASE INFLECTED ADVERB		
	<i>wash</i> 'front'		'ahead', 'first', 'earlier'	DAT, ABL, COP, GEN
	<i>na</i> 'place'		'here'	NOM, AKK, DAT, ABL, COP, OBL
	<i>yasⁱ</i> 'truth'		'true', 'really'	DAT, COP
	[nominal source no longer known]	<i>edá</i>	'alone'	COP, GEN
		<i>muny^u</i>	'all', 'completely'	OBL, COP
		<i>jík^e</i>	'always'	DAT, GEN
		<i>koóke</i>	'there'	DAT invariable
RELATIONAL NOUNS		CASE INFLECTED POSTPOSITION/PREPOSITION		
	<i>ai^a</i> 'side'		'from'	all
	<i>akw</i> 'palm (of hand)', 'sole'		'inside'	all
	<i>búbú</i> 'stomach'		'under'	all
	<i>gwari</i> 'surface'		'top'	all
	<i>kann</i> 'back'		'behind'	all

In Table 7 the source items are listed in the left column, the grammaticalized target is listed in the middle and the case inflexions of the target are listed in the right column. As can be seen the relational nouns used as prepositions or postpositions are the most flexible with regard to case; among the case inflected adverbs the degree to which an adverb may show case inflexion varies. Some are basically used in one fossilized form, only in relics their ability to be case inflected remains. This holds true for *éda* ‘alone’, *muny^u* ‘all’, *jík^e* ‘always’ and *koóke* ‘there’. Others are still used in a variety of different forms. This holds true for *wash* ‘first, earlier’ and *na* ‘place’ (further discussion see also König 2002).

3. Conclusions

We have proposed a fairly far reaching explanation for the categorial misbehavior found in Ik. The kind of data here considered here is by no means unusual. On the contrary, it is to be expected, considering the development of the structures concerned. All the examples relate to only one kind of categorial misbehavior. It is expected, moreover, that further research will reveal an extensive number of additional examples. It is reasonable to hypothesize that they all will have followed a similar path of grammaticalization as outlined here.

List of abbreviations

1.	1. person	NEG	negation
2.	2. person	NOM	nominative
3.	3. person	NP	noun phrase
<i>a</i>	final <i>a</i> without meaning	OBL	oblique
ABL	ablative	OPT	optative
ACC	accusative	PL	plural
COP	copulative	PRON	personal pronouns
DAT	dative	REL	relative pronoun
GEN	genitive	SBJ	subjunctive
INF	infinitive	SC	subordinate clause
IPS	impersonal	SG	singular
MC	main clause	VEN	venitive
N	noun		

Notes

1. Ik is spoken in Uganda. According to Heine (1976), Ik belongs to the Kuliak languages. The external classification is still controversial: Greenberg (1963:86) assigned it to the Eastern Sudanic branch of the Nilo-Saharan family, and so did Ehret (1981a, b). Tucker (1967a,

1967b, 1971–3) pointed to Afroasiatic (“Hamito-Semitic”) connections, and others, like Laughlin (1975), suggested to leave Kuliak unclassified.

2. Ekkehard Wolff (p.c.) observes that there are case inflected conjunctions in Finnish and Z.Frajzyngier (p.c.) says that the same applies to Slavic.
3. See Heine (1983:70 & 139).

References

- Andersen, T. (1988). Ergativity in Pāri, a Nilotic OVS language. *Lingua*, 75, 289–324.
- Blanchon, J.-A. (1998). Semantic/pragmatic conditions on the tonology of the Kongo noun-phrase: A diachronic hypotheses. In L. M. Hyman & C. W. Kisseberth (Eds.), *Theoretical Aspects of Bantu Tone* (pp. 1–32). Stanford CA: CSLI.
- Böhm, G. (1983). Zwei Typen ‘Ergativsprachlicher’ Prädikation in Tschadischen Sprachen. In E. Wolff & H. Meyer-Bahlburg (Eds.) *Studies in Chadic and Afroasiatic Linguistics* (pp. 95–113). Hamburg: Buske.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge: Cambridge University Press.
- Ehret, C. (1981a). The classification of Kuliak. In T. C. Schadeberg & M. L. Bender (Eds.), *Nilo-Saharan: Proceedings of the First Nilo-Saharan Linguistics Colloquium, Leiden, September 8–10, 1980* (pp. 269–90). Dordrecht: Foris.
- Ehret, C. (1981b). Revising Proto-Kuliak. *Afrika und Übersee*, 64 (1), 81–100.
- Frajzyngier, Z. (1984). Ergative and nominative-accusative features in Mandara. *Journal of African Languages and Linguistics*, 6, 35–45.
- Greenberg, J. H. (1963). *The Languages of Africa*. The Hague: Mouton.
- Heine, B. (1976). *The Kuliak Languages of Eastern Uganda*. Nairobi: East African Publishing House.
- Heine, B. *The Ik Language. Preliminary draft*. Cologne: University of Cologne.
- Heine, B. (1990). The dative in Ik and Kanuri. In W. Croft & S. Kemmer (Eds.), *Studies in Typology and Diachrony: Papers presented to Joseph H. Greenberg on his 75th birthday* (pp. 129–49). Amsterdam: John Benjamins.
- König, C. (2002). *Kasus im Ik*. Köln: Rüdiger Köppe.
- Klimov, G. A. (1974). On the character of languages of active typology. *Linguistics*, 131, 11–25.
- Laughlin, C. (1975). Lexicostatistics and the mystery of So ethnolinguistic relations. *Anthropological Linguistics*, 17, 325–341.
- Mallinson, G., & Blake, J. (1981). *Language Typology*. Amsterdam: North-Holland.
- Miller, C. & L. Gilley (2001). Evidence for ergativity in Shilluk. *Journal of African languages and linguistics*, 22 (1), 33–68.
- Reh, M. (1996). *Anywa Language: Description and internal reconstructions*. Köln: Rüdiger Köppe [Nilo-Saharan, 11].
- Rude, N. (1983). Ergativity and the active-stative typology in Loma. *Studies in African linguistics*, 14 (3), 265–283.
- Tucker, A. N. (1967a). Fringe Cushitic. *Bulletin of the School of Oriental and African Studies*, 30 (3), 656–680.
- Tucker, A. N. (1967b). Erythraic elements and patternings: Some East African findings. *African Language Review*, 6, 17–26.
- Tucker, A. N. (1971–73). Notes on Ik. *African Studies*, 30, 341–54; 31, 183–201; 32, 33–48.

The typology of relative clause formation in African languages*

Tania Kuteva and Bernard Comrie

Our main goal is to establish the range of expression possibilities African languages have for encoding relative clauses. For this purpose, we put forward a six-way classification of relativization strategies identified in the languages of the world. The criteria relevant for us here are *where* and *how* the head noun is marked in the relative construction of the individual African languages of our sample.

The present investigation contributes to the research done so far in two ways:

- (i) it represents the first large-scale typological study of African relativization;
- (ii) it identifies – at this stage, only very tentatively – areal typological phenomena in relativization on the African continent.

1. Introduction

Relative clause formation has been studied from different perspectives. Lehmann 1984 offers a comprehensive typology of relative clauses based on a number of criteria including the linear sequencing of the constituents of the relative construction, i.e. the head noun and the relative clause. Hence the typology of prenominal, postnominal, circumnominal embedded relative clauses, and preposed, postposed adjoined relative clauses. Relativization has also been treated from a cognitive point of view, focusing on the cognitive motivation of relativization strategies. Thus Kibrik 1992 proposes that relativization involves only two global strategies, with a clear cognitive motivation, the combining strategy and the inserting strategy. The combining strategy reflects the process of composing two propositions that existed in the speaker's memory beforehand and happened to share an argument; the inserting strategy "corresponds to a process where a referent is first conceived through its participation in a certain event, and coded by a nominalized proposition, and then this complex nomination as a single whole, is inserted into another, broader event" (Kibrik 1992: 143–144). Kibrik proposes this two-way distinction as something relevant not to any particular occurrence of a particular relative construction in a particular language, but to the prototypical usages of relative constructions.

In the present study, we will be concerned neither with the linear order of the head noun and the relative clause (like in Lehmann 1984), nor with the global cognitive mechanisms identified in Kibrik 1992 as underlying relativization on a language-universal level. We will focus, instead, on the *formal*, morphosyntactic (as well as suprasegmental) means of encoding relative clauses as manifested in actual, language-specific relativization strategies.

Our object of investigation are African languages. In this paper we will only examine the strategies African languages use to relativize the subject. Our main goal is to establish the range of expression possibilities African languages have for encoding relative clauses. For this purpose – extending the classification proposed in Comrie 1981 and Comrie 1998 – we will put forward a six-way classification of relativization strategies identified in the languages of the world. The criteria relevant for us here will be *where* and *how* the head noun is marked in the relative construction of the individual African languages of our sample.

The present investigation builds on previous work on relative clauses in individual languages as well as groups of genetically related languages in Africa. It contributes to the research done so far in two ways:

- (i) it represents the first large-scale typological study of African relativization;
- (ii) it identifies – at this stage, only very tentatively – areal typological phenomena in relativization on the African continent.

2. Classification of cross-linguistic relativization strategies

The classification of relativization strategies that we will be using here builds upon Comrie 1981 and 1998, according to which there exist four major types of such strategies in the languages of the world. The first one is the *relative pronoun* strategy. In languages with this relativization strategy, the position relativized (subject, as contrasted to object, oblique, etc.) is indicated inside the relative clause by means of a pronominal element; this pronominal element is case-marked to indicate the syntactic/semantic role of the head noun within the relative clause. Note that the mere presence of a pronoun that is restricted to relative clauses, and is thus in some intuitive sense a relative *pronoun*, is not sufficient to define an instance of the relative pronoun strategy (Comrie 1998:61–62). Such a relative pronoun can be case-marked, for instance, not to indicate its role in the relative clause, but rather to agree in case with the head noun in the main clause. Thus in the following example from Modern Standard Arabic, the relative pronoun is nominative, like the head noun, whereas the position relativized in the relative clause is direct object (which would require the accusative case in Arabic):

- (1) Arabic (Comrie 1998:62)

‘al- yulaam- aani l- musiiqiyy- aani llaḏ- aani
 the boy DU.NOM the musical DU.NOM REL DU.NOM
 ‘The two boy musicians (whom Cyrano sent ...)

The following example illustrates the relative pronoun strategy:

(2) Bulgarian

Măžăt, [kojto me pozdravi včera], beše nemec.
 man.NOM REL.NOM me greet.3SG.AOR yesterday be.3SG.PAST German
 ‘The man who greeted me yesterday was a German.’

The second strategy is *non-reduction*, with two subtypes: *correlatives*, and *head-internal relative clause*. In the case of correlatives, “the head noun appears as a full-fledged noun phrase in the relative clause and is taken up again at least by a pronoun or other pronominal element in the main clause”, Comrie 1998:62, cf.

(3) Pirahã (Everett 1986:276)

boitóhoi bog- ái- hiab- i s aoaxái boito báosa xig i sai
 boat come ATEL NEG EP ? INTER boat barge bring EP NOML
 (hix)
 COMPL
 ‘Might it be that the boat (which) tows barges is not coming?’

In internally-headed relative clauses, “the head is represented by a full noun phrase inside the relative clause, and has no explicit representation in the matrix clause”, Comrie 1998:62–63)

(4) Maricopa (Gordon 1986:255)

aany=lyvii=m ‘iipaa ny- kw- tshqam- sh shmaa m
 yesterday man 1 REL slap+dist SJ sleep real
 ‘The man who beat me is asleep.’

The third major type of relativization strategy involves *pronoun retention*. In this case, the position relativized is explicitly indicated by means of a resumptive personal pronoun, cf.:

(5) Babungo (Schaub 1985:34)

ma yè wa ntia fán nwó si sà n ghó
 I see+PF person that who he P2 beat+PF you
 ‘I have seen the man who has beaten you.’

Note that we define the pronoun retention strategy in a narrow sense, as one which involves cases where a pronoun or pronominal marker referring to the head noun of a relative clause can be analyzed as a resumptive pronoun only if its occurrence in the corresponding independent clause is impossible or optional. Thus Babungo is a good example of pronoun retention not only because the head noun is referred to by means of a resumptive element within the relative clause cf. *ɲwá* in ex. (5) above) but also because using a resumptive element for the subject is not characteristic of simple declarative clauses, cf.:

- (6) a. Babungo (Schaub 1985:23)

Làmbí !sáŋ ɲwǎ

Lambi beat+IMPF him

‘Lambi beat him.’

- b.
- Làmbí sáŋ !ɲwǎ*

Lambi beat+PF him

‘Lambi has beaten him.’

Finally, the *gap* relativization strategy involves cases where there is no overt reference to the case role of the head noun within the relative clause, cf.

- (7) Turkish (Comrie 1998:82)

[*kitab- ı al- an*] *öğrenci*

book ACC buy PRT student

‘The student who bought the book’

In addition to all the assumptions underlying the classification presented in Comrie 1981, the present classification is based on the assumption that *all* natural languages can relativize subjects since we are including functional equivalents of relative clauses, as in Khmer and Walpiri. For instance, Walpiri has no specific subordination construction whose sole, or even prototypical function is to encode a relative clause. It uses, instead, a general unified, noun-modifying construction which – depending on context – may be interpreted as either a subordinate temporal or a relative clause:

- (8) Walpiri (Comrie 1981:137)

ɲatʷulu -ɭu -ɳa yankiri pantunu kutʷa -lpa ɲapa ɲaɲu

I ERG AUX emu speared CONJ AUX water drank

‘I speared the emu while it was drinking water.’

‘I speared the emu that was drinking water’

Instead of regarding languages such as Walpiri as irrelevant to relativization strategies, on the present account we treat them as a manifestation of the paratactic strategy (see below).

Furtheron, on the present classification, the “non-reduction” strategy covers not only the two subtypes of correlatives and internally-headed relatives but one more relativization strategy that we propose to term the *paratactic* (cf. Engl. *That man just passed by us, he introduced me to the Chancellor of the University yesterday*). The paratactic relativization strategy involves cases where the ‘relative’ clause contains the full-fledged head and is the same as an unmarked simple (declarative) clause; the relative and main clauses are only very loosely joined together, cf. also:

- (9) Amele (John Roberts, p.c.)

Mel mala heje on

((*mel*) eu) busali nu-

boy chicken illicit take3SG.SU.REM.PAST boy that run away go-

i- a
 3SG.SU- TOD.PAST
 ‘The boy that stole the chicken ran away.’

In the above example, *mel* ‘boy’ is the ‘relativized’ noun in the ‘relative’ clause. This nominal can be optionally referred to in the following ‘matrix’ clause either by the demonstrative *eu* ‘that’ or, if clarification is needed, *mel eu* ‘boy that’. What links the two clauses is the rising intonation at the end of the first clause. This indicates that it is not a final clause and is in either a subordinate or coordinate relationship with the following clause.

At this point, it should be noted that a language can have different relativization strategies for different semantic/case roles. A paradigm example of this are languages where relativization upon the subject involves the gap strategy, and relativization upon obliques pronoun retention. Persian is one such language:

- (10) a. Persian (Comrie 1981:140–141; Comrie 1998:63)
*Mard- i [ke (*u) bolandqadd bud] juje- râ kost.*
 man- that he tall was chicken ACC killed
 ‘The man that was tall killed the chicken.’
- b. *Mardhâi [ke ketâbhâ râ be ânhâ dâde bud- id]*
 men that books ACC to them given were 2SG
 ‘the men that you had given the books to’ (lit. ‘the men that you had given the books to them’).

In sum, here we will distinguish between six relativization strategies upon the subject: (i) relative pronoun, (ii) correlatives, (iii) internally-headed relatives, (iv) paratactic, (v) pronoun retention, and (vi) gap.

3. The empirical aspect

Our language sample consists of 54 languages, and even though it is a convenience sample, it covers all major genetic families (see Appendix I: Language Sample). In these languages we could identify only three of our proposed six types of relativization strategy, namely gap, pronoun-retention, and correlative (see Appendix II). In none of the languages examined could we identify a head-internal strategy. That it is hard to come across internally-headed relative clauses in African languages has already been observed, cf. Creissels (2000:256).

Another strategy for which we have negative evidence is the paratactic one. A possible exception here comes from Koyaga – a Southern dialect of the Manding cluster, spoken in Mankono (The Ivory Coast) – which is closely related to one of our sample languages, namely Bambara. According to Creissels (forthc.), the strategy used for relativization in Koyaga comes closest to what we have termed here the paratactic strategy. More precisely, the only way of expressing relativization in Koyaga is to use a sequence of two clauses that can be used separately as independent assertive sentences; in the

first clause the demonstrative *mən* is used to refer to an entity which, in the second clause, is also referred to by some anaphorical means. Thus the only way to render the Koyaga sentence ‘The lion killed the cow that the man saw’ is:

- (11) Koyaga (Denis Creissels, p.c.)
Cε wa nisi mən ye jra w’ o fya
 man TAM COW DEM/REL see lion TAM PRO kill
 i. ‘The man saw this cow, (and) the lion killed it.’
 ii. ‘The lion killed the cow that the man saw.’

Crucially, the two component clauses of the above sentence can be used separately – with no modification of any morpheme – as canonical simple sentences:

- (12) Koyaga (Denis Creissels, p.c.)
cε wa nisi mən ye
 man TAM COW DEM see
 ‘The man saw this cow.’
 (13) Koyaga (Denis Creissels, p.c.)
jra w’ o fya
 lion TAM PRO kill
 ‘The lion killed it.’

Hence, we are well justified to treat the Koyaga example as a manifestation of the paratactic strategy. Moreover, having compared Creissels’ data on Koyaga to the Bambara data, we believe that in this particular case we may assume that the synchronic variation of formal means of relativization here are suggestive of the grammaticalization path of a paratactic relative construction developing into a correlative one. The paratactic relative construction in Koyaga can be placed at the beginning point of this path, whereas the Bambara relativization strategy represents an advanced stage along the same path. Thus a closer look at the Bambara data reveals that in this language:

(i) the relativization strategy is the correlative one, cf.

- (14) Bambara (Denis Creissels, p.c.)
Cε ye misi min ye jara y’ o faa
 man TAM COW REL see lion TAM PRO kill
 ‘The lion killed the cow that the man saw.’

(ii) the relative clause is unambiguously identified as such by the presence of the relativizer *min* immediately after the head noun.

Now, the morpheme *min* is obligatory in relativization, but it never occurs in independent clauses:

- (15) Bambara (Denis Creissels, p.c.)
 **Cε ye misi min ye*
 man TAM COW REL see
 not acceptable as an independent clause

- (16) Bambara (Denis Creissels, p.c.)

jara y' o faa
 lion TAM PRO kill
 'The lion killed it.'

In other words, the Bambara *min* structure is a specific relative construction.

Crucially, the Bambara relativizer *min* can be shown to originate from a demonstrative, and its cognates in Koyaga (cf. the morpheme *men* in the Koyaga examples above) and other Southern dialects still function in independent clauses as demonstratives. That is, what in Southern Manding dialects such as Koyaga functions as a demonstrative in simple independent sentences – and as a demonstrative/relativizer in complex, paratactic sentences – has specialized, and grammaticalized, into the relativizer *min* in Bambara.

That demonstratives develop into relative clause markers is certainly nothing unusual (cf. Heine and Kuteva (2002) for numerous examples of a *demonstrative* → *relative clause marker* development in the languages of the world). What is interesting in the Koyaga-Bambara case is that here we can observe a new structural “environment” for the demonstrative to come to function as a relative clause marker. Thus in most cases familiar to us, we are dealing with the grammaticalization of the demonstrative within a particular paratactic construction into a relative clause marker within a particular *syntactic* construction (embedded relative clause). For example, one of the ways to express the relative in Old English was to use the simple demonstrative *se, seo, þat*:

[T]he use of the demonstrative as a relative appears to have come about simply by the subordination of the second of two originally consecutive sentences to the first; thus, ‘he came to a river; that (or this) was broad and deep’ whence ‘he came to a river that was broad and deep.’ (OED:25)

In the Koyaga-Bambara case, however, we are dealing with a demonstrative within a particular paratactic construction grammaticalizing into a relative clause marker within a specialized *hypotactic* construction (the correlative one).

To put it in a nutshell, on the basis of Creissels’ data, we could say that the correlative relativization construction in Bambara could possibly be traced back – historically – to a paratactic strategy such as the one presently used in Koyaga. Hence, there is at least one instance of the paratactic relative structure attested on the African continent, the Koyaga relative construction. However, we can say that with respect to relativization on the object, since object relativization is all we have data on. Given that in the present paper, object relativization remains beyond the scope of our investigation we have not counted Koyaga as a sample language.

We have not been able to come across clear cases of the relative pronoun strategy, either.¹ A particularly interesting case in this respect are some Bantu languages. On the traditional account, Bantu languages have been regarded as having a subject relative pronoun. The very term “subject relative pronoun” may be misleading at first sight because it is suggestive of a grammatical distinction between relativization upon the subject and relativization upon non-subject case roles. A closer look at the linguistic

data shows, however, that the subject relative pronoun codes not the case role of the head noun within the relative clause – which would be a case of the relative pronoun strategy – rather, it codes Nominal Class agreement between the head noun and some element (very often the verbal complex) in the relative clause. Let us take, for example Swahili, where the most frequent way to form a relative clause involves the use of the invariable morpheme *amba-*, and what has been traditionally called the “relative pronoun” attached to it, followed by the verbal complex of the relative clause:

(17) Swahili (Watters 2000: 227)

m- toto amba- ye a- na- lala
 CL1- child REL ye s/he PRES sleep
 ‘The child who is sleeping’

In this example, the morpheme *-ye* has been traditionally referred to as a relative subject pronoun. There are two reasons for this. First, it is a pronominal element referring back to the head noun *mtoto*, and agreeing with it in Nominal Class (CLASS 1). Second, it is a form specialized for use only in relative clauses. Note that we are not dealing with a straightforward repetition of the Nominal Class marker of the head noun (which is marked by the prefix *m-* in the main clause). We have a special morpheme *-ye*, instead, which corresponds to the Nominal Class 1 marker *m-*, and is only used with the relative clause. On the basis of these facts, we might wish to conclude that this is a case of a relative pronoun specialized for indicating the case role of the subject within the relative clause. Examples like the following one show, however, that this is not the case because the same morpheme *-ye* appears as a referring element to the head noun also in cases where it is not the subject within the relative clause, cf. the following example, where we have relativization upon the object:

(18) Swahili (Bernd Heine, p.c.)

m- toto amba- ye ni- li- mw- ona
 CL1 child REL YE I PAST OBJ see
 ‘The child whom I saw’

Therefore, the morpheme *-ye* is to be regarded as a Nominal Class agreement marker but not a marker of the case role of the head noun within the relative clause. In other words, we are only dealing with an agreement marker and not a case-marking relative pronoun. How shall we code languages like Swahili then?

One possibility is to regard subject relativization in such languages as a pronoun retention because of the existence of the subject prefix in the verb complex (cf. the subject prefix *a-* in (17) above). In fact, in some Bantu languages such as Zulu, one could even speak of “double” pronoun retention. Thus, in Zulu the relativization upon the subject involves a relativizer morpheme *a-*, an initial vowel (IV), which varies with the Nominal Class of the subject, and a relativization suffix (Poulos 1982). Note that the initial vowel is nearly always identical with the vowel of the subject prefix/noun class prefix. There is a coalescence of the relativizer vowel *a-* with the initial vowel concerned whereby the following three kinds of coalescence can be identified: (i) $a + a > a$; (ii) a

+ i > e; and (iii) a + u > o. (The relativizer *a-* is – most likely – a historical residue of a demonstrative pronominal marker, Erhard Voeltz, p.c.) The resultant vowel is then either prefixed to the Subjectival Concord (in cases where the Subjectival Concord commences in a consonant) or it replaces the Subjectival Concord (in cases where the latter consists of a vowel only). In the former case then, there is a sense in which one could arguably speak of a retained “trace” of referring to the head noun by means of (i) the vowel of the Subjectival Concord; and (ii) the Subjectival Concord itself:

- (19) a. Zulu (Erhard Voeltz, p.c.)

abantu abambonayo

aba- ntu a a ba- m- bona- yo

CL2 person REL₁ IV CL2 OBJ see REL₂

‘People who see him’

- b. Zulu (Poulos 1982:70)

Isitshudeni esifunda kakhulu sizophumelela.

Isi- tshudeni a- i- si funda kakhulu si zo phumelela

CL7 student REL IV CL7 study a lot CL7 FUT pass

‘The student who studies a lot will pass.’

However, since the subject prefix in the verb complex within the relative clause is obligatory, and since here we have adopted a narrow definition of pronoun retention (see Section 2), our classification of most Bantu languages on our sample is a gap rather than pronoun retention.² The reason is that – if we go back to Swahili, for instance – strictly speaking, there is no overt, case-marked reference to the head noun: *amba-* is an invariant morpheme, *ye-* is a Nominal Class agreement marker – as we have argued above – and none of the morphemes belonging to the relative clause is case-marked.

Cibak (Chadic, Afroasiatic) also represents a case of a language which does not fit straightforwardly into any of the types of the present typological classification. Even though it comes closest to the relative pronoun strategy, there are some characteristics of the case-marked morphemes in Cibak relative clauses which are suggestive of a strategy other than the relative pronoun one. More precisely, in sentences with relativization upon the subject, a special case marker for the subject *tara* is used:

- (20) Cibak (Frajzyngier 1987:436)

zər nam tara si

boy REL SUBJ come

‘The boy who came’

If the head noun is the patient of the relative clause, both the patient and the agent are appropriately marked by the object marker *tə*, and the subject marker *tara*, respectively:

- (21) Cibak (Frajzyngier 1987:436)

mwala nam tə zər ni tara tsar- ba

woman REL OBJ boy DEF SUBJ choose VENT

‘The woman whom the young man has chosen’

Since the morphemes *tara* and *tə*, which mark the case roles of the subject versus the object, seem to be just case markers and not pronominal elements (Frajzyngier 1987:437), we have coded the relativization strategy in Cibak as “other”.

Another result of the present study is the negative evidence for a particular sub-type of the gap strategy, the unified, noun-modifying construction (Comrie 1998), which functions as a single formal means for marking not only what translates English relative clauses but also a number of other types of clause, e.g. the Fact-S construction (as in “The fact that he doesn’t know me...”), as illustrated below:

- (22) a. Karachay-Balkar (Comrie 1998:81)
 [*kitab- i al- ʒan*] *oquwču*
 book ACC buy PRT student
 ‘The student who bought the book’
 b. Karachay-Balkar (Comrie 1998:81)
 [*oquwču al- ʒan*] *kitab*
 student buy PRT book
 ‘The book that the student bought’
 c. Karachay-Balkar (Comrie 1998:81)
 [*prezident kel- gän*] *hapar*
 president come PRT news
 ‘The news that the president has come’
 d. Karachay-Balkar (Comrie 1998:81)
 [*et biš- gän*] *iyis*
 meat cook PRT smell
 ‘The smell of meat cooking’

General noun-modifying constructions are very common in some parts of the world, e.g. in South East Asia, but – on the basis of our present knowledge of relativization in African languages – they seem to be very rare on the African continent.³ In our sample we came across a single language, Yoruba, where we could possibly speak of a unified modifying clause construction but only with respect to non-subjects. For subjects, Yoruba seems to employ pronoun retention since, according to Awobuluyi (1982:94): “When the noun qualified is identical with the subject of the sentence that is functioning as a qualifier, that subject is replaced by *ó* [the impersonal subject ‘it’, see Williams (1977:475), TK & BC]”, cf.:

- (23) a. Yoruba (Awobuluyi 1982:94)
 Ọkùnrin tí ó pè mí
 man REL IMPERS.SUBJ.PRO call me
 ‘The man who called me’

With non-subjects, however, only *tí* is used. Crucially, *tí* is also used as a complementizer, as well as a factive nominalizer and a Fact-S construction, as the following examples illustrate:

- (23) b. Yoruba (Williams 1977:476)
ìwé tí mo rà dára
 book REL I buy be good
 i. 'The book that I bought is good'
 ii. 'The fact that I bought a book is good.'
- c. Yoruba (Williams 1977:475)
ó dára tí mo ra ìwé
 it be good COMP I buy book
 'It is good that I bought a book.'
 (i.e. 'The fact that I bought a book is good.')
- d. Yoruba (Williams 1977:475)
ríra tí mo ra ìwé dára
 buying COMP I buy book be good
 'The fact that I bought a book is good.'
- e. factive
 Yoruba (Williams 1977:475)
ó ɛ ohun àjẹ̀jì tí ìlẹ̀kùn yẹn tí
 it is thing strange COMP door that be-closed
 'It is strange that the door is closed.'

versus:

- (23) f. non-factive
ó dàbí ẹní- kpé ìlẹ̀kùn yẹn tì
 it seems as- if door that be-closed
 'It seems that the door is closed.'

The present results are significant not only with respect to the negative evidence related to the head-internal, the relative pronoun and the gap strategy but also with respect to positive evidence.

Our main piece of positive evidence also has a theoretical significance because it relates to the Keenan and Comrie (1977) Accessibility Hierarchy of Relativization (SUBJECT > DIRECT OBJECT > INDIRECT OBJECT > POSSESSOR). One of the generalizations made regarding the accessibility hierarchy is that the pronoun retention strategy is preferred at the lower end of the hierarchy (whereas gapping is preferred at the higher end), cf. ex. (10a) and ex. (10b) above, and that it is hard to come across languages that employ the pronoun retention strategy with subjects. Thanks to the present study of African languages, we are now in a position to acknowledge the fact that there are languages, at least on the African continent, where pronoun retention is employed with subjects. As pointed out in Section 2 already, Babungu, a Bantu language spoken in Cameroon, cf. ex. (5) above, exhibits subject relativization by means of pronoun retention. Note that for object relativization, Babungu employs a gap, which is usually optional, but with a few verbs in the perfective aspect this strategy is reported to be obligatory (Schaub 1985:34):

(24) Babungo (Schaub 1985:34)

a. optional gap strategy

mà yè wěembwā fáŋ tĩ wĩ sɛ sǎŋ (ŋwà)

I see-PF child who father his P2 beat-PF (him)

‘I have seen a child whom his father had beaten.’

b. obligatory gap strategy

mà yè ŋkáw ŋkíə fáŋ Làmbí kò

I see-PF chair that which Lambi give-PF

‘I have seen the chair which Lambi gave.’

In addition, our data from African languages enable us to distinguish between a number of subtypes of the gap strategy, which – as mentioned above – seems to always involve a specific relative construction in African languages. Thus, depending on the number of relativization markers used to encode the gap strategy, we have been able to identify the following three gap situations on the African continent.

A. **Covert (= gap with zero relativization marker)**, e.g. Koyra (Omotic, Afroasiatic), where there is a specific subordinate clause construction used for relativization on subjects; this construction employs a perfect verbal suffix *-a*, and there is no element referring to the head noun within the relative clause:

(25) a. Koyra (Hayward 1982:255)

ha mud- a mātay ĵiletakko

this sprout- PERF grass-NOM green-COP

‘This grass which has sprouted is green.’

b. Koyra (Hayward 1982:255)

zine hand- a geri hay’uttolso

yesterday go PERF people-NOM die-3POL.PERF

‘The people who went yesterday died.’

c. Koyra (Hayward 1982:255)

harre wond- a ‘indoy yōdonikko

donkey buy PERF woman-NOM come-3F.SG-PERF

‘The woman who bought the donkey has come.’

B. **Gap with one relativization marker**, cf. the gap construction with an invariable relativizing morpheme (i.e. a conjunction) *si* which (Anglo) Ewe uses for relativizing on both subjects and non-subjects:

(26) a. (Anglo) Ewe (Lewis 1985:198)

ame si fiε agbalē- a

person REL buy book DEF

‘The person who bought the book’

b. (Anglo) Ewe (Lewis 1985:198)

agbalē si Kofi fiε

book REL Kofi buy

‘The book Kofi bought’

- c. (Anglo) Ewe (Lewis 1985:198)
ame si Kofi fie agbalẽ- a na
 person REL Kofi buy book DEF for
 'The person Kofi bought the book for'

Another example is Dawuro (Omoti, Afroasiatic), where *-a* (an invariable affix) is used as a relativizer with both subjects and non-subjects, and irrespective of the number and gender of the head noun:

- (27) a. Dawuro (Hirut Woedemann, p.c.)
hawaa y- ee- dd- a imatʼaa ta ʔer- ay
 here come PER PAS REL guest I know IMP
 'I know the guest who came here.'
- b. Dawuro (Hirut Woedemann, p.c.)
ta uš- ee- dd- a maatʼa- y loʔa
 I drink PER PAS REL milk NOM good
 'The milk which I drank was good.'

C. Gap with two relativization markers, e.g. Kxoe, a Khoisan language spoken in Namibia, where subject relativization involves, in most cases, (i) nominalizing person/gender/number agreement marker placed at the end of the relative clause, and identical with the person/gender/number suffix of the head noun, and (ii) a participial morpheme *ko*, suffixed to the main verb within the relative clause, cf.:

- (28) a. Kxoe (Christa Kilian-Hatz, p.c.)
llgẽẽ kxòè.hè kx'ó.xò 'à tcéka.kà.rá kò té.hā.hè, |èẽ
 woman.3SG.F meat OBJ prepare.2NDJUNC PART stay.PAST.3SG.F now
|í.yé.tè
 sing.1ST.JUNC.PRES
 'The woman who prepared the meat is singing now.'
- b. Kxoe (Christa Kilian-Hatz, p.c.)
llgẽẽ kxòè.hè kx'ó.xò 'a tcéka.kà.rá kò
 woman.3SG.F meat OBJ prepare.2NDJUNC PART
hì.yé.hā.hè, |èẽ |í.yé.tè
 do.2NDJUNCT.PAST.3SG.F now sing.1ST.JUNC.PRES
 'The woman who prepared the meat is singing now.'

Another example is Giziga (Chadic), where one of the relativization markers involves the prefix *mu* as the form of a special, relative tense, and the other a postrelative, subordinating marker *ná*, separating the entire relative clause from the remainder of the sentence, cf.:

- (29) Giziga (Frajzyngier 1986:420)

Mbùr mú sáwà vrà 'à Dlàgò ngá dàambó ná 'á.r.lè
 man PART come out PREP Dogba POSS yesterday SUB 3S.GO.PERF
'à Kòzà
 PREP Koza
 'The man who came from Dogba yesterday has gone to Koza.'

4. Conclusion

In sum, our investigation of relative clause formation in African languages reveals clear cases of three out of the six relative strategies we have been able to identify in the languages of the world. Even though we have come across only half of all attested strategies, it is noteworthy to draw attention to the diversity that African languages exhibit with respect to the parsimony/abundance with which they encode relativization strategies. Some languages employ absolutely no relativization markers, e.g. Maale (Omotic, Afroasiatic, spoken in Ethiopia) cf.:

- (30) Maale (Amha 2000:161)

ʔíní [[ziginó muk -é] ʔatsi] za-é-ne
 3MS:NOM yesterday come PF person:M:ABS see-PF-A:DCL
 'He saw the man who came yesterday.'

Here the relative clause precedes the head noun and it contains no pronominal element co-referential to the relativized noun.

Other languages double their relativization marking. This type of marking can be readily identified in all those languages which employ the so-called *bracketing* device in relativization. In "bracket" relativization we are dealing with relative clauses which are enclosed by two elements. These elements are, usually, formally identical – or similar – to deictics, cf.:

- (31) Mbum (Hagège 1970, cited in Chumbow 1977:288)

úi àí mì zàhzáh nú bèlbél
 woman REL I met (DET/REL) is beautiful
 'The woman I met is beautiful.'

The forms *ái* and *nu* which delineate the relative clause are deictics of different types. In Ngbaka (a Central African language), bracketing constitutes a repetition of the same relativization marker, *ne...ne*, which, again, is a deictic (Thomas 1963:270, cited in Chumbow 1977:288). Similarly, relative clauses in Bari (a Nilotic language) use a double marking strategy if the head noun is definite (Dimmendaal, *forthc.*): along with gender-sensitive markers introducing a relative clause in Bari, there is a set of demonstrative copy elements at the end of the relative clause, which are formally identical to the proximate demonstratives *na*, *lɔ*, *kone*, *kolɔ*, cf.:

- (32) Bari (Gerrit Dimmendaal, *forthc.*)

ɲutu lɔ (gwon)a pɔ ni kaje lɔ
 man REL PERF come here yesterday REL
 'The man who came yesterday'

Comrie and Kuteva (2005) discuss cases of African languages employing up to five relativization markers. (33) below gives an example of a language, Ngemba (a Mbam-Nkam language of Cameroon), using no less than five distinct morphemes which encode the relative clause construction (a relative conjunction/determiner (varying for number and nominal class), optional complementizer marker *-bah*, verbal suffix *-ne* (a multipurpose marker for topicalization, nominalization and relativization), pronoun retention, sentential definitizer *-la* (related to the determiner system), cf. Chumbow (1977:296–297: 302):

- (33) Ngemba (Chumbow 1977:290)

nyung wá bah a- keshung- ne mung wa la a- kung atsang
 man REL bah he TNS.beat ne child DET la he enter prison
 'The man who beat the child went to prison.'

How are we to explain this diversity with regard to the elaborateness of expressing relativization in African languages?

Our answer is that different languages have different degrees of elaborateness because just as "variety is the spice of life", so too is diversity the "spice" of language reality. The *abundance* of marking is – in a sense – very similar to the *redundancy* of marking since they are both there without actually being needed; some languages are just more precise in their encoding of relativization, whereas others are more vague, or indeterminate. The former are more context-independent, and the latter more context dependent.

Notes

* The authors wish to thank Bernd Heine and Erhard Voeltz for their insightful suggestions and valuable comments on this paper. We are especially grateful to Denis Creissels for kindly offering us access to his unpublished data on Koyaga. We are also greatly indebted to Gerrit Dimmendaal, Zygmunt Frajzyngier, Christa Kilian-Hatz and George Poulos.

1. Note that one of the languages on our sample, Luganda, may possibly turn out to have a relative pronoun strategy, but at the present stage of our investigation the data on this language is inconclusive. Denis Creissels (p.c.) has drawn our attention to another language, Mina, which – possibly – also employs a relative pronoun strategy but in this case, again, we have no conclusive data.

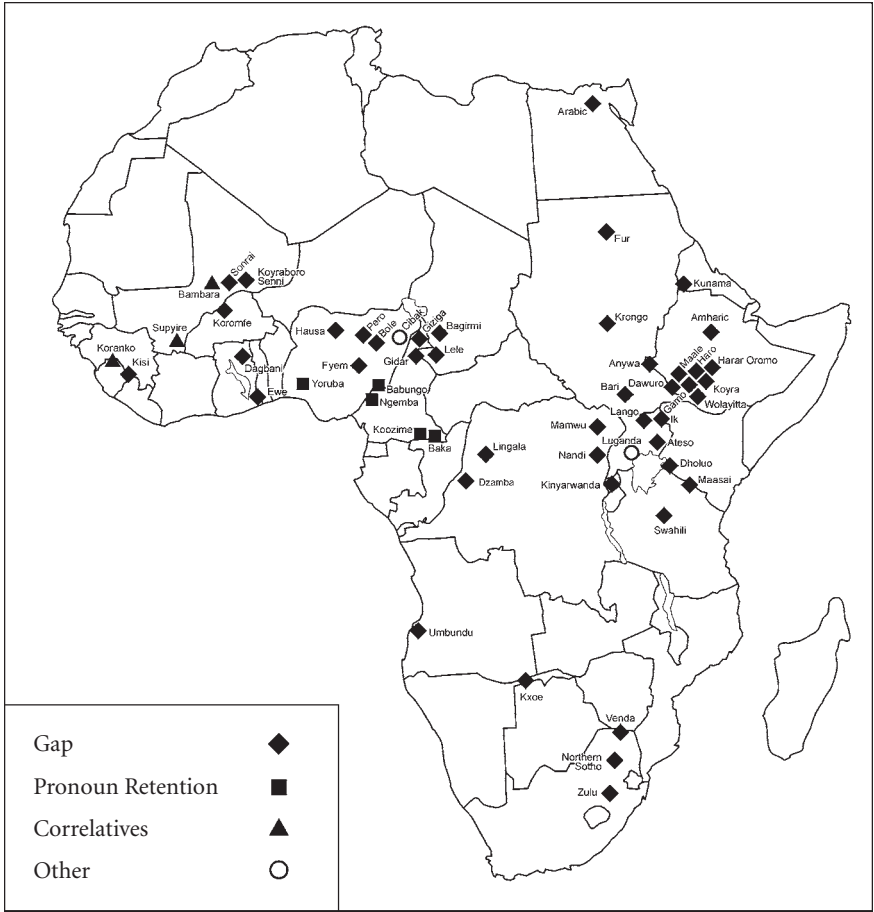
2. Denis Creissels (p.c.) is also of the opinion that due to the obligatory indexation on the verb (which is very often the case with the relativization of subjects) the subject relativization strategy in Bantu languages should be treated as a gap.

3. There are works which are suggestive of the existence of something similar to a general, noun-modifying clause construction, cf. Watters (2000:223), where it is mentioned, et passim, that “In conditional clauses, conditional morphemes are commonly used, but some languages use relative clauses (Efik, Benue-Congo)”. Note, however, that this statement still presupposes the existence of a specific relative construction rather than a general noun-modifying clause.

References

- Amha, A. (2000). The Maale language. *PhD Dissertation*. University of Leiden.
- Awobuluyi, O. (1982). *Essentials of Yoruba grammar*. Ibadan: University Press Limited.
- Chumbow, B. (1977). Relatives as determiners: A case from Ngemba. In P. F. A. Kotey & H. Der-Houssikian (Eds.), *Language and Linguistic Problems in Africa. Proceedings of the VII conference on African linguistics* (pp. 283–302). South Carolina: Hornbeam Press.
- Comrie, B. (1981). *Language Universals and Linguistic Typology. Syntax and morphology*. Chicago, IL: University of Chicago Press.
- . (1998). Rethinking the typology of relative clauses. *Language Design*, 1, 59–86.
- Creissels, D. (2000). *Typology*. In B. Heine & D. Nurse (Eds.), *African Languages. An introduction*. Cambridge: Cambridge University Press.
- Dimmendaal, G. J. (Forthcoming). *Nilotic languages*.
- Everett, D. (1986). *Pirahã*. In D. Derbyshire & G. Pullum (Eds.), *Handbook of Amazonian Languages* (pp. 201–325). Berlin: Mouton de Gruyter.
- Frajzyngier, Z. (1986). From preposition to copula. In V. Nikiforidou et al. (Eds.), *Proceeding of the Twelfth Annual Meeting of the Berkeley Linguistics Society, February 15–17, 1986* (pp. 371–386). Berkeley CA: BLS.
- . (1987). The relative clause in Proto-Chadic. In H. Jungraithmayr & W. W. Müller (Eds.), *Proceedings 4th International Hamito-Semitic Congress* (pp. 425–450).
- Gordon, L. (1986). *Maricopa Morphology and Syntax*. Berkeley, CA: University of California Press.
- Hagège, C. (1970). *La langue mbum de Nganha (Cameroun)*. 2 tomes. Paris: SELAF [Bibliothèque de la Société pour l'Étude des Language Africaine 18–19].
- Hayward, R. (1982). Notes on the Koyra language. *Afrika und Übersee*, Band LXV.
- Heine, B., & Kuteva, T. (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, B., & Nurse, D. (Eds.). (2000). *African Languages. An introduction*. Cambridge: Cambridge University Press.
- Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8, 77–100.
- Kendall, M. (1971). Sonrai relative clauses. In *African Language Review*. 1970/1971, Vol. 9 (pp. 34–50).
- Kibrik, A. (1992). Relativization in polysynthetic languages. *International Journal of American Linguistics*, 58 (2), 135–157.
- Kotey, P. F. A. & Der-Houssikian, H. (Eds.). (1977). *Language and Linguistic Problems in Africa. Proceedings of the VII conference on African linguistics*. South Carolina: Hornbeam Press.
- Lehmann, C. (1984). *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Gunter Narr Verlag.

- Lewis, M. (1985). *Précis from the 15th Conference on African Linguistics*. UCLA, March 29–31, 1984. Los Angeles, CA: University of California Press.
- Poulos, G. (1982). *Issues in Zulu Relativization*. Communication No. 11. Department of African Languages. Rhodes University.
- Schaub, W. (1985). *Babungo*. London: Croom Helm.
- Thomas, J. (1963). *Le parler ngbaka de Bokanga*. The Hague: Mouton
- Watters, J. (2000). Syntax. In B. Heine & D. Nurse (Eds.). (2000). *African Languages. An introduction* (pp. 194–230). Cambridge: Cambridge University Press.
- Williams, W. (1977). The so-called relativized and cleft predicates in Krio: One step closer to an understanding of creolization. In P. F. A. Kotey & Der-Houssikian, H. (Eds.), *Language and linguistic problems in Africa. Proceedings of the VII conference on African linguistics* (pp. 467–478). South Carolina: Hornbeam Press.



© 2001 Cartography: Monika Feinen, Institute for African Studies, University of Cologne

Figure 1. Relativization strategies in African languages (with subjects)

Appendix I: Sample of African Languages

Anywa	(Nilotic, Nilo-Saharan)	Koozime	(Bantu, Niger-Congo, Niger-Kordofanian)
Amharic	(Ethiopic; Semitic, Afro-Asiatic)	Koranko	(Mande, Niger-Congo, Niger-Kordofanian)
Arabic	(Semitic; Afro-Asiatic)	Koromfe	(Gur, Niger-Kordofanian)
Ateso	(Nilotic; Nilo-Saharan)	Koyra	(Omotic, Afroasiatic)
Babungo	(Bantu, Niger-Congo, Niger-Kordofanian)	Koyraboro	(Songhai of Gao; Songhai, Nilo-Saharan)
Bagirmi	(Central Sudanic, Nilo-Saharan)	Krongo	(Kordofanian, Niger-Kordofanian)
Baka	(Ubangian; Niger-Congo, Niger-Kordofanian)	Kunama	(Chari-Nile, Nilo-Saharan)
Bambara	(Mande; Niger-Congo, Niger-Kordofanian)	Kxoe	(Khoisan)
Bari	(Nilotic; Nilo-Saharan)	Lango	(Nilotic, Nilo-Saharan)
Bole	(Chadic, Afro-Asiatic)	Lele	(Chadic, Afro-Asiatic)
Cibak	(Chadic, Afro-Asiatic)	Lingala	(Bantu, Niger-Congo, Niger-Kordofanian)
Dagbani	(Gur; Niger-Congo, Niger-Kordofanian)	Luganda	(Bantu, Niger-Kordofanian)
Dawuro	(Omotic; Afro-Asiatic)	Maale	(Omotic, Afro-Asiatic)
Dholuo	(Nilotic, Nilo-Saharan)	Maasai	(Nilotic, East Sudanic, Nilo-Saharan)
Dzamba	(Bantu, Niger-Congo, Niger-Kordofanian)	Mamvu	(Central Sudanic, Nilo-Saharan)
(Anglo)	(Togo; Niger-Congo, Niger-Kordofanian)	Nandi	(Nilotic, Nilo-Saharan)
Ewe	(Omotic; Afro-Asiatic)	Ngemba	(Benue-Congo, Niger-Kordofanian)
Gamo	(Chadic, Afro-Asiatic)	Northern	(Bantu, Niger-Congo; Niger-Kordofanian)
Gidar	(Chadic, Afro-Asiatic)	Sotho	(Chadic, Afro-Asiatic)
Giziga	(Fur, Nilo-Saharan)	Pero	(Songhai, Nilo-Saharan)
Fur	(Benue-Congo, Niger-Congo)	Sonrai	(Gur, Niger-Congo, Niger-Kordofanian)
Fyem	(Chadic, Afro-Asiatic)	Supyire	(Bantu, Niger-Congo, Niger-Kordofanian)
Hausa	(Cushitic, Afro-Asiatic)	Swahili	(Bantu, Niger-Congo, Niger-Kordofanian)
Harar	(Omotic, Afro-Asiatic)	Venda	(Bantu, Niger-Congo, Niger-Kordofanian)
Oromo	(Kuliak, East Sudanic, Nilo-Saharan)	Umbundu	(Bantu, Niger-Congo, Niger-Kordofanian)
Haro	(Bantu, Niger-Congo)	Wolayitta	(Omotic, Afro-Asiatic)
Ik	(Atlantic, Niger-Kordofanian)	Yoruba	(Kwa Niger-Congo, Niger-Kordofanian)
Kinyarwanda		Zulu	(Bantu, Niger-Congo, Niger-Kordofanian)
Kisi			

Appendix II: Database (African Languages) Subjects

Anywa	Gap	Koozime	Pronoun Retention
Amharic	Gap	Koranko	Correlative
Arabic	Gap	Koromfe	Gap
Ateso	Gap	Koyra	Gap
Babungo	Pronoun Retention	Koyraboro Senni	Gap
Bagirmi	Gap	Krongo	Gap
Baka	Gap	Kunama	Gap
Bambara	Correlative	Kxoe	Gap
Bari	Gap	Lango	Gap
Bole	Gap	Lele	Gap
Cibak	Other	Lingala	Gap
Dagbani	Gap	Luganda	Other
Dawuro	Gap	Maale	Gap
Dholuo	Gap	Maasai	Gap
Dzamba	Gap	Mamvu	Gap
(Anglo) Ewe	Gap	Nandi	Gap
Gamo	Gap	Ngemba	Pronoun Retention
Gidar	Gap	Northern Sotho	Gap
Giziga	Gap	Pero	Gap
Fyem	Gap	Sonrai	Gap
Fur	Gap	Supyire	Correlative
Hausa	Gap	Swahili	Gap
Harar Oromo	Gap	Venda	Gap
Haro	Gap	Umbundu	Gap
Ik	Gap	Wolayitta	Gap
Kinyarwanda	Gap	Yoruba	Pronoun Retention
Kisi	Gap	Zulu	Gap

Deictic categories in particles and demonstratives in three Gur languages*

Kézié Koyenzi Lébikaza

University of Lomé, Togo

The following assumptions on deixis offer some guidelines for the present paper. ‘The term ‘deixis’ (which comes from a Greek word meaning ‘pointing’ or ‘indicating’ is used in linguistics to refer to the function of personal and demonstrative pronouns, of tense and of a variety of other grammatical and lexical features which relate utterances to the spatio-temporal co-ordinates of the act of utterance’ (John Lyons 1977:636). As outlined by Levinson (2000:62), the main categories in the organization of deictic systems in human languages are person, place, time, discourse (or text) deixis and social deixis.

The present paper is concerned with the various manifestations and functions of the deixis in languages belonging to two subgroups of Gur languages, namely Kabiye and Tem (both Gurunsi languages) on the one hand, and Nawdm, a Mosi-Gurma language on the other hand. The investigation is focused on the deictic particles and the demonstratives, in order to explain how they come to operate at several syntactic levels in those languages to play out purely deictic functions as well as specific non-deictic functions.

1. Deixis and semantic functions

1.1 The categories of the local deixis in particles and demonstratives

1.1.1 *The deictic particles as demonstrative identifiers*

The following short utterances will give a cue to identify the different forms used to mark the two deictic distinctions (considered from the deictic center constituted by the speaker) in the three languages: proximal and distal. They have the status of demonstrative identifiers.

(a) The particles *yɔ* in Kabiye, *qɔ* in Tem, *na* in Nawdm indicate nearness to the deictic center and carry therefore the deictic meaning ‘proximal’. The particles *lé* in Kabiye, *mú* in Tem and *álé* / *álé* in Nawdm, express remoteness to the deictic center. This makes their deictic meaning ‘distal’ explicit. The two series of particles mark also distance, and the speaker who is the local reference of the utterance, constitutes the deictic center. (1)–(6).

Proximal	Distal
Kabiye	Kabiye
(1) <i>má'á yó</i> I Proximal 'Here I am'	(4) <i>ma-haló lé</i> my-wife Dist 'My wife is over there!'
Tem	Tem
(2) <i>ma-ná dɔ́</i> I-Cop Proximal 'Here I am'	(5) <i>mɛn-dɛlú wɛ ɲna mó</i> my-wife be Dist 'My wife is over there'
Nawdm	Nawdm
(3) <i>ma-fɔ́gá ní</i> my-wife Proximal 'Here is my wife'	(6) <i>ma-fɔ́gá be álé</i> my-wife be Dist 'It is my wife over there!'

(b) In the following utterances, these particles carry their respective deictic meanings 'proximal' and 'distal', but in a way that differs from the situation described in (a). The deictic center is no more the speaker's local point, but the coding time of the utterance. (7) and (8). So, in Kabiye, these particles have been grammaticalized to temporal clause particles.

- (7) Kabiye
agɔma taláá yó dɪ-háy wɛ lím kɪnídíím
guests arrive.PF PROX/as we-give.IPF them water fresh
'As the guests have arrived, we give them fresh water.'
- (8) *agɔma ká-taliy lé sóná pɪwá*
guests FUT.arrive.IPF DIST/when beans cook.PF
'When the guests arrive, the beans will be cooked.'

Furthermore, the distal particle *lé/lé* at the end of a Kabiye clause-sentence has another specific temporal function: It indicates that the event that takes place at the coding time was also expected for that time. So it expresses the already known event time.

- (9) a. *kɔɔ dɔ́-yɔ́dɪ lé*
come.IMPER we-speak TEMPORAL
'The expected moment has come that we talk.'
- b. *mɛn-dɛ-y sukúli 'lé*
I-go-IPF school DEIC
'I am going to school as expected.'

Whether for the local or the temporal deixis, I assume that the encoding of the distance distinctions is the basic functions of the deictic particles of these languages. But it must

be noted that they display diverse syntactic features. They can occur as predicate-word in a verbless sentence, the case for proximal particles ná in the three languages (1)–(3) and also for the distal particle lé of Kabiye (4). The distal deictic particles mó in Tem (5) and álé in Nawdm (6) occur as predicate-words after a copula. Furthermore deictic particles are found in the role of temporal clause markers with an embedding property (7) and (8); (also see 2.1.1), or just as deictic morphemes within demonstratives (see 1.1.2). Finally, as it will appear later, they have other values at syntactic as well as pragmatic levels.

Ben (Moba), another Gur language uses the particle *nyan* that serves just as deictic marker. Another particle, *daa*, displays the role of a distance marker ‘distal’ and is associated with the deictic marker *nyan* for the local deixis ‘distal’. The deixis marker in absence of the distance marker expresses the proximal local deixis. The distinction between nearness and farness is made clear through the use of *nyan* vs. *daa nyan*. So there is no overt distance marker for proximal, the speaker’s location and the lack of the distal marker *daa* imply the proximal category in Ben.

1.1.2 *The demonstratives and the local deixis*

In Kabiye, Tem and Nawdm, the structure of demonstratives shows that their deictic properties are explicitly marked. Each demonstrative is made of a deictic morpheme associated with an anaphoric pronoun referring to the nominal class of the determined noun. The anaphoric pronoun establishes then the grammatical concord with the two related units, the head noun and the demonstrative.

The demonstratives constitute a complex system based primarily on the distinction LOCAL/SPATIAL vs. TEMPORAL. The local and the temporal deixis are divided into the proximal and distal subcategories. The three languages have for the local proximal, *-na* as deictic morpheme, which the anaphoric pronoun is prefixed (10a)–(10c). But only Kabiye and Nawdm mark the distance opposition in the local deixis by using different demonstratives for proximal and distal. Kabiye uses an attributive distal demonstrative made of a homorganic nasal consonant *N-* that is prefixed to the anaphoric pronoun of the determined noun; and the anaphoric pronoun precedes the suffixed deictic morpheme *-yɔ* (11a). Nawdm has a distal demonstrative as distance marker *a-* that is prefixed to the anaphoric pronoun (11c).

The particles *yɔ* in Kabiye and *dɔ* in Tem that are combined with the demonstratives, have no longer their proximal meaning. That means, they do not serve as distance marker when they appear in association with the demonstratives. The particle *yɔ* in the distal demonstrative of Kabiye is obligatory, but serves only as complementary deictic morpheme. Every demonstrative, even without any deictic particle, is marked for distance.

1.2 The categories of the temporal deixis in particles and demonstratives

The following definition of time deixis by Levinson (2000:62) gives the most important aspects of time deixis: ‘Time deixis concerns the encoding of temporal points and

	Local Proximal	Local Distal
Kabiye	(10) a. <i>lím pU-né</i> (yó) water Pro-LocProx (Deic) 'this water'	(11) a. <i>lím m-bó-yó</i> water LocDist-Pro-Deic 'that water'
Tem	b. <i>lím bɪ-nà</i> (qó) water Pro-LocProx (Deic) 'this water'	b. = (10b).
Nawdm	c. <i>nyá:lm mí-ná</i> water Pro-LocProx 'this water'	c. <i>nyá:lm mí bé á-mí</i> water Pro be LocDist-Pro 'that water. ...'

spans *relative* to the time at which an utterance was spoken (or a written message inscribed). This time, following Fillmore (1971), we shall call coding time or CT, which may be distinct from receiving time or RT, ... Thus, just as place deixis encodes spatial locations on co-ordinates anchored to the time of utterance. Time deixis is commonly grammaticalized in deictic adverbs of time (like English *now*, and *then*, *yesterday* and *this year*), but above all in tense'. As it seems to be the tendency in works on deixis to my knowledge, this definition ignores the demonstratives as one of the domain of expression of time deixis. The investigated Gur languages will give evidence for another view about the relation between time deixis and demonstratives.

In the literature on deixis, the temporal deixis has been generally studied in the domains of adverbs and verbs. And the local deixis has been considered as being the only deictic domain of demonstratives. This appears clearly in the following assumption made by Diesel (1999:36): 'Demonstratives are place (or spatial) deictics'. Such a view concerning demonstratives cannot be true for every language. In former works on Kabiye, I have assumed that the description of demonstratives requires a spatio-temporal frame, distinguishing demonstratives of the local deixis from those of the time deixis (Lé bikaza 1985, 1999). This paper will show that, apart from the demonstratives of the local deixis (1.1.2), Tem has temporal demonstratives falling into two distance scales, proximal and distal and even Kabiye adds to these scales, a third temporal dimension that should better be labeled as 'actual'. For the temporal deixis 'actual' expressed by the demonstrative constituted by the reduplicated anaphoric pronoun in Kabiye, the referred time matches with the coding time (13b). The indexical value does not refer to a distance on the time axis rather to the time point zero, the coding time, which serves as point of departure to fix the distinguished distances 'proximal' and 'distal'.

The proximal temporal demonstrative in Kabiye has the structure *Nasal as distance marker plus Anaphoric pronoun*. The corresponding form in Tem is: *Anaphoric pronoun plus Distance marker lé* Such demonstratives in Kabiye and Tem refer to the time preceding the coding time (but may coincide with the moment of the act of communication).

The structure of the distal temporal demonstratives consists of an anaphoric pronoun and the morpheme for remote distance; in Kabiye: *Anaphoric pronoun* plus *lé*; in Tem: *Anaphoric pronoun* plus *mú/-m*. Distal temporal demonstratives mark the referred time as remote from the coding time.

So it appears in the two Gurunsi languages a clear set of demonstratives whose semantic meaning matches with the time deixis encoded differently according to different temporal distinctions. At pragmatic level they behave also differently from the demonstratives of local deixis as only they can get a recognition function (3.2).

The Mosi-Gurma language, Nawdm, shows different typological features for the corresponding temporal demonstratives in Kabiye and Tem, Nawdm displays only one type of demonstratives referring to the common shared knowledge. In Nawdm such demonstratives are made of the anaphoric pronoun that a vowel *-i* is suffixed. They refer to the time prior the coding time, but without any notion of distance. We have then to do with referential demonstratives that have in common with temporal demonstratives of Kabiye and Tem, the value of ‘known, definite’ (12).

(12) Nawdm

dám mi-i

beer PRO-DEFINITE

‘The beer in question’

Temporal Actual

(13) a. Kabiye

kóná-m kpelaý cíkpelaý

bring-me chair little

‘Bring me the little chair!’

A possible reaction to (13a) is:

b. *kpelaý káká peláa ‘lé*

chair DEM.ACT broke PRES

‘The chair (you are talking about) is broken!’

Temporal Proximal

(14) a. Kabiye

lím mbó

water TP.DIST.DEM

‘This water (that has been talked about some time ago)’

b. Tem

lím brín

water TP.DIST.DEM

‘This water (that has been talked about some time ago)’

Temporal Distal

- (15) a. Kabiye
 lím pólé
 water TP.DIST.DEM
 ‘That water (that had been talked about a long time ago)’
 b. *lím brín*
 water TP.DIST.DEM
 ‘That water (that had been talked about a long time ago)’

1.3 Time deixis and calendar expressions

The nouns that designate parts of calendar divisions can be determined by the proximal or distal demonstratives within nouns phrases. For noun phrases expressing the proximal deixis refer to an event time, which is located in the span of the coding time; the calendar nominal is determined by a proximal demonstrative, which is taken from the local deixis. When the calendar expression gives an event time, the calendar nominal is determined by a proximal demonstrative of the temporal deixis.

One gets following examples in Kabiye:

Span of the coding time (Actual)

- (16) a. *tanaŋ koné* ‘this morning’
 ɖoó ané ‘this night’
 fenaý kané ‘this month’
 pínáý kané ‘this year’
 yolím poné ‘this rainy season’

Out of the span of the coding time

- b. *tanaŋ ígú* ‘that morning’
 ɖoó áná ‘that night’
 fenaý ígá ‘that month’
 pínáý ígá ‘that year’
 yolím mbú ‘that rainy season’

The proximal demonstratives of the local deixis determine calendar nouns to express the event time that coincides with the coding time or is located in the reference time span, that is a referred time of the time category ‘actual’ (17) and (18).

- (17) Kabiye
 cíca talí-na Lome tanaŋ ku-né
 teacher arrive-PF.DIR Lome morning PRO.PROX
 ‘The teacher has arrived from Lome this morning’

- (18) *Kao wóki egbándu wísi si-né*
 Kao go.IPF hunting sun PRO-PROX
 'Kao is going hunting this afternoon.'

If the event time is different from the coding time or the reference time span, which is located in the past or the future, it will be referred to by the use, as determiner of the calendar noun, of the proximal demonstrative of the time deixis (19)–(20).

- (19) *dá-caa talá hódo wiye dānáy yɔɔ.*
 our father arrive.PF Monday day evening on
 'Our father arrived Monday evening.'

dānáy ígá é-kpéélí dǔ
 evening TP.PROX.DEM he-bring together.PF us
 'That evening he brought us together.'

- (20) *fáda káy kóm kujuká wiye dānáy yɔɔ,*
 priest FUT come Sunday day evening on
 'The priest will come on Sunday,

né é-lá cɔ́cɔ́ dānáy ígá kudɔndmáa yɔɔ.
 and he-do.AOR church evening TP.PROX.DEM sick persons on
 and celebrate a service for sick people.'

Demonstratives of the category proximal of the local deixis and those of time deixis are employed with non-deictic words to express different time relations. A possible explanation for the correlation of the local proximal category with the event time that matches with the coding time (or located in the reference time span) is certainly the more concrete character of this temporal relation. Similarly remoteness from the coding time point can be considered as more abstract and therefore may correlate with the more abstract character of the temporal demonstratives. We are facing a concrete situation where time and place show their complicated relations. This is not a language specific but rather a universal problem. Levinson (2000:72) outlined with right 'Both time and place are greatly complicated by the interaction of deictic co-ordinates with non-deictic conceptualization of time and space'.

Another specific kind of time reference is realized by distal demonstratives of the time category, which determine a calendar noun. The so constructed nominal phrases refer to a time, which is located in the past with the features 'known, definite' and is related to a known event. This is why these calendar terms have a recognitional use (21).

- (21) *pínáy kálé pɪ-lab-á ezímá té*
 year TP.DIST.DEM it-occur-PF how Q-FOCUS
 'What did occur that year?'

2. The syntactic properties of deictic particles and demonstratives

2.1 Deictic particles and subordination

2.1.1 *Deixis in temporal clauses*

In the three languages, Kabiye, Tem and Nawdm, the temporal clauses are marked by deictic particles and they precede always the matrix clause. This fixed syntactic linearity reflects the linearity on the time axis, then the event in the temporal clause precedes necessarily the event in the matrix clause. The proximal and distal particles, which have been identified in the preceding section, get new functions: they indicate a temporal relation and function syntactically as clause particles. The following Kabiye utterances are illustrations for the proximal particle *yɔ́* and the distal one *lé* as temporal clause markers as well as subordinating particles.

(22) Kabiye

nyé-píya wí-y yɔ́ nn-ní-ý kée
 your-children cry-IPF PROX/as you-NEG-hear-IPF Q-FOCUS
 ‘Don’t you hear your children crying?’

(23) *wíyau kɔm-á yɔ́ dí-la ezímá*

chief come-PF PROX/as we-do what
 ‘Now the chief has come, what should we do?’

(24) *dómáý táli-y yɔ́ n-nyɔɔz-á nyá-hayím na?*

seed-time arrive-PF PROX/as you-prepare-PF your-field Q
 ‘As the seed-time is approaching now, have-you prepared your field?’

(25) *wíyau kɔm-á lé í-ta-há-í lim sewe?*

chief come-PF DIST/when you-give-PF-him water Q
 ‘Why didn’t you give the chief water when he arrived?’

(26) *nyé-píya ká-kɔ lé dɔ́-yɔɔdɪ tɔ́m*

your children FUT-come-AOR DIST we-JUSS-speak matter
 ‘We will discuss the matter when your children are there.’

In these utterances the temporal clauses show that the proximal particle *yɔ́* views the event in the embedded clause as being near the reference point, which is the coding time. The nearness can be interpreted as coinciding with the coding time (22), or in direct proximity of the coding time (22), preceding (23) and (24) or following (26). The distal particle *lé* expresses remoteness on the time axis and puts the time of the event in the embedded clause far from the coding time (utterance time). The time of the event as expressed by *lé* may refer to past (25) or future (26). The time of the event is just put far away from the reference time-point. Therefore it must be pointed out that the temporal values of the deictic particles, unlike the oppositions we have by tempus morphemes, do not refer to tempus distinctions past, present and future but rather nearness or remoteness from the referring point of time.

In Kabiye these particles have been grammaticalized to clause markers that still continue to express a relative deictic opposition proximal vs. distal. On the other hand in Tem, only the particle *ná*, that has been given a proximal value in the local deixis, has been grammaticalized to a clause marker, but without any deictic opposition ‘proximal’ vs. ‘distal’.

(27) Tem

ábóní wɔɔ-góni ná dǎ-ń-lám ɲíní?
old lady PF-COME PROX/as WE-IPF-do how
‘As the old lady has come, what should we do?’

(28) *pídé wíre ábóní wɔɔ-góni ná ɲíní ge nyáá-lá?*
it is day old lady PF-COME PROX/as how FOCUS you-PF.do
‘What did you do that day when the old lady came?’

In Kabiye and Tem, the deictic particles do not only give a temporal relation in reference to the coding time, but they locate an event as part of the utterance within the whole utterance, and serve as articulator between propositions. At syntactic level this has as result to treat the concerned utterances as embedded clauses within a sentence instead of an independent sentence.

The deictic particles have not only been grammaticalized to temporal clause markers as already attested in Kabiye and Tem, but they appear also in adverbial expressions, for example in Kabiye, for temporal reference, *piwayí* ‘lé ‘after that’, *púcǎ...lé* ‘before’, or for comparison, *ezi ... yǎ* ‘like’. Further discussion will be needed to capture other syntactic functions of the particles in Kabiye and Tem (2.1.2).

2.1.2 Deixis and relativization

Another question that arises in syntactic structures where deixis is involved, in the investigated languages, concerns the domain of the relative clause.

The relative clause follows the antecedent, and is embedded between the relative pronoun, which begins the clause, and the proximal deictic particle, which ends the clause, as *yǎ* in Kabiye and *ná* in Tem. The following examples of Kabiye show that there is a strategy of relativization that consists in transforming a sentence into a clause. The proximal particle *yǎ* appears together with the relative pronoun. The latter functions as relativizer at the beginning of the relative clause whereas *yǎ* appears at the end of the clause and embeds the clause within the matrix clause. Finally the proximal deictic particles in these Gurunsi languages serve to transform a whole clause to a determiner.

(29) Kabiye

kpelaý wengá pa-há-m yǎ, ké-peláa
chair REL they-give-PF DEIC PRO-break-PF
‘The chair they gave me is broken.’

(30) Tem

kpeɔ́ wɛ́ŋá pa-vá-má ná, wɛɛ-pɛ́lɪ.
 chair REL they-gave-PF-me DEIC PF-break
 ‘The chair they gave me, is broken.’

Nawdm, the Mosi-Gurma language, embeds the relative clause between a relative morpheme and the anaphoric pronoun referring to the antecedent NP. It does not make use of any clause particle.

2.2 Determining demonstratives and concord

At morphosyntactic level in the three languages, demonstratives are determiners and they follow the head noun and assume the attributive function. When there are other determining constituents (adjectives, numerals), the demonstrative still takes the final position in the nominal phrase. In a nominal phrase that has a demonstrative as the determining constituent, it is even possible to have a proper name as head noun (32a)–(34).

(31) Kabiye

yalá cikpéná áné (yɔ́) a-wɛ líidiyé
 eggs little those (PROX) they-be money
 ‘Those little eggs are expensive.’

(32) a. *Naka ɛ-né, ɛbɛ lákɪ-í té*
 Naka PRO-PROX what do-IPF-her so
 ‘As for Naka in question, what is wrong with here?’

b. *Naka únú ɛ-wɛ lé-yɔ́?*
 Naka TP.PROX she-be where-DEIC
 ‘Where actually is Naka (who has been spoken about)?’

c. *Naka élé ɛ-kɔm-á kɛ́yɔ́*
 Naka PRO.TP.DIST she-come-PF Q-FOCUS
 ‘Has Naka (who had been talked about) come now?’

(33) Nawdm

Naka wéé-ná bɛ té?
 Naka PRO-PROX be where
 ‘Where is Naka (it has been talked about)?’

(34) Tem

Abu cee-ní
 Abu PROX
 ‘Abu (has been talked about)’
 vs.
Abu mó
 Abu DIST
 ‘Abu (had been talked about)’

Through its nominal constituent, the demonstrative takes the class concord with the determined head noun, which can be a substantive or a proper noun. Furthermore the demonstrative in the noun phrase provokes the concord of the verb with the subject or object nominal phrase: the verb then takes the prefixed anaphoric pronoun when the NP assumes the subject function and the suffixed anaphoric pronoun when it is an object NP.

2.3 Deixis and predicative function

In Kabiye as well as in Tem and Nawdm, there are several semantic types of determining demonstratives in both local and temporal domains, with specific pragmatic functions.

If we consider the examples (1)–(6) in the Section 1.1.1, we realize that no verb is needed in these sentences. The proximal deictic particles *yɔ́* (Kabiye), *ná dɔ́* (Tem) and the reduplicated proximal particle *náná / ní* (Nawdm), act as predicates. The distal deictic particle *lé* together with the demonstrative *mbv* has these predicative properties in the following example (35). Through their capacity to assume a predicative function deictic morphemes (particles and some demonstratives) may enable a verbless construction to have the status of a sentence.

- (35) Kabiye
 nyé lé mbú
 you DIST.TP DIST-G6
 ‘You are so.’

3. Deixis and pragmatic functions

Concerning the pragmatic functions of deixis, it will be interesting, in the three languages to investigate their use for referential functions, as well as focus markers, and finally their capacity of encoding social functions.

3.1 Deixis and reference

The demonstratives of the spatio-temporal domain are employed to express several referential and pragmatic functions. Depending on whether they are semantically analyzed as local or time deixis they are differently encoded and assume different kinds of referential functions from which the exophoric use is one of the most important one.

‘Exophoric demonstratives focus the hearer’s attention on entities in the situation surrounding the interlocutors. They have distinctive features: first, they involve the speaker (or some other person) as the deictic center; second, they indicate a deictic contrast on a distance scale (unless they belong to the small minority of demonstratives that are distance neutral...); and third, they are often accompanied by a

pointing gesture.’ (Diesel 1999:93). In the languages concerned dealt with here, the so defined exophoric demonstratives match with the demonstratives of the local deixis distinguishing the distance scale ‘proximal’ and ‘distal’ (36a) vs. (36b).

- (36) a. Kabiye
háSI SI-né (yɔ) SI-nyási-ɣ eyáa
 dogs PRO.PROX DEIC they-bite-IPF people
 ‘These dogs bite people.’
 b. *háSI ín-zí-yó ...*
 dogs LOC.DIST-PRO-DEIC
 ‘Those dogs over there...’

According to the inherent syntactic features of exophoric demonstratives, they appear in an endocentric construction as determiners. Though such demonstratives can be used without the determined noun, behaving as anaphoric pronouns but still retaining some deictic features, so far as the speaker can point at the referent.

- (37) Kabiye
ene (yɔ) e-we nyáŋ
 this DEIC he-be wickedness
 ‘This one (this person) is wicked’

The following examples from Kabiye and Tem (40a)–(41) make it clear that the pronominal anaphoric function that is assumed by the simple anaphoric pronouns can be displayed by the proximal temporal demonstratives. But in this case such demonstrative pronouns can represent only a focused subject or object NP, (38) vs. (39).

- (38) Kabiye
n-ta-yáa nyé-neyáa se we?
 you-NEG-call.AOR your-youngster that what
 ‘Why have you not invited your youngsters?’
 (39) *mbá kizi-na se paa-kɔŋ*
 TP.PROX.DEM refuse-PF.SUGJ.FOCUS that they-NEG-COME-IPF
 ‘They have refused to come.’
 (40) **mbá kizaá se paa-kɔŋ*
 TP.PROX.DEM refuse-PF that they-NEG-COME-IPF
 (41) Tem
belé bee-gizi-ná
 TP.PROX.DEM they-PF-refuse-SUBJ.FOCUS
 ‘They have refused.’

In the Kabiye language it is the demonstrative of the temporal deixis ‘actual’ that is absent from the deictic systems of Tem and Nawdm, which is profoundly involved in the process of grammaticalization through which it comes to function as a specific anaphoric. It does not serve only as substitute to a nominal phrase, but it allows

also avoiding of referential ambiguity when the anaphoric pronoun is preceded by more than one nominal phrase. In fact, when the demonstrative for ‘actual’, for example *káká*, gets an anaphoric use, it represents the directly preceding noun phrase, that means the last one (46). So it is opposed to the simple anaphoric pronoun, for example here *ka-*, that occurs when instead of the last NP, the first NP is substituted (45). It is also used when there is a kind of intensified anaphoric representation of the referent of the substituted NP (43) and (44). In the example (43), it is possible to have *káká* or *ka-*, that is the demonstrative for ‘actual’ or the simple anaphoric pronoun as substitute. So it is clear that both the demonstrative ‘actual’ and the simple anaphoric pronoun have the same origin, the nominal class marker that indicates the nominal class of the substantive.

- (42) a. *Kabiye*
kóná kpelaʔ kisémaʔ wɛŋɡá ma-yab-á-ŋ yɔ́!
 bring.IMPER chair red REL.PRO I-buy-PF-you PROX
 ‘Bring the red chair that I bought for you!’

A possible reaction to (42a) is:

- (42) b. *kpelaʔ káká pɛ́láa n-ta-ná kée?*
 chair TP.PROX.DEM break-PF you-NEG-see-AOR Q.FOCUS
 ‘That one (you are talking about) is broken, don’t you know?’

Another possible reaction to (42a) is:

- (43) *káká pɛ́láa n-ta-ná kée?*
 TP.PROX.DEM break.PF you-NEG-see-AOR Q.FOCUS
 ‘That one (you are talking about) is broken, don’t you know?’
- (44) *ɔ́a-caa nyíni-ɐ kpelaʔ kisémaʔ kɛŋɡú káká pɛ́láa*
 1P-father look for-IPF chair red but TP.PROX.DEM break.PF
 ‘Our father looks for the red chair but that one is broken.’
- (45) *pɛ́láa káta evebíya né pé-héyi-wéʔ tóm kibánda*
 girls meet.PF young boys and they-tell.AOR-them news good
 ‘The girls met young boys and told them good news.’
- (46) *pɛ́láa káta evebíya né pába-héyi-wéʔ tóm kibánda*
 girls meet.PF young boys and they-tell.AOR-them news good
 ‘The girls met young boys and the latter told them good news.’

3.2 Recognitional use of temporal demonstratives

In the light of cross-linguistic works on demonstratives, especially Diesel (1999), it is clear that the demonstratives I described as demonstratives of temporal deixis, have when playing their pragmatic function, a recognitional use. In fact the examples (47) and (48) show that these demonstratives used always adnominally, and displaying also

a recognitional function, are not only connected with ‘definiteness’, but in Kabiye and Tem languages they are related with time deixis. In the two languages, it is specified whether the referent has become a shared knowledge a short or a long time ago. There is therefore a formal distinction between the categories ‘proximal’ and ‘distal’ on the time axis. So the recognitional function is based on the temporal value of such demonstratives.

‘The shared knowledge’ that characterizes the recognitional function is due to the fact that the referent is known by the speaker and the hearer because it has been the topic of a communication at a previous time. But if both the proximal and the distal temporal demonstratives can have this function, only the distal one can be employed with an additional shared feeling of the interlocutors about the referent (48). As only the speaker and the hearer know the reason of their common opinion about the referent, they may show a kind of complicity against the referent. This confirms partly what Diesel (1999: 107) observed quoting Lakoff who used the term ‘emotional deixis’, in fact ‘recognitional demonstratives are often used to indicate emotional closeness, sympathy, and shared beliefs’.

(47) Kabiye

haláa mbá pɔ-kóŋ kelé
 women those.TP.PROX they-come.IPF PRESENTATIVE

‘Those women (it was talked about, a short time ago) are coming now.’

(48) *haláa pálé pɔ-kóŋ kelé*

women those.TP.DIST they-come.IPF PRESENTATIVE

‘Those women (it was talked about, a long time ago) are coming now.’

It is also possible for the proximal or distal temporal demonstratives, which get a recognitional use, to determine a proper noun (49).

(49) *Naka únú ɛ-we lé té*

Naka that.TP.PROX.DEM she-be where Q.FOCUS

‘Where is Naka it has been talked about?’

3.3 Discourse deixis

Some studies on deixis have outlined that the specific characteristics of discourse deixis lie in the fact that it does not refer to a text, but rather serves as an overt link between two propositions (Lyons 1977), while referring back to an element of the preceding discourse and anticipating an upcoming information (Diesel 1999). The description viewing temporal demonstratives as text deictics in Kabiye (1985, 1999) must be revised. The demonstratives for temporal deixis have an adnominal use carrying the concord with the determined noun. Not all of them can have a discourse deictic use. Discourse demonstratives like *mbú lé* ‘So...; That is...’ appear as linking between two utterances (51)–(56), the discourse demonstrative *ndú’ó yó* ‘In fact...’ is employed

for reminding preceding information as transition to the upcoming information (54), about the content or the following of the old information (54), *ńdú lé* 'That's all' used for closing an utterance (52), *ńbú elé* 'So...' (52). Although they have the structure of temporal determining demonstratives, discourse demonstratives are not used as determiner within a nominal phrase with any determined noun. For discourse deictic functions such demonstratives occur as unvarying forms in Kabiye (52a)–(55) and Tem (56) where they have developed from temporal demonstratives: the forms *mbú* and *ndú* in these demonstratives are respectively the determining demonstratives referring to nominal classes whose main characteristic is to be made of nouns that refers to abstract concepts. This is obvious in the case of *ńdú lé* developed from the elliptic use of *tóm ńdú lé* 'That is the matter'.

(50) Kabiye

nyá-caa se má-yá-ń
 your-father that/COMP i-call-AOE-you
 'Your father told me that I call you.'

(51) *ń-bú kée?*

TP.PROX-PRO6 Q.FOCUS
 'So?'

(52) a. *ń-bú 'lé*

TP.PROX-PRO6 DIST
 'That is so.'

b. *ń-bú elé*

TP.PROX-PRO6 DIST
 'So, ...'

(53) *ń-bú món-sóólaá*

PROX-PRO6 I-like.PF
 'That is what I like.'

(54) *ń-dú yó dóóyé nyá-tumíye tén kée?*

PROX-PRO6 PROX when your-work finish.IPF Q
 'In fact, when will your work finish?'

(55) *tótú labí wé? pítíña pi-káv nyɔɔzúu.*

TP.ACTUAL do.PF what all that PRO6-FUT repair
 'What has happened? (That has no importance.) Everything will be alright again.'

(56) Tem

bɪ-lé yaá

PRO-TP.PROX Q
 'So?'

3.4 Deixis and focus

In the three languages the proximal deictic particles display other functions than those which are typical to them. So they occur at the end of a nominal phrase and put it in focus. In a verbless utterance the proximal particles *yó* (Kabiye), *dó* (Tem), *ná* (Nawdm) express the presentative focus assuming also a predicative function (1)–(6). Apart from these values attested in the three languages, the deictic particle *yó* in Kabiye expresses also the contrastive focus in Kabiye (57) at clause level opposing the referent to an element or to a group. The distal particle *lé* follows a constituent as a contrast-to-the-whole marker, opposing the referent as specific element to the whole (58).

(57) Kabiye

mo-wó-ki kiyaku, nyé-yó, n-lá-kí we?
 I-go-IPF market you-PROX you-do-IPF what
 ‘I go to the market, and you, what are going to do?’

(58) *nyá-caa lé, ε-sɔɔl-á agɔma.*
 your-father DIST he-likes-PF foreigners
 ‘As for your father, he likes foreigners.’

The deictic particles *né* (proximal) and *lé* (distal) play an important role in several focus constructions, as focus marker or as part of the focus marker together with a copula. The deictic particle *ná* is suffixed to the verb and serves as subject focus marker in Kabiye and Tem.

(59) *nyé-píya kizí-na tón*
 your-children refuse-PF.SUBJ.FOCUS hones
 ‘Your children have refused the honey.’

(60) Tem

máá-mó-ná kembízi síná
 IPF-buy-SUB.FOCUS chicken these
 ‘I bought these chicken.’

The predicate focus is expressed in Kabiye by the form *kéné* that follows the predicate (61). The focus marker is then constituted with the proximal deictic particle *né* and the copula *ké*. The presentative focus affects the whole sentence. Its focus marker consists of the copula *ké-* and the distal particle *-lé*, that is *kélé*, appears at the end of the sentence (62).

(61) Kabiye

píya léyí-y kileeyú kéné
 children play-IPF ADVERB FOCUS
 ‘The children are just playing.’

- (62) *man-talaá kélé*
 I-arrive PF.FOCUS
 'I have indeed arrived.'

3.5 The social functions of the deixis

Neither the Kabiye language, nor Tem, or Nawdm expresses social distance through forms of personal pronouns. However there is an interesting use of demonstratives in utterances expressing thanks, a certain mark of respect, distance or disrespectfulness or despise. So the addressee or referent can be subject to these kinds of considerations through specific use of deictic expressions.

In terms, which require respect the head noun that is a honorific term is determined by a proximal local (63a) or a distal temporal demonstrative (64a). In Kabiye such an expression may be used for the complete form of respectful thanks (63b) and (64b). For simple thanks *n-lab-á lé* in (63a) and (64a) is the normal form.

Expression of thanks with respect in Kabiye:

- (63) a. *man-caa ε-né, n-lab-á lé.*
 my-father PRO.LOC-PROX you-do-PF DIST
 'Thank you, respectful!'
 b. *man-caa ε-né,*
 my-father PRO.LOC-PROX
 'Thank you, respectful!'
- (64) a. *man-caa ε-lé, n-lab-á lé.*
 my-father PRO.TP.DIST you-do-PF DIST
 'Thank you, respectful!'
 b. *man-caa ε-lé,*
 my-father PRO.TP.DIST
 'Thank you, respectful!'

In some contexts there is a clear social motivation for the use of the proximal *vs.* the distal deixis. The contrasting use of the proximal and distal particles appears in the expressions in which a person is designated by noun for insult or praising. The proximal particle occurs in the insult (65) and the distal one in the praising (66). Further negative or even insulting expressions with this structure are attested, for example in Kabiye (67) and (68) as well as in Tem (69).

- (65) Kabiye
háɣ ka-né
 dog PRO.LOC.PROX
 'You, despicable person!'
- (66) *tóóyíw kólé*
 lion PR.TP.DIST
 'You, brave person!'

- (67) *kimeleŋ* *kuné*
 stupid.person this
 ‘What a stupid person!’
- (68) *afélaa* *pané*
 soul.eaters these
 ‘What for nasty people!’
- (69) Tem
fáa ceŋga
 dog this
 ‘What a despicable person!’

In Kabiye, expression of a request with insistence (for persons of equal status), or disrespectfulness or desire of humiliation (when speaker and hearer are from different social status, are unknown for each other): by putting the proximal deictic particle *yó* at the end of the imperative sentence. So the use of the proximal particle of the local deixis correlates with the lack of respect by a request that is then interpreted as commandment.

- (70) Kabiye
kóná-m *mo-hulaŷ*
 bring-me my-hat
 ‘Bring me my hat!’
- (71) *kóná-m* *mo-hulaŷ* *yó*
 bring-me my-hat DEIC
 ‘Bring me my hat!’ (I order you!)

There is another strong argument for assuming a correlation between the distance marker ‘proximal’ and the expression of less consideration. The Kabiye language possesses a reduplicated form of the proximal particle *né*, that is *nné* preceded by the noun class anaphoric. The head noun which is known or can be identified by the hearer is deleted from the utterance. The whole nominal phrase is reduced to the proximal demonstrative followed by the proximal particle. They are used with a pointing gesture as mark of despise. Compare the following examples in Kabiye (72)–(73), Tem (74) and Nawdm (75).

- (72) *abalú* *ene* *yó* *ń-sev* *mbí* *yó’ɔ?*
 man this DEIC you-fear IPF so
 ‘Is that this man that you are afraid of?’ (Just with surprise)
- (73) *ε-nne* *yó* *ń-sév* *mbí* *yó’ɔ*
 PRO-PROX.LOC PROX.LOC.DEIC you-fear IPF so
 ‘Is that this man that you are afraid of?’ (With great despise)
- (74) Tem
Cèèní *ge* *nyé-ń-zée*
 ‘Is that this man that you are afraid of?’ (Without respect!)

(75) Nawdm

wéé-né ben sǒlí?

PRO.PROX you fear

‘Is that the one you fear?’ (With despise!)

Conclusions

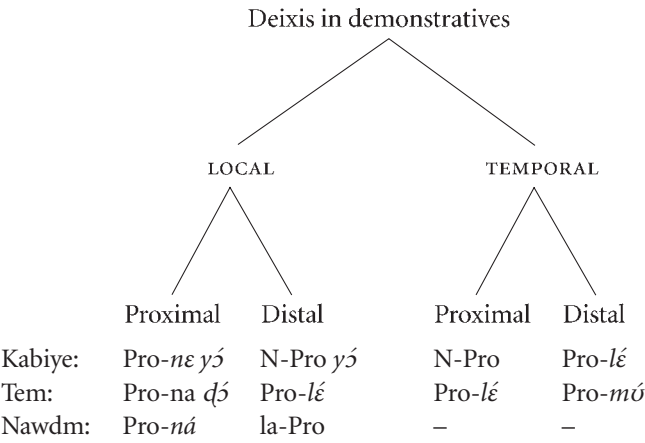
The two schemata give an overview of the deictic system of the three Gur languages Kabiye, Tem and Nawdm discussed in this paper.

Schema 1:

Deictic particles								
Proximal					Distal			
	Kabiye		Tem		Nawdem	Kabiye	Tem	Nawdem
	<i>ná</i>	<i>ɣɔ́</i>	<i>ná</i>	<i>dɔ́</i>	<i>ná/ní</i>	<i>lé</i>	<i>lé</i>	<i>álé/lé</i>
LOCAL	+	+	+	+	+	+	+	+
PROXIMAL	+	+	+	+	+	–	–	–
DISTAL	–	(–)	–	–	–	+	+	+
TEMPORAL	+	+	–	–	–	+	+	–
PREDICATIVE	+	+	+	+	+	+	+	+
SUBORDINATIVE	–	+	+	–	–	+	–	–
PRESENTATIVE	–	+	+	+	+	+	–	+
CONTRASTIVE	–	+	+	–	+	+	–	+
SUBJECT FOCUS	+	–	+	–	–	–	–	–

In the Gurunsi languages, Kabiye and Tem, the systematic opposition of the forms of demonstratives of local deixis to those of temporal deixis makes clear that these languages do not need to derive the expressions of time from those of space. They have demonstratives that are differently encoded not only for distance as noted for several languages, but also for space and time.

Schema 2:



Although one can say that Kabiye has the most complex deictic system, it shares with Tem the same typological features: their deictic system has basically an explicit encoding for the local deixis and another one for the time deixis so that there is no evidence to derive the temporal functions from the local ones. The syntactic and pragmatic functions of the particles and demonstratives show that, it is the deictic morphemes carrying out a temporal value that have been also grammaticalized to syntactic markers (clause markers). Another feature that characterizes these languages is the fact that the recognitional function is connected with the temporal deixis whereas particles and demonstratives of local deixis express also focus or forms of social deixis that associates the categories ‘proximal’ with lower consideration and ‘distal’ with higher consideration.

The temporal demonstratives of the category ‘actual’ in Kabiye dissociate themselves from other demonstratives: they don’t contain any segmental deictic morpheme. The reduplicated morpheme is identical to the class anaphoric that functions also as anaphoric pronoun or concord morpheme. As general common feature of the investigated Gur languages, all other demonstratives consist basically of a deictic morpheme (demonstrative morpheme) that is added a class anaphoric for concord with the determined noun. This structure enables such grammatical units in these languages to behave as demonstrative pronouns. Furthermore it appears that there are different origins, for on the one hand, the deictic morphemes (particles and demonstrative morphemes) that belong to the basic units of the grammatical system of the investigated Gur languages, and on the other hand the anaphoric pronouns, which have evolved from the noun class markers that appear in substantives. Another important common feature that characterizes these Gur languages is the attributive use of temporal demonstratives and other demonstratives in recognitional function in noun phrases that can have a proper noun as head noun.

The Mosi-Gurma language, Nawdm, belongs to a typological system that is different from that of Kabiye and Tem. The absence of temporal demonstratives in Nawdm

and the lack of grammaticalization of the deictic particles to clause markers offer sufficient arguments for this. Furthermore one can argue that, contrary to Nawdm, the encoding of time distinctions through particles and demonstratives in Kabiye and Tem has been in favor of grammaticalizing the proximal particles to clause markers.

Abbreviations

1p, 2p, 3p	1st, 2nd, 3rd person plural	LOC	locative
AOR	aorist	N-	prefixed homorganic nasal consonant
COP	copula	NEG	negation
DEIC	deictic morpheme	PF	perfective
DEM	demonstrative	PL	plural
DIR	directional	POSS	possessive
DIST	distal	PRO	anaphoric pronoun
G1, G2, G3, G4, G5, G6	grammatical gender	PROX	proximal
IMPER	imperative	REL	relative
Q	interrogative	SG	singular
IPF	imperfective	TP	temporal
JUSS	Jussive		

Note

1. While I could easily find the Kabiye data myself as native speaker and verify some utterances by other Kabiye speakers, I did have to get data from informants in the case of the other languages. I therefore wish to express my gratitude to Fatima Djiba, Abou Sama and Laré Kantchoa, all of them students in African Linguistics at the University of Lomé, respectively for data of Nawdm, Tem and Ben, respectively. I thank also my friend Emile Koussanta Amouzou for additional data from Nawdm.

References

- Diesel, H. (1999). *Demonstratives. Form, function, and grammaticalization*. Amsterdam: John Benjamins [Typological Studies in Language 42].
- Fillmore, C. (1971). Types of lexical information. In D. D. Steinberg & L. A. Jakobovitz (Eds.), *Semantics* (pp. 370–392). London: Cambridge University Press.
- Heine, B. (1997). *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
- Lévikaza, K. K. (1985). *Phonologie, Tonologie und Morphosyntax des Kabiye*, Inaugural dissertation University of Cologne.
- — —. (1999). *Grammaire kabiye: Une analyse systématique. Phonologie, tonologie et morphosyntaxe*. Köln: Rüdiger Köppe.
- Levinson, S. C. (2000). *Presumptive Meaning*. Cambridge, MA: The MIT Press.
- Lyons, J. (1977). *Semantics*. Volume 2. Cambridge: Cambridge University Press.

Preprefix or not – that is the question

The case of Kwangali, Kwanyama and Ndonga*

Karsten Legère

Göteborgs Universitet

The linguistic varieties constituting the Ambo cluster (Oshiwambo) in Northern Namibia/Southern Angola share a number of similarities with the so-called Kavango languages (in these days frequently referred to as Rukavango).¹ This is not only due the fact that all these languages belong to the Bantu family, but there must have been strong historical links among those speaking these linguistic varieties which support an amazingly far-reaching genetic relationship. Thus, in particular neighbouring varieties such as Kwanyama (henceforth Kw.) and Ndonga (Nd.) on the one hand, and Kwangali (Kg.), on the other hand, display striking lexical resemblances which probably amount to 70 percent of cognates.² Notwithstanding this fact linguistic classifications do not reflect this situation, as e.g. Guthrie (1948) and subsequent publications put the Ambo cluster into Zone R, while Kg. and the other Kavango languages are grouped into Zone K. In addition, Guthrie (1948:64) makes even the categorical statement for Zone R. “This zone is sharply distinguished from its neighbours, but it is not easy to indicate the features which are peculiar to it.”

1. The structure of the noun class prefixes

The following discussion focuses on a typological issue which plays a prominent role in Bantu linguistics in general and in Guthrie’s classification in particular, i.e. the structure of the noun class prefix (NCLP). The question whether this NCLP is a single or a complex form is for Guthrie and other scholars a prominent feature for allocating a language to a zone.

For Zone K languages Guthrie (1948:53) gives the following description:

“1. The independent nominals have single prefixes in Groups 10–20, but double prefixes in Groups 40–50.”

One of the characteristic features of Zone R according to Guthrie (1948:64)³ is:

“4. Double independent prefixes occur throughout the zone. In Group 10–30 the first part of the prefix is usually o-, but there are ... exceptions.”

The two statements by Guthrie about the NCLP characteristics will be discussed below with reference to examples from Zone R (Kw. R21 and Nd. R22 in Guthrie 1948:64 as well as in Guthrie 1970a) and Kg. (which is not included in Guthrie 1948:52 in the K zone, but has more recently been listed in Guthrie 1970a: 14 as K33 of K.30 Luyana Group).⁴

The following table summarizes the noun class system of the three languages mentioned before to which the reconstructed Proto Bantu (PB*) noun classes are added.

Table 1. Noun classes⁵

Noun class	Ndonga	Kwangali	Kwanyama	PB*
1	ɔ - m̥-	mu-	ɔ - m̥-	mu-
2	a - a-	βa-	ɔ - va-	ba-
1a	Ø	Ø	Ø	Ø
2a	ɔ - ɔ-	βa-	ɔ - ɔ-	ba-
3	ɔ - m̥-	mu-	ɔ - m̥-	mu-
4	ɔ - mi-	(nɔ-mu-)	ɔ - mi-	mi-
5	ε -	ε -	ε -	q-
6	ɔ - ma-	ma-	ɔ - ma-	ma-
7	ɔ - fi-	si-	ɔ - fi-	ki-
8	i - i-	ji-	ɔ - i-	bq-
9	ɔ - N-	N-	ɔ - N-	n-
10	ɔ - ɔN-	nɔ-N	ε - εN-	n-
11	ɔ - lu-	ru-	ɔ - lu-	du-
12	ɔ - ka-	ka-	ɔ - ka-	ka-
13	—	tu-	—	tu-
14	u - u-	u-	ɔ - u-	bu-
15	ɔ - ku-	ku-	ɔ - ku-	ku-
16	pa-	pa-	pa-	pa-
17	ku-	ku-	ku-	ku-
18	mu-	mu-	mu-	mu-

The Kw. and Nd. columns of Table 1 display double prefixes, the Kg. column a single prefix. In Kw. and Nd. there is for class 1–15 (excluding 1a) an initial vowel (IV) which judging from Kw. seems to have been *o-* (at an earlier stage probably also for Nd. throughout). This affix is called preprefix (PRP) by various authors, which is also used in this paper, although in larger parts the term “initial vowel” will be preferred. This will be done for methodological reasons of not pre-empting its existence in Kg. According to Maho (1999) in recent years the term “augment” (to express the idea that the class identifying morpheme [Cl] is augmented) has come in vogue. However, this term does not seem to suit well, as it lends to confusion with the term “augmentative” in particular when abbreviated.

As can be seen in Table 1 some changes have already taken place in the shape of the preprefix in Nd. compared to Kw. which seems to be more “conservative” for retaining almost throughout *o-* as the initial vowel. Obviously, in Nd. regressive

vowel assimilation has resulted in creating identical vowel sequences for NCLP of class 2, 8 and 14.

The structure of nouns in Kw. and Nd. (here also relevant for adjectives) is

NCLP [PRP + CL] + STEM

The preprefix is a constituent of any noun in both linguistic varieties. There is no clearly marked grammatical role for it in Kw. or Nd.⁶ which probably only becomes evident when it is not present. This is also suggested by Hyman & Katamba (1993, quoted in Maho 1999:61–2). These authors argue that the preprefix-less (augment-less) construction is the marked one and therefore displaying a specific function in a given utterance. By discussing examples from Luganda particular reference is made to negation and focus. Both aspects are obviously relevant for Kw. and Nd. too, as illustrated below.

Here are some examples from Kw. and Nd. of sentences in the negative where the preprefix/initial vowel does not occur:

- | | | | | | | |
|-----|-----------|-----------|-------------|-----------|---|--|
| (1) | <i>ha</i> | <i>fi</i> | <i>téja</i> | | <i>hashíténýa</i> → <i>oshéténýa</i> (guy) | |
| | NEG | CL7 | guy | | ‘It is not the guy.’ | |
| (2) | <i>ha</i> | <i>m</i> | <i>páði</i> | | <i>hampádhi</i> → <i>ompádhi</i> (leg) ⁷ | |
| | NEG | CL9 | leg | | ‘It is not the leg.’ | |
| (3) | <i>ha</i> | Ø | <i>jśka</i> | <i>li</i> | <i>néne</i> | <i>hayóka linéne</i> → <i>eyoka</i> (snake) ⁸ |
| | NEG | CL5 | snake | CM5 | big | ‘It is not the big snake.’ |

The focussing role of the initial vowel becomes evident in examples (4) and (5). In the latter example the initial vowel clearly has a topicalizing function:

- | | | | | | | | |
|-----|------------|------------|-----------|-----------|-------------|-----------------|--|
| (4) | <i>tse</i> | ɔ | <i>tá</i> | <i>tu</i> | <i>léja</i> | <i>ɔmbílive</i> | <i>Tse otátu lésha ombílive</i> (Nd.) |
| | we | IV | PRES | we | read | letter | ‘We are reading the letter.’ |
| (5) | ɔ | <i>tsé</i> | <i>ta</i> | <i>tu</i> | <i>léja</i> | <i>ɔmbílive</i> | <i>Otsé tatu lésha ombílive</i> (Nd.) |
| | IV | we | PRES | we | read | letter | ‘It is us who are reading the letter.’ |

In Kw. and Nd. the preprefix/initial vowel is dropped in the vocative. The following examples from Nd. illustrate this:

- | | | | |
|-----|-------------------|-----------------|---|
| (6) | <i>pulákéna</i> | <i>mámáti</i> | <i>Pulákéna, mumáti!</i> (→ <i>omumáti</i>) |
| | | | ‘Listen, boy.’ |
| (7) | <i>pulákenéni</i> | <i>ántu</i> | <i>Pulákenéni, antu!</i> (<i>aántu</i>) |
| | | | ‘Listen, people.’ |
| (8) | <i>dána</i> | <i>thíthi</i> | <i>Dána, thíthi!</i> (<i>ethíthi</i>) |
| | | | ‘Dance, monster.’ |
| (9) | <i>danéni</i> | <i>mathíthi</i> | <i>Danéni, mathithi!</i> (<i>omathíthi</i>) |
| | | | ‘Dance, monsters.’ |

- (10) *íla fiténa* *Íla. shiténya* (\rightarrow *shiténya*)
 ‘Come, guy.’
- (11) *iléni iténa* *Iléni, iténya* (\rightarrow *iiténya*)
 ‘Come, guys.’

When personification (a frequent feature in fables and animal tales) takes place, the noun is shifted from its original class to class 1a (plural 2a). In this process the initial vowel is dropped, e.g.:

- (12) ɔ *fí mbúngu* \rightarrow \emptyset *fimbúngu Shimbungu* (cf. *oshimbungu*)
 IV CL7 hyena CL1A hyena Hyena the hyena
- (13) ϵ \emptyset *kóla* \rightarrow \emptyset *kóla Kóla* (cf. *ekóla*)
 IV CL5 crow CL1A crow Crow the crow
- (14) ɔ *m kwíju* \rightarrow \emptyset *mkwíju Mukwíyu* (cf. *omukwíyu*)
 IV CL3 fig tree CL1A fig tree Fig Tree the fig tree

Note the specific form of class 5 nouns when personified – the elision of ϵ clearly points to its status as preprefix, which has been retained in the citation form while the original class identifying morpheme CL is gone.

2. Nominal structure of Kwangali

For Kg. and other closely related linguistic varieties such as Manyo (Gciriku + Shambuyu) or Mbukushu the noun structure is

(NCLP+) NCLP + STEM

Consider the following examples: *sí-tji* (tree, cl. 7), *yi-díra* (birds, cl. 8), *mu-kúro* (river, cl. 3), *ru-súmo* (song, cl. 11), *ku-fu* (winter, cl. 15), etc. The plural NCLP is superimposed on NCLPs marking class 3, class 14, 15 and partly class 11 respectively, e.g. *no-mu-kúro* (rivers, NCLP class 10+NCLP class 3), *ma-ru-pátji* (ribs, NCLP class 6+NCLP class 12).

In Kg. (and Manyo) there is a NCLP resemblance to a preprefix as found in Nd. and Kw. This is in class 5 where in Kg. the vowel ϵ is commonly regarded as the NCLP. This position is also supported here as a point of departure. There are cognates like *e-kísi* (monster), cf. Kw. *e-kishi(kishi)*, or *e-zímo* (belly), cf. Kw. *e-dímo*, *e-ího* (eye), cf. Kw. *e-ísho*, *e-yóka* (snake), cf. Kw. *e-yóka*, etc. In class 5 no trace of the PB prefix *i-, exists in Kg. (nor Kw. or Nd.) while Mbukushu has retained *di-* which is referred to in Guthrie (1970b: 221) as CS 2204 a/b; cf. *di-kíthi* (monster), *di-nóta* (thirst).

3. Noun modifiers

The following section looks from a typological perspective at the position of nouns in combination with grammatical morphemes and function words. Both in R21/R22 and K33 adnominal constructions have the structure

$$NP_1 + CM \text{ (of } NP_1) a + NP_2$$

CM is the concord morpheme (also called “pronominal prefix”), *a* (PB CS 2267 a*) is the “linking element” (Guthrie 1970b:253). It will be noted in each case that the *nomen regens* appears without the iv. Examples (15) through (19) are from Nd.

- (15) $\text{ɔ má psa ya-a } \emptyset \text{ tate} \rightarrow \text{omápya gatate}$
 IV CL6 plot CM6-of CL1A father ‘father’s plot’
 NB: $\text{ya} + a \rightarrow \text{ya}$ (the vowels are coalesced)
- (16) $\text{ɔ ɾi ði yu-a } \varepsilon \emptyset \text{ ráka} \rightarrow \text{omúdhí gweráka}$
 IV CL3 root CM3 IV CL5 tongue ‘tongue root’
 NB: $\text{yu} + a \rightarrow \text{ywa}$ (which subsequently assimilates to the iv of *eraka*)
- (17) $\varepsilon \emptyset \text{ ðína d}_3\text{-a } \acute{\varepsilon} \emptyset \text{ mbɔ} \rightarrow \text{edhina lyémbo}$
 IV CL5 name CM5-of IV CL5 book ‘book title’
 NB: $\text{li} + a \rightarrow \text{lya} \rightarrow \text{d}_3a$ (which subsequently assimilates to iv of *émbɔ*)
- (18) $\text{ɔ } \eta \text{ gúwɔ j-a } \text{ɔ } \text{fí tépa } \acute{\text{ɔ}} \rightarrow \text{ónguwo}^9 \text{ yoshiténya } \acute{\text{o}}\text{sho}$
 IV CL9 cloth CM9-of IV CL7 guy DEM7 ‘this guy’s cloth’
 NB: $i + a \rightarrow ja$ (the *a* subsequently coalesces with the following *o*)
- (19) $\text{ɔ ka pále ka-a } \acute{\text{ɔ}}\text{akáti} \rightarrow \text{okapále kóshakáti}$
 IV CL12 field CM12-of Oshakati Oshakati Airport
 NB: $ka + a \rightarrow ka$ (the *a* and *o* of *Oshakati* then coalesce)

As noted in the above examples, the linking *-a* undergoes a number of morphophonological changes.

The following clauses illustrate *adnominal constructions* in Kwangali:

- (20) $\text{mu góŋgɔ gu-a } \emptyset \text{ hómpa} \rightarrow \text{mugóngo gwahómpa}$
 CL3 back CM3-of CL1A chief ‘the chief’s back’
 NB: $\text{gu} + a \rightarrow \text{gwa}$
- (21) $\beta a \text{ r}^{\text{ó}}\eta \text{gi } \beta a\text{-a } \emptyset \text{ simbúngu} \rightarrow \text{varongi vaSimbunu}$
 CL2 teacher CM2-of CL1A hyena ‘Mr. Hyena’s teachers’
 NB: $ka + a \rightarrow ka$

In those cases where N_2 is a noun other than of class 1a the following Kg. examples present a different picture for the linking part of the clause, e. g.:

- (22) $\text{si pápa si-a } \acute{\varepsilon} \emptyset \text{ hɔ} \rightarrow \text{sipápa sého}$
 CL7 skin CM7-of IV CL5 eye ‘eyelid’
 NB: $si + a \rightarrow sa$; $sa + \acute{\varepsilon} \rightarrow sé$

- (23) *u húngu u-a ε Ø zóka* → *uhúngu wezóka*
 CL14 poison CM14-of IV CL5 snake ‘snake poison’
 NB: *u + a* → *wa*; *wa + ε* → *wε*
- (24) *mu síngi gɔ sí haúɔ* → *musíngi gosíhaúto*
 CL1 driver CM1.of CL7 car ‘driver of the car’
 NB: *gu + a* → *go?*
- (25) *ka kámbe kɔ ka zéra* → *kakámbe kókazéra*
 CL12 horse CM12.of CL12 horse ‘little white horse’
 NB: *ka + a* → *ko?*
- (26) *ru súmɔ rɔ pa u pólítika* → *rusúmo rópaupólítika*
 CL11 song CL11.of CL16 CL14 politics ‘political songs’
 NB: *ru + a* → *ro?*
- (27) *Ø níme go ku kurúpa* → *nyíme gokukurúpa*
 CL1A lion CL1A.of CL15 be old ‘aging lion’
 NB: *gu + a* → *gwa* → *go?*

In the examples above the linking morpheme *-a* combines with the NP₁ concord morpheme CM in various ways, i.e.

- The vowel *a* is retained in the adnominal constructions in Kg. whenever N₂ is a member of noun (sub) class 1a as examples (20) and (21) demonstrate.
- The clauses (22) and (23) contain a class 5 noun in the adnominal position as N₂. Earlier reference was made to the noun initial vowel *ε-* in Kg. which was said to resemble the Kw. and Nd. preprefix. In these two examples (and with all other class 5 nouns) the linking morpheme *-a* cannot be identified, since obviously its elision (or assimilation with the noun initial *ε-*) has taken place.
- In (24) to (27) the existence of an underlying *a* for linking NP₂ with the antecedent NP₁ is assumed, but the surface form is a combination of the NP₁ concord morpheme with the vowel *o*. From the examples above (and the general experience about this issue) it may be concluded that when N₂ is any noun (excluding nouns of class 1a having a zero noun class prefix and class 5 nouns) the linking *-a* changes to *-o* in Kwangali (as well as in other linguistic varieties spoken in Kavango).¹⁰

The use of *-o* as the major component in linking NP₁ and NP₂ has been observed both by Dammann and Westphal in their Kwangali descriptions. The former speaks here of the emphatic pronominal stem (“emphatischer Pronominalstamm” – Dammann 1957:35), while Westphal (1958:16c–d) lists all recurrent forms with *-o* separately as entries of the “introductory prefix-complex ... before secondary ad-nominal stems”. None of them gives a plausible reason for the fact that the use of PB **-a* is limited to a single subclass of nouns.

The observations and comments on vowel assimilation (or elision) in Ambo and in Kw. can be further corroborated by the following examples where nominals combine with various function words.

The function word “*na*” (prep, conj - Guthrie 1970b:CS 2264 **nà*) combines with nouns, adjectives, etc. in Kw. and Nd.

- (28) *na* Ø *méme* → *naméme* (Kw., Nd.)
and CL1A mother ‘and/with mother’
- (29) *na* Ø *kavándʒe* → *nakavandje* (Kw.)
and CL1A jackal ‘and/with Mr. Jackal’
- (30) *na* ε Ø *kúja* → *nekúya* (Kw., Nd.)
and IV CL5 axe ‘and/with the axe’
- (31) *na* ɔ *shi lóngɔ* → *noshilónggo* (Kw., Nd.)
and IV CL7 country ‘and/with the country’
- (32) *na* ɔ *mú néne* → *nomúnéne* (Kw., Nd.)
and IV CL1/3 big ‘and/with the tall one’

While in (28) and (29) “*na*” is retained unchanged, in (30) through (32) the vowel of the function word *na* has obviously undergone regressive assimilation or elision.¹¹ In Kg. the function word “*na*” combines with nouns in the following two ways:

In class 1a no preprefix appears:

- (33) Ø *simbúngu na* Ø *mbánze* → *Simbúngu naMbánze*
CL1A hyena and CL1A jackal ‘Mr. Hyena and/with Mr. Jackal’
- (34) *mu róngi na* Ø *hómpa* → *muróngi naHómpa*
CL1 missionary and CL1A chief ‘missionary and/with Chief’

But in all other classes a trace of the preprefix can well be observed:

- (35) *m buru na* ε Ø *témba* → *mburu netémba*
CL9 Boer and IV CL5 wagon ‘the Boer and/with the wagon’
- (36) *ji kómbɔ na* ɔ *nɔ nzwi* → *yikómbɔ nonónzwi*
CL8 goat and IV CL10 sheep ‘goats and/with sheep’
- (37) *na* ɔ *tu kámbe* → *notukámbe*
and IV CL13 horse ‘and/with small horses’
- (38) *na* ɔ *ma gúni* → *nomagúni*
and IV CL6 wild oranges ‘and/with wild oranges’

Similarly, we find with “*nga*” (‘like’, Guthrie 1970b:243 – CS 2263 **nga*)) examples such as:

- (39) *nga* Ø *nuyóma* → *ngaNuyóma* (Kw., Nd.)
like CL1A Nuyoma ‘like Nuyoma’
- (40) *nga* ɔ *í ta* → *ngoíta* (Kw.)
like IV CL8 war ‘like war’

- (41) *nga* ɔ *si* *púndi* → *ngosipúndi* (Kg.)
 like IV CL7 chair 'like a chair'
- (42) *nga* ɔ Ø *súre* → *ngosúre* (Kg.)
 like IV CL9 school 'like school'
- (43) *nga* ɔ *nɔ* *mu* *síra* → *ngónomusíra* (Kg.)
 like IV CL9 CL3 tail 'like tails'
- (44) *nga* ε Ø *témɔ* → *ngétémo* (Kg.)
 like IV CL5 hoe 'like a hoe'

Compare for Kwangali also the locatives such as:

- (45) *mu* ɔ *ji* *róŋgɔ* → *moyiróngo* (cf. Kw.: *mɔilóŋgɔ*)
 LOC IV CL8 country 'in the countries'
- (46) *ku* ɔ *ku* *twi* → *kokútwi*¹² (cf. Kw.: *kɔkútsi*)
 LOC IV CL15 ear 'to the ear'

When it comes to adnominal constructions or antecedent *na* and *nga* the examples above display some striking typological similarities between Kw. or Nd. and Kg. In particular the occurrence of -o after the concord morpheme is a fact which must be properly accounted for and in a more convincing way than the shallow statements by Dammann or Westphal. This -o which is a recurrent feature in Kg. can be nothing else than a reflex of an otherwise obsolete preprefix in this language since it occurs in the same paradigms as in Kw. and Nd. While in Ambo varieties the structure of adnominal constructions can be well explained as the linking -a being assimilated (or elided), the traditional Kg. approach to this issue does not describe, but just postulates the existence of an -o as a link morpheme in classes other than 1a (and 5). However, it makes much more sense to assume that even in Kg. the underlying structure of adnominal clauses and for selected function words is -a as the antecedent of an initial vowel which is in fact the otherwise obsolete preprefix (mainly o-) where the former is assimilated by this preprefix in analogy with Kw./Nd.

This conclusion is further corroborated by Kg. class 5 nouns whose initial vowel has been treated as a NCLP so far, but which is obviously a reflex and remnant of the preprefix in Kg., again similar to Kw. and Nd. With this argumentation in mind and replacing IV with PRP accordingly, the examples above can be reanalyzed, thus e. g.

- (22) a. *si* *pápa* *si-a* *é* Ø *hɔ* → *sipápa sého*
 CL7 skin CM7-of PRP CL5 eye 'eyelid'
- (35) a. *m* *buru* *na* ε Ø *témba* → *mburu netémba*
 CL9 Boer and PRP CL5 wagon 'the Boer and/with the wagon'
- (24) a. *mu* *síngi* *gu* ɔ *sí* *haúɔ* → *musíngi gosíhaúto*
 CL1 driver CM1 PRP CL7 car 'driver of the car'

- (25) a. *ka kámbe ka ɔ ka zéra* → *kakámbe kókazéra*
 CL12 horse CM12 PRP CL12 horse 'little white horse'
- (37) a. *na ɔ tu kámbe* → *notukámbe*
 and PRP CL13 horse 'and/with small horses'
- (45) a. *mu ɔ ji róngɔ* → *moyiróngo* (cf. Kw.:) *mɔilóngɔ*
 LOC PRP CL8 country 'in the countries'

For Kw. and Nd. it was demonstrated earlier that PRP is not an irrevocable constituent of nouns in Ambo. This argument was supported by examples for vocatives, negation and personification. Hence, the canonical form of most Kg. nouns which is preprefix-less does not contradict the analysis above. A particular case is noun class 5 where the initial vowel *e-* which has been regarded as the NCLP so far is in fact the sole overt reflex and remnant of the preprefix system in Kg.

4. Conclusion

In concluding it may be taken for granted that at an earlier stage of its history Kg. had a system of double prefixes (i.e. preprefix + noun class identifying morpheme) which was much similar to that of current Kw. These preprefixes were gradually given up in the canonical forms except for class 5. They are still operational in some syntagmatic constructions like adnominal clauses and in combination with function words and the locative.

The erosion of the noun class system as illustrated with regard to Kg. is not strange to Bantu languages. Above, the case of some Nd. noun classes which had replaced the preprefix *o-* with harmonized vowels was mentioned. Even in Kw. which was described as having retained the preprefix *o-* to a large extent, this erosion of the class system is partially observed in current speech. Thus, the young generation does no longer pronounce *ɔvan^hu* (people), but has made its pronunciation more similar to Nd. by applying regressive vowel assimilation as in Nd., i.e. *aan^hu*.

Suffice it to note here that the erosion issue has been accounted for e.g. by Richardson (1967:386) who gave the following summary:

...the continued exposure of double prefixes to such erosive practices will result in the eventual elimination of initial vowels and the emergence of other means of indicating syntactic and emphatic connotations. It would seem that this stage of evolution has already been reached by single prefix languages.

It remains to reject Guthrie's statement about the sharp distinction of zone R languages from other zones. The typological similarities outlined in this paper are indeed an expression of the strong genetic relationship of Ambo and Kavango varieties, in particular Kw./Nd. and Kg. More morphosyntactic details could be provided, but would be out of place here. Nonetheless, time is ripe for reconsidering the position of Kg. and other Kavango varieties with regard to zone R in a genetic classification as sug-

gested by Bernd Heine (1972/3). In his classification a number of languages spoken in Southern Angola/Northern Namibia have been included in Luchazi-Chokwe (11.7, from Guthrie's K zone) and the West Highland Group (Westhochland-Gruppe, 11.8, a merger of H and R languages). The latter would definitely be adequate for the Kavango languages too.

Notes

* The author expresses his gratitude to Karel Nairenge (Education Officer for Rukwangali, National Institute for Educational Development, Okahandja, Republic of Namibia) and Freddie Kaulinge (Kwanyama Reader, Gamsberg Macmillan Publishers, Windhoek, Republic of Namibia) for their support and patience when discussing linguistic issues pertaining to Kg. and Kw./Nd. respectively.

1. For further discussion see Lusakalalu (2001). The distribution of the languages discussed in this paper is portrayed in the attached map which was compiled by Monika Feinen (Institut für Afrikanistik, University of Cologne).

2. Details in Legère (forthcoming).

3. It should be borne in mind that when Guthrie published his classification in 1948, almost no linguistic material for the Kavango languages was available. This may also explain the lack of Kavango language names except Mbukushu (spelled Mpukusu – K 33) therein.

4. The classificatory labels coined by Guthrie (1948) have been retained and still serve widely as a useful reference. Changes were made in Dalby & Mann (1987) who gave up the two digit numbers and instead wrote R2 for Ambo languages and K5 for Kavango languages (thus slightly revising Guthrie's material) without further dividing the groups. In recent years Maniacky (1997) presented his subclassification of K where in sub-group K.50 Kwangali is K.54 (Mbukushu K.52, Gciriku K.53).

5. PB is based on Meeussen (1967) who differs from Guthrie (1970 b) mainly in class 5 *-di* or *-dq* and 9/10 *-ny/ny*. In the table the pre-prefix/initial vowel is separated from the class-identifying morpheme (henceforth: CL) in Kw. and Nd. by a hyphen.

6. Identical with (IV).

7. For Nd. Fivaz (1986:33) describes the function of PRP as a definitiser thus assuming a functional value which, however, is neither adequate for the given language nor comprehensive enough to account for its rather complex role.

8. Example from Zimmermann & Hasheela (1998:28).

9. The structure of this class 9 noun with voiced velar plosive in syllable one of the stem is IV + NCLP₉ + stem (with compulsory homorganic nasal for voiced consonants in stem-initial position), e.g. ɔ + N + ηgúwɔ → ɔη + ηgúwɔ → ɔηgúwɔ (the nasal class marker assimilates to the stem-initial homorganic nasal).

10. Comp. in Mbukushu e.g. *dihám̐ba dyomého* (eye disease) which comes from *dí + hamba d̐3* ɔ ma + ih̐3, CL5 disease CM₅ IV CL6 eye d̐3 ← dj + a ← dí + a.

11. Even with *na* in compounds these changes can be traced, e.g. ɔ́névúmbɔ ← omún(a) + egumbo (cl. 5 – house) 'house owner'.

12. NCLP for class 10 could be probably another example, which displays the existence of this strange *o*.

References

- Dalby, D., & Mann, M. (1987). *A Thesaurus of African Languages*. London: Hans Zell Publ./K. Saur.
- Dammann, E. (1957). *Studien zum Kwangali. Grammatik, Texte, Glossar*. Hamburg: De Gruyter [Abhandlungen aus dem Gebiet der Auslandskunde 63, Reihe B: Band 35].
- Fivaz, D. (1986). *A Reference Grammar of Oshindonga*. Windhoek: Academy [African Studies of the Academy, 1].
- Guthrie, M. (1948). *The Classification of the Bantu Languages*. London: OUP.
- — —. (1967) 1, (1971) 2, (1970a) 3, (1970b) 4. *Comparative Bantu*. Farnborough: Gregg Intern. Publ.
- Heine, B. (1972/3). Zur genetischen Gliederung der Bantu-Sprachen. *AuÜ*, LVI (3), 164–185.
- Kloppers, J. K. (1994). *Bukenkango Rukwangali – English = English – Rukwangali dictionary* compiled by the late J. K. Kloppers; expanded by D. Nakare L. M. Isala. Windhoek: Gamsberg Macmillan.
- Lusakalalu, P. (2001). Languages and Glossonymic Units. *AAP*, 66, 47–65.
- Maho, J. (1999). *A Comparative Study of Bantu Noun Classes*. Göteborg: Acta Universitatis Gothoburgensis [Orientalia et Africana Gothoburgensia 13].
- Maniacky, J. (1997). Nouvelle classification de la zone k. Paris: http://www.bantu.ovh.org/fr/jmk_50.html (consulted 8 March 2001).
- Meeussen, A. E. (1967). Bantu grammatical reconstructions. *Africana Linguistica*, III, 79–121. Tervuren: Musée Royal de l'Afrique Centrale [Annales, Sciences Humaines 61].
- Richardson, I. (1967). Linguistic Evolution and Bantu Noun Class Systems. In *La Classification Nominale dans les Langues Négro-Africaines* (pp. 373–390). Paris: CNRS.
- Westphal, E. O. J. (1958). *Kwangari: An index of lexical types*. London: School of Oriental and African Studies.
- Zimmermann, W., & Hasheela, P. (1998). *Oshikwanyama Grammar*. Windhoek: Gamsberg Macmillan.

Nonverbal and verbal negations in Kabyle (Berber)

A typological perspective*

Amina Mettouchi

LLING Nantes (EA3827), INALCO Paris (EA3577) and IUF

Our aim in this paper is to show that there is a correlation between tense-aspect asymmetries with respect to negation, and the encoding of distinct non-verbal negations: semantically and morphologically, perfect/perfective negations are related to locative-existential nonverbal negative forms (of the type “there is not”), whereas imperfective negations are related to equative-attributive nonverbal negative forms (of the type “it is not”). Enlarging the scope of the study from Kabyle (Berber) to other African languages, we ultimately propose to consider emphasis on this opposition between attribution-equation and location-existence as a typological feature linked to the predominance of aspect over tense and mode in the organisation of the verbal system.

Introduction

Our aim in this paper is to show that there is a correlation between tense-aspect asymmetries with respect to negation, and the encoding of distinct non-verbal negations. In fact, the values taken by these non-verbal negations correspond to the semantic core values of aspectual forms: in Kabyle (Berber, Afroasiatic) for instance, perfect/perfective negatives are related to locative-existential nonverbal negation (of the type “there is not”), whereas imperfective negatives are related to equative-attributive nonverbal negation (of the type “it is not”). This correlation will be investigated briefly in some other African languages belonging to different phyla and presenting similar language facts (Hausa (Chadic, Afroasiatic), Yoruba (Kwa, Niger-Congo), San (Mande, Niger-Congo), Tashelhiyt and Tuareg (Berber, Afroasiatic)).

We will first present the problem of tense-aspect asymmetry with respect to negation, then move on to the analysis of non-verbal negations in the abovementioned languages. We will then show that verbal negation, in its interaction with tense-aspects in Kabyle, semantically mirrors the opposition between attributive and existential nonverbal negations.

We will ultimately propose to consider this opposition between attribution-evaluation and existence, which underlies the negative asymmetries under consideration, as a typological feature linked to emphasis on aspect rather than tense and mode in the organization of the verbal system.

1. Position of the problem

It is common among the languages of the world to find tense-aspect systems that are asymmetrical with respect to affirmative vs. negative. Such asymmetries are of various kinds: generally, tense-aspect distinctions are fewer in the negative than in the affirmative; some systems have tense-aspects restricted to the negative subsystem, others use different types of negative markers or auxiliaries to negate different tense-aspects. The variation in that respect is considerable.

A useful and insightful reference on this problem is Contini-Morava (1989). This book deals with affirmative-negative asymmetry in Swahili, in a pragmatic and semantic framework. It presents a detailed account of the various meanings of Swahili verb forms, ascribing the asymmetrical pattern to (1989: 171–174): difference in semantic categorization (dependency and location in time *versus* temporal limitation and probability of the affirmative) and difference in related pragmatic values (describe states of affair that actually occur *versus* forestall a possible expectation of the contrary by mentioning a state of affairs that fails to occur).

This mapping between semantics and pragmatics appears to be also related to the values, or number and kind of oppositions inside each subsystem. Contini-Morava (1989: 171–172) mentions such distinctions as:

for the affirmative domain: “main event/secondary event, potential/actual, simultaneous/sequential, iterative/unique, temporally connected/logically connected, etc. [...] state/activity, habitual/ongoing, completed/uncompleted, present relevance/lack of present relevance, etc.”

and for the negative domain: “restricted vs. unrestricted opportunity to occur, high vs. low contrast with expectation of the affirmative, context-free, “state”-like vs. context-bound, “event”-like negation, change of state likely vs. unlikely, etc.”.

Our study of tense-aspect asymmetry in Kabyle in a different framework (Mettouchi 1995, 1996a, 1996b, 1998, 2000, 2001) has led us to similar conclusions, and to the hypothesis that, beyond differences linked to the complex interplay of context, verb-type, subject-type, clause-type, discourse type, etc., tense-aspects have semantic core values with which negation interacts to produce various meanings in context.

An investigation of the functions and meanings of nonverbal negations (Mettouchi 1996a and 1996b) has led us to define basic values for nonverbal negations that are ultimately related to those we have discovered for imperfective and perfect/perfective¹ verbal negations.

It is this correlation that we are going to present in this paper, in order to support the view put forward by many linguists (among whom Givón 1984), that negation in natural language is not a logical marker that allows a neat symmetrical distinction between truth and falsity, but a complex cognitive operator, which we think is at the basis of (and interacts with) language categories based on systematic oppositions, such as, in our case, the category of aspect.

2. Nonverbal negation in Kabyle

Kabyle is a Berber language spoken in the North of Algeria, by approximately 3 million speakers, many more if we add the emigrated communities in France, Canada, etc. It belongs to the Afroasiatic phylum. The basic word order is VSO, which alternates with a SVO order in topicalized contexts. It is primarily a spoken language, but several novels have been published since the 1970s, and there is a very dynamic Kabyle press which started mostly in the 1980s.

For Kabyle data, we have worked on corpora which we have collected ourselves, composed of face-to-face conversation, political speeches, and a novel.

In actual use of language, more than one third of negations are nonverbal, and among those nonverbal negations, approximately half are attributive and half existential.²

	NEGATIVE	AFFIRMATIVE
	<i>mači d argaz i d'iri ny...</i>	<i>d argaz i d'iri</i>
Attribution	ATTRNEG COP man that COP'bad or... ³ 'It was not the husband who was bad or...'	COP man that COP'bad 'It was the husband who was bad'
	<i>ulaš⁴ msakit ašu ara č-nt</i>	<i>y-la wašu ara č-nt</i>
Existence	EXISTNEG poor-PL what that eat(aorist)-3PLF 'There was not, poor girls, any- thing for them to eat (poor girls, there was nothing to eat for them)'	3MS-be(perf) what that eat(aorist)-3PLF 'There was something for them to eat'

Nonverbal negations are used for constituent negation and focalization, as in the examples above, or for sentential negation, of the type just below.

mači nk ad xDm-γ, nitnti ad smuqul-nt
ATTRNEG me AIM work(IMPERF)-1S they-F AIM watch(IMPERF)-3PLF
'No way I'm going to work while they're going to watch...'

The characteristic feature of non-verbal negations is that they are invariable as far as tense-aspect or agreement are concerned. These negations often occur with relative clauses which contain finite verb forms, and which therefore provide referential anchoring differing from the speech situation (which is the default anchoring when non-verbal negations are not completed by a subordinate clause).

Semantically speaking, attributive negation indicates conflicting views on a given theme. On a pragmatic level, there is debate between the speaker and his co-speaker as to the degree of validity of the statement, and this debate involves modal standpoints. Attributive-equative negation is used for metalinguistic negative judgments.

On the other hand, the semantic characteristic of existential negation is to assess the lack of coincidence between the situation which is referred to and prior expectations. On a pragmatic level, this type of negation seldom involves debate between speakers and conflicting viewpoints. The speaker only provides information as to the fact that the expected situation does not hold, or that the expected person or thing is absent or not located in a given place.

This contrast between interactive, modal and sometimes polemical negation on the one hand, and descriptive, informational negation on the other, is the essential distinction that we will bear in mind when we broach the subject of aspectual negations.

Other Berber languages⁵ also distinguish between attributive and existential negations.

	VERBAL NEG	ATTRIBUTIVE NEG	EXISTENTIAL NEG
Tashelhiyt	<i>ur ... (yat)</i>	<i>ur d</i>	<i>lah</i>
Tuareg (Ahaggar)	<i>ur</i>	<i>ur ġiy</i>	<i>aba</i>

The three other African languages under study⁶ also distinguish between attributive and existential negations.

	VERBAL NEG	ATTRIBUTIVE NEG	EXISTENTIAL NEG
Ader Hausa (CHADIC)	<i>bà/bàa ... ba</i>	<i>bàa ... ba</i>	<i>baabù</i>
Yoruba (NIGER-CONGO-KWA)	<i>kò</i>	<i>kó/kí se</i>	<i>kò sí</i>
San (NIGER-CONGO-MANDE)	<i>bā/bēè ... wā</i>	<i>... bēè wā</i>	<i>...bām ba</i>

If the semantics of those non-verbal negations have been described by linguists, they have seldom been analyzed in depth. Horn (1989:448–452) mentions a few examples of such nonverbal negations, and relates them to former philological or philosophical analyses of negation, such as the Hegelian dichotomy between significant and insignificant negation (1989:451). Horn (1989:448) also makes the following observation “The negator used in nonexistence statements and other verbal environments is often formally distinct from the one used in negative identity statements and/or for constituent (especially nominal) negation.” This observation is rephrased at various points of his development: “in many languages, as touched on above, a special negative existential form can be isolated from both the general predication negation and the special emphatic or constituent negator, if any” (1989:451), thus forming a “recurring

morphosyntactic split between one negation employed for straightforward negative predications (predicate denials) and for nonexistence claims and another employed for negating identity statements or non-verbal constituents" (1989:451).

Those remarks, which were not further developed in Horn (1989) seem to reinforce our claim that beyond the variety of negative forms, an underlying system that transcends morphosyntactic distinctions can be brought to light.

Various approaches can be chosen to reach this aim. Ours is based on the following hypothesis: existential and equative-attributive negations are marked encodings of the two semantic poles that underlie the notion/operation of negation in general, namely subjective assessment, and referential stocktaking. In that respect, we follow Culioli's hypotheses as stated for instance in (1988:112): "Il existe une opération primitive de négation liée d'un côté à la valuation subjective (bon/mauvais, d'où rejet, refus) et de l'autre à la localisation spatio-temporelle (présence/absence; vide; apparition/disparition; itération)".⁷

In some languages this distinction is not marked morphosyntactically while in others, such as Kabyle, it is.

Further remarks can be made about this binary system of nonverbal negations, which allow to consider the possibility of bridges between the verbal and the nonverbal domains:

Attributive negation is closer than existential negation to verbal negation in general.

The affirmative counterpart of existential negation (in Berber in general) is a verbal predicate in the perfect/-ive.

And in fact, the study of verbal negation in Kabyle, which we will now present, shows striking similarities with this binary organization of the nonverbal negative system.

3. Verbal negations

In Kabyle, two thirds of negations are verbal, and use the preverbal negator *ur*.

	AFFIRMATIVE SUBSYSTEM	NEGATIVE SUBSYSTEM
AORIST	<i>a(d)</i> + AORIST ≈30% (<i>ad y-Du- ad y-krz</i>) AORIST (IMPERATIVE) ≈1%	<i>a wr</i> ⁸ + AORIST <0.5% (negative optative) ----- <i>ur</i> + IMPERFECTIVE ≈35% (<i>ur i-ṭDu – ur i-kRz</i>) (including ≈5% IMPERATIVES)
IMPERFECTIVE	----- IMPERFECTIVE ≈15% (<i>i-ṭDu – i-kRz</i>) <i>a(d)</i> + IMPERFECTIVE <1% (<i>ad i-ṭDu - ad i-kRz</i>) <i>la</i> + IMPERFECTIVE <1% (<i>la i-ṭDu – la i-kRz</i>)	
PERFECT/-IVE	PERFECT/-IVE ≈53% (<i>y-Da – y-krz</i>)	<i>ur</i> + NEGATIVE PERFECT/-IVE ≈65% (<i>ur y-Di – ur y-kriz</i>)

NB: Percentages are to be read (and added) vertically. Figures are approximations of counts made on our conversational corpus (30 minutes).

Examples are based on two verb stems, *D* (come) and *KRZ* (plough). A basic utterance is composed of a personal affix and a radical. The radical combines a root and an aspectual scheme. Verbs cannot appear without their personal affixes, nor in a non-finite form.

y-Da: 3SM-COME(PERF.): ‘he came/has come’
Basic utterance = personal affix (*y/i*: 3rd SG.MASC) + RADICAL (*Da*)
RADICAL: *Da* = ROOT (*D*) + aspectual scheme (*a*⁹).

There are four aspectual themes (Aorist, Perfect/-ive, Negative Perfect/-ive, Imperfective). The first three themes are opposed on the basis of apophonia, and the fourth is prefixed or has tensed (or geminated) consonants. Some of those aspectual themes (aorist and imperfective) can be preceded by preverbs: *ad* marks the potential quality of the predication. It is a modal preverb. Its range of meanings covers future, generic, potential, plausible, habit in the past, complement clauses. *La* indicates simultaneity (only with imperfective).

The perfect/-ive is used in narratives, and in the assessment of situations or actions (Mettouchi 2000). In independent and main clauses, it refers to past or present with states and stative predicates, and to past with dynamic predicates. In subordinate clauses, the temporal distinction disappears. It is mostly a non-dynamic aspect whose interpretation also depends on diathesis: in the perfect/-ive, a basic utterance can be interpreted differently according to the status of the subject (agent or experiencer).¹⁰ This has led us to consider that the perfect/-ive in Kabyle is very permeable to the way the referential event or situation is construed. The perfect/-ive indicates that the situation or event or state is construed as “being the case”.

The bare imperfective has habitual and progressive meanings (Mettouchi 1998, 2001). The progressive is the marked value, currently renewed by the introduction of preverb *la*. The habitual is the unmarked value, and some of its uses compete with those of *ad*+aorist, especially in past contexts. For both values, emphasis is put on the subject as agent or topic. In fact, the imperfective is incompatible with stative predicates, and never allows the stative reading observed with the perfect/-ive, in cases where the verb is compatible with both diatheses.

In the negative subsystem, it takes on in many contexts a prospective meaning, which explains why in most grammatical descriptions the form *ur*+imperfective is considered as the counterpart of the *ad*+aorist form (which very often refers to future time) in the negative subsystem. This has led us to consider that the imperfect in Kabyle is dynamic (agent-oriented), and thematic in its information structure. The negative perfect/-ive (Mettouchi 2000, 2001) is synchronically the counterpart of the perfect/-ive in negative contexts, but it also has residual counterfactual meanings in positive subordinate hypothetical clauses.¹¹

We can notice that in the perfect/-ive, the system is more or less symmetrical, with a perfect and a so-called negative perfect, which are not always morphologically distinct. On the contrary, in the imperfective, there is a high degree of asymmetry:

Preverbed forms do not appear in the negative (except for the special case of *a wr*+aorist)

In the negative, the aorist represents less than 0.5% of occurrences, some of its pragmatic values being conveyed by the imperfect.

A hypothesis to account for this “asymmetrical” asymmetry would be that the negative morpheme *ur* behaves like a preverb and directly modifies the aspectual scheme. This view (Mettouchi 1995, 1996b and 2001) is supported by the fact that *ur* is a former verb (Prasse 1972:244) which has lost its inflections and has therefore become a particle, and also by the fact that it triggers the same word-order changes as preverbs.¹²

Our view is that the range of pragmatic values in negative contexts is conveyed thanks to this interaction between preverbal negation (*ur*) and aspect, and in relation to speech situations and contextual influence. We will argue that the range of pragmatic values taken on by verbal negation in relation to aspect¹³ originates in a narrower and more abstract system of oppositions, which is based, in Kabyle, on a non-durational and non-temporal definition of aspect. This different definition, based on Culioli’s utterer-centred theory,¹⁴ has been formulated in Danon-Boileau (1991: 19) “l’aspect sanctionne, non une quantité d’action, mais la relation entre le but du procès et l’état de la réalité au moment de l’énonciation”.¹⁵ Such a definition of aspect is compatible with the “uncompleted vs. completed” terminology, if those terms are not taken as describing the unfolding of a process along a time-line, but the assessment of the coincidence, or non-coincidence between the representation of the completed process or stabilized state and the situation of reference.

Along those lines, and thanks to a corpus-based study of more than 700 affirmative and negative sentences, we have been able to characterize semantically the various aspects of Kabyle as follows:

The perfect/-ive marks the coincidence between the representation of the completed process or stabilized state and the situation of reference, the emphasis being laid on the latter. Hence the “objective” undertones of that aspectual form.

The negative perfect/-ive marks the disjunction between the representation of the completed process or stabilized state and the situation of reference, the emphasis being also laid on the latter, just like for perfect/-ive. This emphasis on the situation of reference for both perfect/-ives explains the fact that the distinction between negative perfect/-ive and perfect/-ive is losing ground.

The imperfective marks the non-coincidence between the representation of the completed process and the situation of reference, the emphasis being laid on the former. Hence the “subjective” undertones of that aspectual form: the process or the predication are construed as intentional, or debatable. In terms of agency, the features of agentivity and intentionality are enhanced; in terms of information structure, the viewpoint of the speaker on the predication prevails, thus triggering possible polemical and modal meanings.

The aorist does not relate the representation of the completed process or stabilized state to the situation of reference, it functions on one level only, that of representation. Hence its compatibility with states as well as processes, and the fact that it is always preceded by an irrealis modal preverb in Kabyle. It cannot be negated in Kabyle, except for very rare cases (optative-prohibitive).

Having described the system of aspectual oppositions in those terms, we can relate the three aspectual forms that appear in the negative to the two non-verbal negations whose semantics we described in part 2:

(negative) perfect/-ive and existential negations share an emphasis on the situation of reference, by which prior expectations are gauged, imperfective and equative-attributive negations share a feature of interpretation, debatable judgment on a theme/topic.¹⁶

The existence of a common territory between the imperfective aspect and attributive negation, and between the perfect/-ive aspect and existential negation is brought to light in the nonverbal glosses of the following examples:

- (1) *Yiw* *ur* *ɤ* *iD* *i-Wit* *ara*, *tis* *sna*
 one NEG ACC 3SF PROX 3SM-hit(NEG.PERF) NEG2 time two
i-dɣr *asn* *aɛkwaz* *i-ruħ*.
 3SM-throw(PERF) DAT3PLM stick 3SM-go(PERF).

First he didn't hit her, and secondly he threw his stick towards them and left.

- (1') *ulaš tiyriṭ.*
 EXISTNEG beating.
 There is/was/has (had) been no beating
- (2) *ur byi-n ara ad eiwn-n*
 NEG want(NEG.PERF)-3MPL NEG2 IRR help(AOR)-3MPL
 They didn't¹⁷ want to help
- (2') *ulaš lbyi*
 EXISTNEG willingness.
 There is/was/has (had) been no will (to help)
- (3) *tlata ny ma rbea tikal aKNi, y-uyal uqbl a t*
 three or if four times thus, 3MS-become(PERF) before IRR ACC3MS
 'Three or four times, it happened that before we let him
- n-Ġ ad i-ruḥ a t iD*
 1PL-left(AOR) IRR 3MS-go(AOR) IRR ACC3MS PROX
n-Sali s imεLmn, wid
 1PL-make-climb(AOR) to chiefs, those
 go, we took him to the chiefs, those chiefs
- Ni ur Kat-n ara zwr-n šwiṭ,*
 ANAPH NEG hit(IMPERF)-3MPL NEG2 be.clever(PERF)-3MPL a.little,
 they don't beat (the prisoners), they are a bit clever,
- Sn-n amk ṭmslay-n.*
 know(PERF)-3MPL how talk(PERF)-3MPL.
 they know how to talk.'
- (3') *wid Ni, maČi d ijḥliyen.*
 those ANAPH ATTNEG COP brutes.
 'Those ones, they are not brutes (even if the opposite view might be sustained).'
- (4) *ur iyi i-ṭyiḍ ara lḥal d lxir i*
 NEG DAT-1S 3MS-upset(NEG.PERF) NEG2 situation, COP good REL
yi bya-n.
 DAT-1S want(PERF)-3MPL
 'The situation wasn't upsetting me, (I knew that) they wished me well.'
- (4') *maČi d aywbl lḥal agi*
 ATTNEG COP trouble situation this
 'it was not a problem, this situation' (even if it should have been, in someone else's opinion).

The glosses that were elicited in relation to verbal negations worked by pairs: attribution and imperfective, existence and perfect/-ive. Of course some of those glosses may seem far-fetched, but our purpose is not to show that verbal and nonverbal negations are synonymous. Rather, our aim is to underline the existence of common semantic features between negated aspectual forms and nonverbal negations.

This language-internal exploration of the case of Kabyle has led us to wonder whether such correspondences between attribution and imperfective on the one hand, and existence and perfect/-ive on the other hand, also appeared in other languages.

The following table,¹⁸ which synthesizes the data in Galand (1994), represents TAM asymmetries in two Berber languages (with a verb meaning “do again”, with first person plural agreement).

We can see that in Tuareg, there are two specific negative themes instead of one for Kabyle and Tashelhiyt. Except for the aorist, there is no identity between the aspectual themes used in the positive and those used in the negative. There is a very clearcut distinction between the negative and the affirmative domains in relation to aspect.

In Tashelhiyt, the distinction only concerns aspectual themes in the perfect/-ive (perfect/-ive vs negative perfect/-ive). But otherwise there is symmetry at that level. It’s only through the order of preverbs that we can draw a line, not between affirmative and negative, but between the modal domain (the TAM particle precedes the negative marker) and the indicative domain (the negative marker precedes the TAM particle).

It is interesting to note that Tashelhiyt is said to be drifting from aspect to tense as a language (Leguil 1982). This shift from aspect (relationship between the representation of the completed process or stabilized situation and the situation of reference) to tense (position of the event-time or situation-time with respect to a deictic or translated origin) apparently has consequences on the symmetry of verbal forms as regards negation vs affirmation. Nonverbal negations are described as part of the grammar of Tashelhiyt, but there is competition with forms bearing TAM distinctions.

	AFFIRMATIVE	NEGATIVE
Tashelhiyt	<i>nuls</i> (perfect)	<i>ur nulis</i> (negative perfect)
	(<i>ar</i>) <i>nTals</i> (imperfect)	<i>ur a nTals</i> (imperfect)
	<i>rad nals</i> (“future”) (aorist)	<i>ur rad nals</i> (aorist)
	<i>ad nals</i> (“optative”) (aorist)	<i>ad ur nals</i> (aorist)
	<i>als</i> (imperative 2p) (aorist)	<i>ad ur talst</i> (aorist)
Tuareg (Ahaggar)	<i>nulðs</i> (perfect)	<i>ur nulis</i> (negative perfect)
	<i>nðta:lðs</i> (affirmative imperfect)	<i>ur nðtilðs</i> (negative imperfect)
	<i>e nalðs</i> (“future”) (aorist)	<i>ur e nals</i> (aorist)
	<i>als</i> (imperative 2p) (aorist)	<i>ur tulisðd</i> (negative perfect)

In the three other African languages¹⁹ under study, verbal systems are also asymmetrical, and some links can be made between nonverbal and verbal negations, and among verbal negations.

	AFFIRMATIVE	NEGATIVE
YORUBA	Ø (perfective)	kò + Ø
	tí (marked perfective)	kò tî
	ń (progressive & habitual)	<i>no negative form</i>
	máa ń (habitual)	kîi or kò kîi
	máa (ingressive-continuative)	<i>no negative form</i>
	òò (future)	kò níí í
	(optative-imperative)	máà
HAUSA (Ader)	taa/ta (perfective)	bà tà ...ba
	yì (aorist)	dàC yì
	zân (future)	bàa zân ...ba
	nàa/kà (imperfective)	baa (imperfective + prospective)
SAN	Ø (injunctive)	bārà + nonverbal pred. + wā
	tá (perfective)	bā ... wā
	n (habitual)	bēè ... wā
	n gwê (progressive)	bēè gwê ...wā

In all those languages, the (unmarked) perfect is treated differently from the other TAM: the only distinction between affirmative and negative is the presence of a negative marker. On the contrary, imperfective and modal forms show much variation between affirmative and negative: the use of different sets of TAM and negative markers draws a clearcut distinction between negative and affirmative. There apparently is less need to use different forms in the perfect/-ive than in the imperfective. This relative symmetry is probably related to the “objective” dimension of the perfect/-ive, which lays emphasis on the situation of reference, and presents the predication as either “being the case”, or “not being the case”. In our opinion, this special treatment of the perfect/-ive in the affirmative/negative opposition should be linked to the existence of a distinct marker for existential negation, which sets it apart from attributive negations.

This brief overview of TAM asymmetries in African languages needs to be carried further, but it points towards a phenomena that seems to have typological implications.

As for the relationship between nonverbal and verbal negation, we notice that in San, where the verb stem is inflected for aspect there is a distinct negative marker in the imperfective (bēè...wā) and in the perfect/-ive (bā...wā), and it is the imperfective negator that we find in attributive negations. This would tend to reinforce our claim about the correlation between attributive negation and negation of the imperfective.

Conclusion

Our aim in this paper was to show that there is a correlation, across languages, between tense-aspect asymmetries with respect to negation, and the encoding of distinct

non-verbal negations. Namely, if a given language presents different TAM markers in the negative as opposed to the affirmative subsystem, and also presents different non-verbal negations, then the core values of aspectual forms are likely to mirror the core values of the non-verbal negations.

We have shown that in detail for Kabyle, and have provided data for other African languages. Further investigations are necessary to specify the details of this binary relationship in languages other than Kabyle, but this opposition between interactive, intentional and thematic on the one hand, and descriptive and informational on the other hand seems to hold on a typological level.

Negation and aspect are therefore intimately related on a semantic plane, and this appears all the more clearly, through asymmetries as regards negation vs affirmation, as the verbal system of the language under consideration is predominantly aspectual.

The hypothesis that we would now like to test is the following: strictly aspectual systems are more likely to make extensive use of nonverbal negations whereas tense-based systems are more likely to replace them in the long run by forms with TAM distinctions.

Notes

* I would like to thank the conveners of the conference, as well as the colleagues who have kindly answered my queries or provided information related to my topic: Ursula Drolc, Karen Ebert, Tom Givón, Tom Güldemann, Christa König, Derek Nurse and Ekkehard Wolff.

1. Perfect and perfective are not morphologically or morphosyntactically distinct in Berber, hence our use of the term “perfect/perfective”. (I thank Derek Nurse and Christa König for their useful remarks on aspectual terminology).

2. A count made on one of our conversational corpora gave the following results: number of verbal negations = 60, number of existential negations (*ulaš*) = 15, number of attributive-equative negations (*mači*) = 17.

3. In the Berber examples, capitals represent geminated or tensed consonants. Table of abbreviations: *SIMULT.*: preverb marking simultaneity with another process, or progressive aspect (*a/la*); *IRR.*: preverb marking prospective (and habit in the past) values (*ad*); *PROX*: proximal particle; *PERF*: perfect/-ive aspect; *EXISTNEG*: existential negation; *NEG.PERF*: negative perfect/-ive; *ATTRNEG*: attributive negation; *IMPERF*: imperfective aspect; *COP*: copula; *AOR*: aorist; *DAT*: dative; *ACC*: accusative; *NEG*: preverbal marker of negation (*ur*); *NEG2*: negative reinforcement (*ara*).

4. Diachronically speaking, *mači* is a loanword of Arabic origin, whereas *ulaš* is considered to be the contraction of “*ur y-Li ša*”: “NEG be(negperf)-3sm thing”.

5. Examples are taken from Galand (1994)

6. Examples are taken from Caron (1990) for Hausa, Sachnine (1990) for Yoruba and Platiel (1990) for San.

7. “There is a cognitive operation of negation linked on the one hand to subjective valuation (good vs. bad, etc.), and on the other hand to spatio-temporal location (presence vs. absence, etc.).”

8. This negative marker is modal, it is composed of a nominal element plus the verbal negator *ur*, and is used with the aorist aspect. Its frequency is very low, and it conveys optative-prohibitive meaning.

A wr D y-uyal!
 Neg-Opt Deict. come-back(aorist)-3sm
 May he not come back!

9. Except for the imperfective, aspectual distinctions are based on apophonia in Berber, so that it is not possible to consider that *a* is the marker of perfect/-ive for all verbs. Sometimes it is a *schwa*, sometimes *u/*.

10. *y-krz yigr* *y-krz wrgz (igr)*
 plough(perf)-3sm field plough(perf)-3sm man (field)
 the field is ploughed the man ploughed/ has ploughed (the field)

A resultative reading can occur when the verb form takes on a passive diathesis (prefix *ṭwa*) in the perfect/-ive:

y- ṭwa-krz yigr
 plough(perf)-PASS-3SM field
 the field has been/was ploughed, and as a result it is ploughed

11. With a special counterfactual hypothetical marker: *IMr y-Li wašu zri-y* (if(counterfactual) be(neg-perf)-3SM what know(perf)-1S = if I had known something).

12. The following word orders are respected in relation to TAM, negation, or subordination markers:

basic predicative utterance	dative pronoun	accusative pronoun	deictic particle.
<i>y-fka (give(perf)-3S)</i>	<i>yas (3S)</i>	<i>t (3SM)</i>	<i>iD (prox)</i>
(He gave it to her/him)			
TAM/NEG/SUB	DAT. pron.	ACC. pron.	deictic part. basic pred. utterance
<i>ur (NEG)</i>	<i>s (3S)</i>	<i>t (3SM)</i>	<i>iD (prox)</i> <i>y-fki</i> <i>(give(neg-perf)-3S)</i>
(He didn't give it to her/him)			

13. For a thorough description of these values, see Mettouchi (1998 & 2000).

14. Itself based on Benveniste's work on the indexation of verb forms and deictics to the speech situation.

15. "Aspect does not measure a quantity of action, but it relates the representation of the expected completion of the process (or the expected state) to the representation of actuality/reference". Linguists who are not acquainted to the utterer-centred signed-based semantics developed by Culioli and his followers should be informed of the fact that this reading is not pragmatic, but semantic: it is not the speaker as a person in a definite situation who makes this assessment, but the utterer. The latter is a theoretical construct: the origin of TAM and determination construals.

16. The other dimension of the imperfective aspect, namely its agentive-intentional feature, does not appear as such in attributive negations, because these negations are non-dynamic, being nonverbal. Nevertheless, the shift in the imperfective, between intentionality of the agent and intentionality of the speaker is a widely acknowledged fact (cf. the ambiguity of *Mrs Smith isn't seeing anyone: I'm reporting her refusal to see you* vs. *I'm objecting to your seeing her*), and we therefore consider that we can relate agentivity and speaker's viewpoint at the level of semantics.

17. The examples containing verbal negations are taken from our corpus, that is why the translation provides unambiguous temporal values.
18. Examples are taken from Galand (1994).
19. Examples are taken from Caron (1990) for Hausa, Sachnine (1990) for Yoruba and Platiel (1990) for San.
20. Christa König (whom I thank for her personal communication) remarked that existential negation was related to perfective aspect in Maa, and Ekkehard Wolff (whom I also thank for his personal communication) underlined the same phenomenon for Lamang. Further investigations, which we have not yet been able to make, should show more precisely the extent of the relationship between aspect and nonverbal negations in those languages.

References

- Caron, B. (1990). La négation en haoussa. *Linguistique Africaine*, 4, 31–46.
- Chaker, S., & D. Caubet (Eds.). (1996). *La négation en berbère et en arabe maghrébin*. Paris: L'Harmattan.
- Contini-Morava, E. (1989). *Discourse Pragmatics and Semantic Categorization. The case of negation and tense-aspect with special reference to Swahili*. Berlin: Mouton de Gruyter.
- Culioli, A. (1988). La Négation: Marqueurs et opérations. In *Pour une Linguistique de l'énonciation, tome 1* (pp. 91–114). Gap-Paris: Ophrys.
- Danon-Boileau, L. (1994). La négation: De l'absence au refus, du refus à l'absence. In P. Attal (Ed.), *La négation* (pp. 177–189). Paris: Numéro spécial de Lynx.
- Galand, L. (1994). La négation en berbère. *MAS-GELLAS n.s.*, 6, 169–181.
- Givon, T. (1984). *Syntax. A functional typological introduction*. Amsterdam: John Benjamins.
- Horn, L. (1989). *A Natural History of Negation*. Chicago, IL: University of Chicago Press.
- Leguil, A. (1982). La naissance des temps en Chleuh. *Bulletin des Etudes Africaines*, 3, 57–84.
- Mettouchi, A. (1995). *Aspect et Négation : recherche d'invariants et étude énonciative de l'incidence de la modalité négative sur l'aspect en berbère (kabyle)*. PhD dissertation. University of Paris III, 2 vols, 500 p.
- Mettouchi, A. (1996a). Remarques sur la négation en berbère (kabyle). *Mémoires de la Société de Linguistique de Paris, nouvelle série*, 4, 105–116.
- . (1996b). La négation dans les langues du Maghreb. Synthèse. In S. Chaker & D. Caubet (Eds.), *La négation en berbère et en arabe maghrébin* (pp. 177–195). Paris: L'Harmattan.
- . (1998). Aspect et négation: Remarques sur l'inaccompli et la négation en anglais et en berbère (kabyle). *Cahiers Chronos*, 2, 191–205.
- . (2000). Accompli et négation en kabyle. In *Etudes Berbères et Chamito-sémitiques, Mélanges offerts à K. G. Prasse* (pp. 281–293). Louvain: Peeters.
- . (2001). Negation and aspect in Kabyle (Berber). The case of the imperfective. Paper presented at the 29th North American Conference on Afroasiatic Languages, Toronto-Canada, March 30–April 1, 2001.
- Platiel, S. (1990). La négation en san. *Linguistique Africaine*, 4, 153–165.
- Prasse, K-G. (1972). *Manuel de grammaire Touarègue (tahaggart)*. Copenhague : Academisk Forlag.
- Sachnine, M. (1990). Aperçu sur la négation en yoruba. *Linguistique Africaine*, 4, 205–221.

Grammaticalization chains of the verb *Kàre* ‘to give’ in Kabba

Rosmarie Moser

La Trobe University, Melbourne

In linguistic terminology grammaticalisation refers to an ongoing process involving the use of lexical items (e.g. verbs or nouns) for grammatical purposes.

It is a phenomenon which has been observed in many languages around the world, including Kabba, a Nilo-Saharan language from the Central African Republic.

The term polygrammaticalisation has been used in this paper to highlight the multi-faceted nature of grammaticalisation in Kabba. It demonstrates how the verb *kàre* ‘to give’ can be used to express benefactive, recipient, causative, purposive, resultative, complement, sequential and connective relationships in a variety of different contexts and by speakers of all ages.

The structure, conjugation, syntax and semantics of the verb *kàre* ‘to give’ is discussed and illustrated, before the grammaticalisation processes are analysed and exemplified. Lexical and grammaticalised structures may occur together in the same structure.

1. Introduction¹

Kabba belongs to the Sara group of the Central-Sudanic branch of the Nilo-Saharan languages. It is mutually intelligible with Laka, Mbay, Ngambay and other languages of the same group in the Central African Republic, Chad and Cameroon, constituting a dialect chain. The Kabba are believed to originate from the Upper-Nile region. They are a proud and dignified people who love their language and ethnic identity and always speak it among themselves; but when people from other ethnicities are present, Sango or French are the preferred languages of communication. However, many younger Kabba married someone from another ethnic group and as a result their children no longer speak Kabba. The Kabba have large families, but very few men are polygamous. Most of them are adherents of the ‘Eglise Evangélique des Frères’. In the Kabba homeland region of the Paoua district church services are conducted entirely in Kabba.

1.1 Typological features of Kabba

Like most African languages, Kabba is rich in proverbs, ideophones, idiomatic expressions, and folktales. It has a three level tonal system and distinguishes between alienable and inalienable nouns, and inclusive and exclusive person markers. Typological features of Kabba include syllabic nasal consonants, vowel harmony, a complex pronominal system, an honorific marker, logophoric pronouns, case markers for genitive, dative, locative and comitative. Kabba also has a causative marker and spatial adverbs related to body parts. It has basically a dependent marking system and its unmarked constituent word order is subject-verb-object. It also has traces of a noun and a verb class. Serial verb constructions are frequent. Grammatical expansions of lexical items are flourishing in this language. The verb 'to give', which will be examined in detail, exemplifies some synchronic processes of grammaticalization, which include markers for recipient, benefactive, causative, purposive, resultative, complementizer, and discourse functions.

The speakers interviewed for this research involved both men and women ranging from the early twenties to the early seventies. All originate from the Paoua region in the north of the Central African Republic. This paper is based on the original fieldwork conducted in the Central African Republic between 1995–2000. The orthography is that which has been officially approved in 1999, except for the schwa, which will be maintained throughout this paper for the verb /kàrə/ and its derivations and for a few minimal pairs. The tonal markings for high tone and low tone are also maintained; the unmarked syllable is always a mid-tone. The contour tones are written as two vowels, although their actual length is only slightly longer than a simple vowel; sometimes the second tone occurs on the following sonorant consonant.

1.2 Grammaticalization chains

The purpose of this paper is to attempt to clarify the multitude of confusing functions and meanings of the verb *kàrə* 'to give' in Kabba. As the various forms are being used by all speakers, it may be assumed that the process of grammaticalization has started a long time ago. As early as 1912 Meillet used this term to refer to "the grammatical character of a previously autonomous word" (Heine, Claudi, & Hünemeyer (1991a: 1). Much later Heine and Reh (1984: 15) saw grammaticalization as "an evolution whereby linguistic units lose in semantic complexity, pragmatic significance, syntactic freedom, and phonetic substance, respectively." Hopper and Traugott (1993: 2), who looked at it from a synchronic and a diachronic perspective, point out that grammaticalization is a "linguistic change through which a lexical item in certain uses becomes a grammatical item, or through which a grammatical item becomes more grammatical". More recently Heine, Claudi and Hünemeyer (1991a: 4) conceived of grammaticalization essentially as a diachronic process, which occurs "constantly and independently in all languages". They defined grammaticalization as a "dynamic, unidirectional, historical process whereby lexical items in the course of time acquire a new status as grammati-

cal, morphosyntactic forms, and in the process come to code relations that either were not coded before or were coded differently".²

In Kabba, the source lexeme *kàrə* 'to give' has given rise to polysemy and a series of grammatical morphemes. With some exceptions, they retain the phonological form of their source lexeme, but function like free grammatical morphemes with inflections for subject and indirect object pronouns. These grammaticalized forms behave syntactically like the verb 'to give', but they have different semantic functions and meanings. When and how this process started is difficult to determine as no diachronic data is available.

Craig (1991:486) defines 'polygrammaticalization' as a "multiplicity of grammaticalization chains that may originate in one particular lexical morpheme". It involves an evolutionary process which eventually gives rise to one or more grammatical morphemes. Hopper (1991:22) outlines five principles of grammaticalization, of which two may apply to Kabba: layering and persistence. With 'layering' "new layers are continually emerging" while older layers remain "to coexist and interact with newer layers". With 'persistence' Hopper meant that "some grammatical traces of its original lexical meaning tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution". As a result of these processes and similar structures, Kabba discourse initially appears highly ambiguous and difficult to analyze.

Figure 1 illustrates the developmental relationships of some grammatical targets with their source lexeme in Kabba. It suggests that the benefactive emerged before the recipient. The causative emerged as a separate track preceding the resultative, the purposive, the complementizer, the sequential and the connective markers. According to T. Givón (p.c.) grammaticalization chains are "well-attested elsewhere around the world". Yap-Foong-Ha (2000), in her diachronic study of the polysemy of 'give' constructions in Malay over 400 years, identified separate paths for the dative-benefactive and the permissive/causative track. In Kabba these processes take place simultaneously and not completely independently of each other. There is some overlap of the grammatical functions, which are structurally and semantically interrelated. The Figure 1 indicates that a morphological reduction is taking place from two pronominal arguments to zero arguments. On these grounds the benefactive has been placed before the

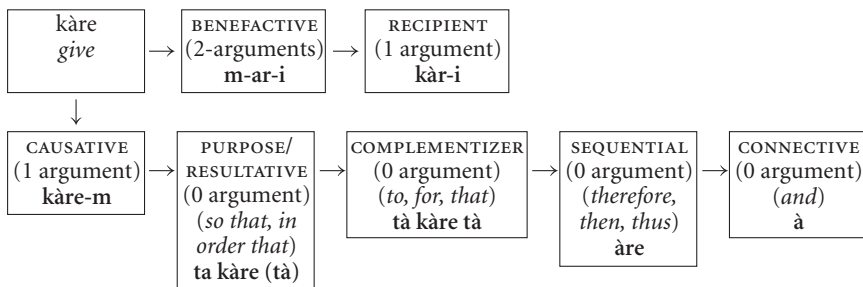


Figure 1. Grammaticalization chains of the verb 'to give'

recipient, although T. Givon (P.C.) believes that “the most common development is from ‘give’ to ‘recipient’ and from ‘recipient’ to ‘benefactive’.

2. The verb /kàrà/ ‘to give’

Before attempting an analysis of these diverse grammatical developments, the meaning, structure and function of the verb ‘to give’ needs to be fully understood in order to be able to distinguish between its lexical meaning and their grammatical realizations. Incidentally, the verb ‘to give’ is not the only lexeme that is undergoing grammaticalization in this language.

2.1 Initial /k-/ verbs

The verb ‘to give’ belongs to the group of verbs with an initial /k-/ consonant, which may be considered a fossilized remnant of a verbal class system. It deletes with conjugation and the verb undergoes tonal root changes. A considerable number of verbs belong to this group. The /k-/ appears to have lost its original purpose and meaning. It could be considered an infinitive marker for this particular group of verbs. However, as no lexical item starts with a vowel, its removal for classificatory purposes becomes phonotactically problematic. Giacalone and Hopper (1998: 3) consider such remnants indicators of tendencies of “certain consonants to be associated with very broad lexico-semantic classes”.³ Kabba has also verbs starting with /k-/ which do not lose their initial /k-/ with conjugation and thus do not undergo tonal or vowel harmony root changes. They include both transitive and intransitive verbs. These verbs conjugate like all the other verbs that do not start with a /k-/. In the following table the verb /kòko/ ‘to laugh’ is conjugated for the perfective, imperfective and future tense/aspects.

The verb /kàrà/ ‘to give’ is both transitive and ditransitive, but never intransitive. It functions frequently as a three-place predicate with three arguments. As well as being a lexeme in its own right, it is the source of a variety of grammatical functions, which will be discussed in this paper after an analysis of the verb ‘to give’ and its inflections.

Table 1. Conjugation of verbs with initial non-deleting /k-/

Perfective		Imperfective		Future	
<i>m-kòko</i>	I laugh	<i>maw kòko</i>	I am laughing	<i>má kòko</i>	I will laugh
<i>e- kòko</i>	you laugh	<i>aw kòko</i>	you are laughing	<i>á kòko</i>	you will laugh
<i>ń- kòko</i>	he laughs	<i>náw kòko</i>	he is laughing	<i>ná kòko</i>	he will laugh
<i>ń- kòko</i>	we laugh	<i>jàw kòko</i>	we are laughing	<i>jà kòko</i>	we will laugh
<i>e- kòko-je</i>	you laugh	<i>aw kòko-je</i>	you are laughing	<i>á kòko-je</i>	you will laugh
<i>ń- kòko</i>	they laugh	<i>dáw kòko</i>	they are laughing	<i>dá kòko</i>	they will laugh

2.2 Conjugation of /kàrà/ ‘to give’

The following table illustrates the conjugated forms of the verb /kàrà/ ‘to give’ in all three forms in which it occurs. It shows how the initial /k-/ drops off in the perfective and imperfective, but is retained with the future tense. The perfective and the imperfective are used for both the present and the past tense. All /k-/ deletion verbs conjugate like the verb ‘to give’, except for changes in the tone patterns, which vary according to the tones on the infinitive root. The tonal pattern for the verb /kàrà/ ‘to give’, which is Low-Mid in its infinitive form becomes: Mid-Mid, Mid-Mid, Hgh-Mid, Low-Mid, Mid-Mid, High-Mid when conjugated, as the following table illustrates for the perfective, the imperfective and the future tense/aspects. The verb ‘to go’ is used to form the imperfective and the marker /á/ is used for the future; both take inflections. For the first person plural and after a noun phrase the future marker takes a low tone:

Table 2. Conjugation of verb with initial deleting /k-/

Perfective		Imperfective		Future	
m-arə	I give	m-aw m-arə	I am giving	m-á kàrà	I will give
Ø-arə	you give	Ø-aw Ø-arə	you are giving	Ø-á kàrà	you will give
n-árə	he gives	n-áw n-árə	he is giving	n-á kàrà	he will give
j-àrà	we give	j-àw j-àrà	we are giving	j-à kàrà	we will give
Ø-arəje	you give	Ø-aw Ø-arəje	you are giving	Ø-á kàrà-je	you will give
d-árəje	they give	d-áw d-árəje	they are giving	d-á kàrà	they will give

The verb ‘to give’ takes bound subject pronouns prefixes. With the future the /k-/ is retained, but if the subject of the verb is a noun phrase the /k-/ is simply deleted. This applies as a general rule to all verbs whose /k-/ is deleted with conjugation. The mid tone in the second syllable does not change. With a subject pronoun prefix, the tone of the pronoun moves onto the root of the verb. The first person plural subject pronoun, which is a syllabic nasal /n̩/, becomes an affricate /j/, and the third person plural subject pronoun /ń/ becomes a voiced alveolar stop /d/ with verbs that delete their initial /k-/ such as the verb /kàrà/ ‘to give’. The second person singular and plural take a zero subject pronoun marker.

Direct object pronouns are the same as the possessive pronouns that are attached to inalienable nouns, except for the tones for the first and second person singular,

Table 3. Indirect object pronoun for the verb ‘to give’

kàrà + m	kàrà-m	give me
kàrà + i	kàr-i	give you
kàrà + é	kàr-é	give him
kàrà + jé	kàrà-jé	give us
kàrà + sé	kàrà-sé	give you
kàrà + dé	kàrà-dé	give them

which become mid. The indirect object pronoun suffixes are directly attached to the final schwa of verb, except those for the second and third person, where the schwa of the root deletes before the pronoun suffixes are attached. The third person /é/ becomes /é/ to comply with vowel harmony constraints.

Very often both subject and indirect object pronouns occur simultaneously on the same root, resulting in the following types of constructions:

Table 4. Subject and indirect object affixes

<i>M -ar -i</i>	'I give you'
<i>1S-give-2S</i>	
<i>Ø-ar -é</i>	'you give him'
<i>2S-give-3s</i>	
<i>n -árə -jé</i>	'he gives us'
<i>3S-give-1P</i>	
<i>j -àrà -sé</i>	We give you'
<i>1P-give-2P</i>	
<i>Ø-arə -dé</i>	You give them'
<i>2P-give-3P</i>	
<i>d -árə -m</i>	'they give me'
<i>3P-give-1S</i>	

2.3 Syntax and semantics of /kàrà/ 'to give'

The basic syntactic structure in which *kàrà* 'to give' occurs is: S+V+IO+DO. Newman, (1996:15) observes that the verb 'to give' is a source of "metaphysical extensions". With its "abundance of non-literal uses", it is a "salient component of human experience, employed to help conceptualize various acts or events". The verb 'to give' makes reference to two human participants. As in many languages, 'to give' in Kabba is syntactically a typical ditransitive verb consisting of three salient entities: an agent, a patient (direct object) and a recipient or benefactive (indirect object). However, the patient and the recipient are not always explicit, as the following examples illustrate. The patient is absent in the imperative:

- (1) Ø- arə- m
2s give- me
'Give-me!'

In example (2) the dowry is the patient. The recipient is implied. The verb takes inflections for the third person plural subject pronoun.

- (2) d -árə nàrè kòlè
3p give money dowry
'They gave the dowry money'

Much more frequent are examples with all three arguments expressed explicitly. The subject prefix is a pronoun, whereas the direct and indirect objects compound noun phrases. The direct object precedes the indirect object.

- (3) *d -ára né kùsà njé mba- je*
 2P give thing to eat persons visitor- P
 'They gave food to strangers'

The verb /*kàrə*/ 'to give' is a frequently occurring verb in conversational discourse where it is often used as an imperative to request something. The agent takes a zero marker. The recipient pronoun precedes the patient pronoun:

- (4) *Ø- are- m nàrè lé*
 2s give me money that
 'Give me that money!'

In the following complex structure the verb 'to give' occurs in the relative clause at the beginning of the utterance. The verb in the main clause is 'to come' and it is followed by the anaphoric pronoun *né*, which means 'thing' and acts as a direct object referring to the item that had been given to them before:

- (5) *né ké nd-é ténn d-ára-je lé, e-ddèe né wà?*
 thing REL day-3s LOC 2P-give-1P DET 2s-come ANA QUE
 'That thing that they gave us the other day, did you bring it?'

In the next example the verb 'to give' occurs in a question and its recipient appears as a noun followed by a relative clause before the interrogative pronoun at the end. This is a serial verb construction consisting of the verbs 'to send' and 'to give'. The transitive verb 'to send' is followed by the anaphoric pronoun /*nè*/:

- (6) *Tà kùlà né kàrà njé ké ddá?*
 to send ANA give person REL which
 'To send it to give to whom?'

In the following example /*kàrə*/ could be interpreted as both the verb 'to give' and the dative case marker 'for you'. The patient constitutes a noun preceding the verb 'to give':

- (7) *m-bbòkò nàrè tà kàr-i*
 1s-stole money PUR give (DAT)-2s
 'I stole the money to give to you' ('I stole the money for you')

The verb 'to give' is rarely used figuratively. Nor is it used with reflexive, impersonal, passive, or intransitive constructions. The following example comes from a proverb where the verb 'to give' has no figurative meaning. The proverb itself implies that a man does not take a wife and give her to another:

- (8) *Bàw bisi ùnn singa àw né àrè njè màré ààng.*
 male dog take bone go ANA give person other not
 'A male dog never takes a bone to go and give to another'

There is an element of cause and effect involved in the verb ‘to give’ which requires a recipient, but not a patient. Unlike other verbs it only takes subject prefixes and indirect object suffixes, never direct object suffixes. However, it may be preceded by a subject noun phrase and followed by an object noun phrase.

3. /Kàrə/ as a dative case marker

The basic syntactic structure of a dative construction consists of two clauses, one followed by a direct object and the other by an indirect object: A+V+D.O.+ A+DAT+IO+(DO). According to Comrie (1981:174) the dative case marker is the “typical exponent of experiencer or recipient”. It expresses indirect object relationships by means of pronominal inflections attached to the dative marker. The Kabba use the verb ‘to give’ as a dative marker of recipient and benefactive relationships. This dative marker occurs in simple constructions and takes inflections for subject and indirect object like the verb ‘to give’. Newman (1996:82) speaks of a dative when “the case is used prototypically to mark a ‘recipient’ phrase in ‘give’ constructions”. After the verbs ‘to give’, ‘to say’, ‘to tell’, ‘to ask’, ‘to want’ and ‘to send’ the dative marker functions as a recipient, but after transitive verbs like ‘to find’, ‘to sift’, ‘to crush’ as a benefactive. This comparative analysis of the benefactive and recipient dative structures reveals some morphological reduction from two pronominal arguments with the benefactive to one with the recipient.

3.1 /Kàrə/ as a benefactive

The benefactive root of the verb ‘to give’ is accompanied by both a subject pronoun prefix and an indirect object pronoun suffix. Example (9) consists of two clauses, each with a subject, verb and object. Literally this sentence says ‘I found a wife I gave him’. Semantically, however, the verb ‘to give’ has become a benefactive marker with the grammatical meaning of ‘on behalf of’. Following the verb ‘to find’, this could be considered a serial verb construction:

- (9) *M-ínga dèné m-ar-é*
 1s-found wife 1s-DAT-3s
 ‘I found a wife for him’

The benefactive is frequently used in questions. In example (10) it forms a syntactic structure consisting of a subject, a verb and an object. The dative construction is preceded by a transitive clause. The verb ‘to sift’ together with the benefactive ‘for the benefit of’ could be considered a serial verb construction:

- (10) *n-dàle ndùju lé n-ár-i-ì?*
 3s-sift flour DET 3s-DAT-2s-QUE
 ‘Did she sift that flour for you?’

In example (11) the dative is used with the first person plural subject pronoun and the third person singular indirect object pronoun. The first person plural pronoun could be considered the agent of the dative construction with a recipient function meaning 'we give him'. However, following the verb 'to crush' the dative also forms part a serial verb construction with a benefactive meaning:

- (11) *ñ-gáji àr j-àr-é*
 1P-crush stone 1P-DAT-3S
 'We crush stone for him'

In example (12) the dative is embedded in a relative clause. It follows the verb 'to send'. As with all the previous examples, the benefactive is fully inflected for both subject and indirect object pronouns:

- (12) *Ddém m-ínga né-je ké d-úla né d-áre-m lé ke màjé njáá*
 also 1s.get thing.P REL 2P.sent ANA 3P.DAT.1S DET VA well really
 'Also I received the things well that they sent me'

3.2 /*kàrə*/ as a recipient

With the future tense, which may imply potentiality, intention or irrealis aspects, the verb 'to give' occurs in its infinitival form *kàrə* preceded by the purposive particle *tà*. In example (13) it is followed by the dative case marker /*àrə*/ to which the second person singular indirect object pronoun suffix /-i/ is attached. This example illustrates its recipient meaning. It forms part of a ditransitive verb phrase which follows a transitive clause whose direct object 'God' becomes the implied subject of the second part of the clause. Following the verb 'to ask' the dative marker could be considered a preposition. A full agent-verb-object clause precedes the dative construction which is followed by a direct object:

- (13) *M-dùjù Lúbba tà kàre àr-i singa*
 1s-ask God PUR give DAT-2S strength
 'I ask God to give you strength'

In example (14) the verb 'to give' occurs in the infinitive with the meaning and function of a dative case marker. As it occurs in a position immediately preceding the recipient 'Dominic', it could be considered a preposition. Following the verb 'to say' or 'to tell' and preceding a proper noun, the dative marker *kàrə* occurs in its uninflected form:

- (14) *N-áw tà pà tàr kàrə Dómìníc*
 3s-go to tell story DAT Dominic
 'He is going to tell the story to Dominic'

Example (15) is a complex construction consisting of two clauses separated by an adverbial phrase. The second clause contains a anaphoric pronoun as well as the pur-

posive marker *mba* which is an abbreviation of *mbatà* ‘because’. The dative marker occurs in the initial subordinate clause and functions as a preposition. It is followed by an adverbial phrase, a main clause and another subordinate clause. The agent is the same throughout:

- (15) *M-ndìkì tà pà kàr-i mè dàmàsì té nèénn m-íla ri-m gál*
 1s.want to tell dat.2s in week LOC this 1s.put name.1s times
joó mba kùsà né nèénn.
 two PUR eat ANA here
 ‘I would like to tell you that for this week I put my name down to eat here twice’

The dative often follows the verb ‘to tell’, which is preceded by the purpose marker *tà*. In example (16) the subject is a noun phrase and the object a proper noun. The dative carries no inflections and acts as a preposition to the recipient Jean. The particle *to* between the two names functions as a complementizer. The initial subject clause is an idiomatic expression consisting of a inflected inalienable noun followed by the verb ‘to die’ meaning ‘I forgot to tell’:

- (16) *mè-ém wòy tà pà kàrà Jean to Háns dá í-ddè pán*
 T-1s died PUR tell DAT Jean COM Hans FOC 3p.come already
 ‘I forgot to tell Jean that Hans has already come.’

The dative marker may also be embedded in an extended relative clause followed by an adverbial phrase. The main clause consists of the verb ‘to escape’ with the meaning ‘to forget’. Its indirect object suffix is followed by an adverb. This is an idiomatic expression with the literal meaning, ‘What you told me the other day completely escaped me’:

- (17) *tàr ké e-pà àre-m ndé té-nn lé dá ìki-m nyòm*
 talk REL 2s.tell DAT.1s day LOC.DEM DET FOC escape.1s completely
 ‘I completely forgot what you told me the other day’

Thus these examples demonstrate the grammaticalization process of the verb ‘to give’ to function as a dative case marker to express benefactive and recipient relationships. The agent and the recipient appear to be always +human. Signs of a morphological reductions occurs when the dative case marker, which functions as a benefactive takes pronominal affixes for both subject and indirect object pronouns, just like its source verbs ‘to give’; whereas the recipient only takes suffixes for the indirect object pronouns. A discussion of whether the benefactive forms part of a serial verb construction and whether the recipient constitutes a preposition exceeds the space allotted for this paper. If a pronoun could be considered a reduction of a noun or noun phrase, then the recipient could be listed before the benefactive on the evolutionary scale.

4. /*kàrə*/ as a causative marker

The causative marker forms part of the basic syntactic structure: clause 1 + *kàrə* + clause 2. The causative has also been called 'allower', 'causer' and 'enabler'. According to Newman (1996: 171) "the meaning of literal 'give' has some connections, semantically, with two important notions in the study of language and logic: causation and enablement. It requires a "human agent acting on a patient" (172–173). The giver is considered the causer and the recipient the causee. In figurative usage the verb 'to give' may refer to non-animate things. In Kabba polysemy takes place when the verb 'to give' retains its form when it is used for grammatical purposes. This may initially cause ambiguity and loss of meaning.

According to Comrie (1981: 158) "any causative situation involves two component situations, the cause and its effect (result)". Causative constructions are structurally independent and express themselves in a number of different ways. Talmy (1976: 47) points out that the "term causative in a semantic analysis of language must first be distinguished from the scientific notion of causation in the physical world". According to Talmy a basic causative situation "consists of a simple event, that which immediately causes the event, and the causal relation between the two" (1976: 52). Comrie (1976: 296) talks about the "morphological synthetic causative" as the "clearest variety of a causative construction where causative and embedded verb are fused into one in derived structure". This appears to be the case in Kabba where the causative construction is a perfect copy of the verb 'to give'.

Chappell's (2001: 262) analysis of Southern Min dialects provides another confirmation of the cross-linguistically attested (Heine et al. 1993) process of grammaticalization of the verb 'to give', involving the "semantic development from give > dative > causative > passive." In Southern Min the causative constructions are of the permissive type, "expressing that the causative agent let, allowed or enabled the situation to happen". In these Chinese dialects the causative marker occurs with verbs that are "semantically stative", such as 'to know', 'to believe', 'to receive' and with the active verb 'do' in the imperative.

In Kabba the causative construction occurs with intransitive stative verbs such as 'to die', 'to fall', 'to go', 'to come', 'to see', 'to sleep', 'to return', and with some verbs that can be both transitive and intransitive such as 'to do', 'to eat', 'to keep' and 'to learn'. The verb *kàrə* 'to give' retains its phonological form with the future tense. Elsewhere it deletes the initial /k-/ and takes affixes for subject or object pronouns. The causative marker always follows a full clause, but is not always followed by a full clause.

Dixon (1982: 122–123) found the verb 'to give' to have "strong grammatical connections" occurring in most languages, indicating a "transfer of ownership from one person to another". He suggests that "at the semantic level this verb is effectively the causative form of a basic grammatical relation", and as "such it stands apart from the other nuclear words in the lexicon". This explains how the verb 'to give' can be grammaticalized as a causative marker. In their typology of causative constructions, Dixon and Aikhenvald (2000b: 62–73) outline nine semantic parameters related to causative

constructions. State/action, transitivity related to the verb, control, volition, affectedness to the causee, directness, intention, naturalness and involvement to the causer. Example (18) illustrates the verb ‘to give’ in a simple causative construction referring to an action which is to take place in the future and which contains an element of warning or a threat. The verb ‘to give’ has become a causal verb in its own right. The cause and effect relationship is explicit. The causee consists of a bound pronoun subject prefix and an intransitive process verb. The causer consists of an agent-verb-object and the causee of a subject-object structure. Since the verb ‘to cause’ is more semantically ‘bleached’ than ‘to give’, we could still consider the process involved to be grammaticalization. In the following example the causative marker *kàrà*, which takes a direct object pronoun, precedes the verb ‘to fall’. The causer of the event is the subject of the first intransitive clause. The causee is the subject of the second intransitive clause:

- (18) *m-á kàrà-é n-óso.*
 1S-FUT CAU-3S 3S-fall
 ‘I will make him fall’

The causer in the next example is a woman who hits a child on the backside to cause it to go to sleep. No inflections appear on the causative marker and the stative verb ‘to sleep’. The patient of the transitive clause is the same as the implied subject of the intransitive verb phrase. The agent of the causative verb ‘to smack’ is the woman. The child is the causee, who is expected to go to sleep after a good hiding. Thus the meaning of *kàrà* is both causative and resultative. The child is both recipient and agent:

- (19) *Dèné àw gól kùtù ngonk kàrà tìbbi*
 woman IMP smack buttocks child CAU sleep
 ‘The woman is smacking the child’s backside to make it sleep’

Another cause and effect relationship is expressed in the following construction involving reported speech. In the introduction the mother, who is the causative agent, is talking to the reporter, who is the causee asked to present his newly acquired wife to his mother. An element of subjunctivity is involved when the mother expresses the desire to see her daughter-in-law. This example contains both a dative and a causative marker and consists of three clauses. The causee is also the agent who enables the mother to see the young woman. The causative construction precedes a transitive clause:

- (20) *kón-je pà àre-m pà nà m-aw kè wùm- nè-je*
 mother-HON said dat.1s say COM 1s.go CON daughter.in.law.3s.HON
m-are n-ò-é
 1s.CAU 3s.see.3s
 ‘My mother told me that I (must) go with her daughter-in-law so that she sees her’

The causer of the following utterance gives reasons for her actions and the results she expects. Here the first causative marker expresses enablement. The enabler is the woman and the seeds the enableee for the wind to blow upon. Thus the wind is both

the enablee and the causative agent. The expected result is that the thing would not go bad. The speaker is a lady in her forties and the place is Bangui, the capital of the Central African Republic. The second causative is followed by an intransitive verb without a subject:

- (21) *m-nàjì dò né té m-arə yél ùlà té kàrə ndù àáng*
 1s-dry on thing LOC 1s-CAU wind blow LOC CAU decompose NEG
 'I dried it on a bag for the wind to blow so that it would not decompose'

God is the causer in a folktale about the origin of death. It was written by a man in his late thirties. *kàrə* is being used as a causative marker with the meaning of 'to make' followed by an direct object pronoun and a intransitive clause:

- (22) *Ngà lúbba pàna nà kàrə-dé dáy ngájí*
 so God said COM CAU-3P 3P-die little
 'So God said that he was going to make them die a little'

In example (23) the causative marker occurs as an elliptical structure or an afterthought at the end of utterance. It functions as a noun and is followed by a logophoric possessive pronoun referring to the mother. Thus, *kàrə nè* simply means 'her cause'. The author reminisces about his mother's life after her death. When the father died she had made the decision to do something before her eight children died from hunger. When the author uses the logophoric low tone possessive pronoun in the ellipsis, he is indicating that the mother would have been the indirect cause of her children's death:

- (23) *ñ-pàna bbo ñ-mbón jì nè dòkù n-ìsì à jéje ngánn*
 3s-said if 3s-assemble finger 3s ten 3s-stay and we children
nè j-à kòsò làbbà kàrə nè
 LOG.3s 1P-FUT fall.down hunger CAU LOG.3s
 'She said that if she stayed with her ten fingers crossed, we, her children, would die from hunger, caused by her'

A pastor in his fifties used the following example during a church message. A son is asking his father's permission to cultivate a small piece of land. The causative marker functions as an enabler and carries the meaning of 'to let', 'to allow' or 'to make'. It occurs at the beginning of the main transitive clause:

- (24) *Ngonn dá nà bàann àáng, Ø-arə ñ-ndò kòbbu éya kété*
 child FOC said like.this not 2s-CAU 1P-dig portion small first
bbá
 please
 'The child said, 'That's not it please. Allow us to cultivate the small portion first.'

In the story of 'The lion, the hunter and the rat', Baro, a man in his late fifties, uses the following complex construction consisting of a main clause followed by a negative

relative clause in which the causative structure occurs. The causative *kàrà* is preceded by *àse* meaning ‘capable’. The lion considers the rat incapable of causing his death. The causer in this structure is the rat and the causee the lion. The structure of the relative clause is as follows:

- (25) *n-pàrà n-á tɔl yàkə lé to dèw ké àse kàre n-tɔl-é*
 3s-said 2s-FUT kill rat DET be person REL suffice CAU 3s.kill.3s
àáng tò
 NEG also
 ‘He said that he will kill the rat, who is a person not capable of killing him’

Example (26) comes from the story about the hare and the hyena. It took place during a time of famine, causing the wild animals to search for food elsewhere. The storyteller is a high school teacher in his late thirties. The causative marker is preceded by the intransitive verb ‘to fall’. The complex structure consists of an idiomatic introductory clause, a discontinued main clause and an embedded relative clause. The causative marker links the introduction to the first part of the main clause:

- (26) *Lew dá bbo làbbə òso àrà da-je ké mè mù té dá*
 long.ago FOC if hunger fell CAU animal-s REL in forest LOC FOC
aw-je njàá bè pàlèpàlè mba né kùsà
 go-P even like elsewhere PUR thing to.eat
 ‘Long ago, there was a famine causing animals to go elsewhere for something to eat’

Yaita is the name of the hero of the story from which example (27) comes. The king issues a challenge and promises a reward to the soldier who succeeds in killing the river monster. Thus he sets up a cause and effect relationship. Yaita takes up the challenge, kills the beast and reaps the reward. Here, the causative *kàrà* has the meaning of ‘to make’ or to ‘bestow’. It is preceded by a pronoun, the future aspect marker, and followed by the stative verb ‘to be’. The causer is the king and the causee the soldier who kills the dragon. *kàrà* occurs in the second main clause in the following structure:

- (27) *Mbáy ké njè kòn bbe pàrà ‘Dèw ké tɔl da màann*
 chief REL person govern place said person REL kill animal water
lé, à n-á kàrà n-to mbáy lé mbámábá-je tóyn’
 DET CON 3s-FUT CAU 3s-be chief DET soldier-s all
 ‘The chief who governs the place said that he would make the person who killed the water dragon the major general of the army’

A lady in her late fifties tells the story of a bird that has the magic power to make flour. The causative has both a purposive and an enabling function. The bird is being asked to demonstrate that it can really make flour so that the people will be able to see it. Thus the bird is the causer and the people are the causee. The causative is followed by the verb ‘to see’. The structure of the main clause is as follows:

- (28) *à n-pà-i nà 'sá e-dda né dda nduju lé àrə j-òo mà'*
 CON 3s-say COM OK 2s-make ANA make flour DET CAU 1P-see then
 'And he said, 'OK you make that flour for us to see, then'

Thus, the causative marker, which resembles the dative case marker, has many functions and occurs in a large number of different contexts. It also occurs in structures consisting of more than one morpheme.

5. *Tà kàre (tà)* as purposive/resultative-causative(-complementizer)

The basic syntactic structure of this construction is: clause 1 + *tà kàrə* + clause 2. To express the meaning of 'so that', 'for' or 'that' the verb *kàrə* 'to give' frequently combines with the purposive marker *tà*, which is also used as a complementizer and is related to *mbatà* 'because' and its abbreviation *mba*. The construction *tà kàrə* occurs in both conversations and narratives. It is found as a link between two clauses, the second one being the consequence of the first one. This construction does not take any inflections and the initial /k-/ is not deleted.

5.1 /*Tà kàrə*/ as a purposive resultative/causative

Example (29) contains an element of purpose and expected result. The purposive marker *tà* precedes the causative marker, which is followed by the complementizer *tà*. The main verb is an intransitive motion verb. The causative applies to the process verb 'to learn'. The causee is Dominic who may or may not participate willingly. The causer is the person whom his parents brought to teach him. *Kàrə* is preceded by a intransitive clause and followed by an ditransitive one:

- (29) *n-ddè kè dèw tà kàrə ndó né Dominic*
 2P-came with person PUR CAU learn thing Dominic
 'They came with a person to make Dominic to learn something'

The construction *tà kàrə* is used to express a purpose or an intention. A shot is fired to make people afraid. *Tà kàrə* separates a ditransitive clause from a transitive clause:

- (30) *n-ddɔ pùrù dɔ-dé té tà kàrə n-bbeél*
 2P-shot fire head-2P LOC PUR CAU fear
 'They fired at them to make them afraid'

The verb 'to give' may occur simultaneously with the dative case marker and the purposive/ causative in the same sentence, as examples (31) illustrates. There is no causer and no causee in this complex structure, which contains the root of the verb 'to give' four times, but only twice with its original meaning. This is a polite inquiry containing three indirect objects or recipients:

- (31) *d-ár-i bàá wàsé d-ár-i tà kàrà e-tél né àre-dé o?*
 3P-give-2s simply or 3P-give-2s PUR CAU 2s-return ANA DAT-3s QUE
 'Did they simply give it to you or did they give it to you to return to them?'

In example (32) an intransitive clause precedes the causative structure with which it form an integral part. It is followed by an transitive clause. The speaker expresses the wish for someone to perform a certain action before it starts raining:

- (32) *m-ndikì tà kàrà e-tóko kubbu-je lèém kété bbá ndi à kàrà*
 1s-want PUR CAU 2s-wash clothe-s my first before rain FUT fall
 'I would like for you to wash my clothes first before it rains'

Example (33) comes from a man in his early forties. He uses *kàrà* together with the reasoning marker *mba* twice, the first time as a complementizer and the second time as a purposive. God is being asked to heal someone and to enable him to do the work. The object of the agent-verb-object clause becomes the implied subject of the (agent)-verb-object clause:

- (33) *M-dùjù Lúbba mba kàrà òru rɔ to rɔ sé mba kàrà*
 1s-ask God PUR CAU remove body suffer body your PUR CAU
e-dda-i-je né kùlà ké n-ùla sé
 2P-do-VOC-P thing work REL 3s-send 2P
 'I ask God to remove the illness from your body, so that you can do the work that he sent you to do'

The speaker of example (34) is a man in his late thirties. He is urging the old lady Deboura to describe how to prepare the gluey sauce, which is a Kabba specialty reserved for special occasions. He gives as a reason the fact that many young women no longer know how to prepare this special dish. The utterance starts with the sequential discourse marker /àrà/ (cf. (7)) *Tà kàrà* separates an intransitive clause, of which the causer is the subject, from an transitive clause with the causee as the agent. The utterance ends with a relative clause:

- (34) *àrà ñ-ndikì tà kàrà e-tɔji-je kúl gém bbe ké lè*
 CAU 1P-like PUR CAU 2s-teach-1P make sauce place REL GEN
Kabba-je ké d-áw kúl gém
 Kabba-s REL 3P-IMP make sauce
 'That's why I want to urge you to teach us the preparation of the gluey sauce of the village which the Kabba make'

The sheep, the dog and the goat travel together. At the end of the journey the sheep pays his fare and gets off, the goat jumps off without paying, and the dog pays, but does not receive his change. This example contains a passive clause. The agent, that is the chauffeur, remains implicit and the subject of the first clause becomes the anaphoric object of the second clause:

- (35) *àkàá kàsə nàrè làá nàyn tà kàrə ní-tél né d-ár-é bbáy*
 but left money his stayed PUR CAU 2s-return ANA 3p-DAT-3s yet
 'But his change remained to be returned to him yet'

Example (36) comes from the story of the tortoise and the toad. It is told by a man in his late thirties. He uses *tà* as a purposive and *kàrə* as a causative. The first clause is intransitive and the second one transitive. The subject of the first clause goes into indirect object function in the second clause:

- (36) *á yàá m-ddèe tà kàrə Ø-are-m kùma kàre lè-i sén*
 FOC REA 1s-come PUR CAU 2s-give-1s medicine strong GEN-2s little
 'That's why I came, so that you can give me some of your strong medicine'

The causative *kàrə* may be followed by the complementizer *tà* which appears to have resultative implications with the meaning of 'as a consequence'. In the following example a young professional in his thirties uses this construction when he writes about the importance of going to school so as to learn to do things like the white men. In other words, the result of going to school is to be able to do things like a white man. The causative construction introduces the main intransitive clause, which is preceded by a lengthy subordinate clause. The causer is also the causee. The main clause has the following structure:

- (37) *Tà dda né toké lè nàsàrà-je lé dá tò mājì bòi kàrə tà*
 pur make thing like gen European-s det foc be good very res com
éyi dèw àw làkšòr té
 you people go school loc
 'To do things like the white men, it is good that you people go to school'

Example (38) comes from the same author. It has both a causative and a purposive meaning. The elephant is trying to find someone to shave his head. The monkey excuses himself hoping to make the elephant look for someone else. The storyteller uses the reasoning marker *mbatà* 'in order to' followed by the idiomatic expression *kùwà gól* 'seize foot' to say 'I am sorry'. The monkey is the causer and the elephant the causee. The elephant is the patient of the first clause and the agent of the second clause:

- (38) *ní-pà bè mbatà kùwà né gól kàrə kàrə tà sáŋge màre*
 3s-spoke like PUR seize his foot elephant PUR COM search other
dèw ke kété
 person VA first
 'He spoke like this to excuse himself to the elephant so that he (the elephant) could find someone else (to shave his head)'

5.2 /Tà kàrà tà/ as purposive-causative-complementizer

This purposive-causative-complementizer construction occurs between two clauses. It functions as a complementizer with the meaning of ‘for’, ‘so that’ and ‘in order to’. It also carries the meanings of ‘to have to’, ‘to make’ and ‘to cause’. In this construction *kàrà* is preceded and followed by the particle *tà*, which takes its origin from the noun ‘mouth’ or ‘opening’. Before *kàrà* it functions as a purposive; after *kàrà*, it serves as a complementizer as example (39) illustrates. The snake is the causer and the young man is the causee. This example comes from a folk story called ‘The Snake’. It is told by a young man in his early twenties. The causative occurs between two transitive clauses:

- (39) *Li dá aw sá ddéw bbàsinè tà kàrà tà n-tɔl naám*
 snake FOC IMP look chance now PUR CAU COM 3s-kill young.man
lé
 DET
 ‘The snake now looked for an opportunity to kill that young man’

The next example expresses the desire and the necessity to remain in touch through the recorded voice. The old lady is the causer; the voice is the causee. The causative takes the meaning of ‘to make or ‘to cause’ the voice to go inside the recorder and to allow it to travel. The complementizer introduces a new clause:

- (40) *á yàá m-aw pà tàr tà kàrà tà ndu-ɾɿ lé tè mè*
 FOC REA 1s-IMP tell story PUR CAU COM voice-1s DET arrive in
né té làá nèén lé n-áw núnn né ndu-ɾɿ lé.
 thing LOC her here DET 3s-go over.DEM ANA voice-1s DET
 ‘That’s why I am talking, so that my voice goes inside her thing (recorder)
 here to go with her over there’

Another speaker is using the /tà kàrà tà/ construction immediately after the sequential marker *bbá* ‘before’. The agent-verb-object clause is followed by an subject-verb clause. Both are independent structures, one using the verb ‘to want’ and the other the verb ‘to go’. The causative construction occurs between the two with the meaning of ‘to have to’. The causer is implicitly the French army:

- (41) *jè sà-á dá n-ndìkì nàa njàá wáké bbá tà kàrà tà*
 1p with-3p FOC 1p-like REC even spontaneously before PUR CAU COM
n-áw lò né ndó té lé à m-ɾru gu-é
 3s-go place thing learn LOC DET CON 1s-follow after-3s
 ‘We liked each other spontaneously before he had to go to the place of learn-
 ing, and I followed him later’

An old pastor in his seventies used this causative construction in his prayers. The causer is God and the causee, who is not mentioned or referred to directly, is the person who has to do what God decides. The causative construction occurs between two intransitive clauses using the verbs ‘to ‘decide’ and ‘to stay’. It functions as the com-

plementizer ‘that’ and carries the meaning of an all-powerful God, in whose hands is both purpose and cause with the ultimate decision making power. A conditional clause with the complementizer *nà* ends with another causative construction with God as the causer:

- (42) *n-ó à tà kàrà tà Ø-ìsì à Ø-á kisi bbo n-ó*
 3s-decide CON PUR CAU COM 2s-stay CON 2s-FUT stay if 3s-decide
nà Ø-aw à tò ngeng nya tà kàrà Ø-ìsì tò
 that 2s-go then be difficult INT PUR CAU 2s-stay also
 ‘If He (God) decides that she must stay, she will stay; but if He decides that she must go, it will be very difficult for her to stay as well’

Lydie is a pastor’s daughter in her sixties, an excellent story teller. She told the story about ‘Esu and Wasp’ to a group of children in her village. This example consists of a discontinued main clause with an embedded relative clause which contains the *tà kàrà tà* structure. It is preceded by an agent-verb-object and followed by (implied agent)-verb-object clause. The causers are Esu and Wasp and the causee the child:

- (43) *à ngonnn ké n-ùnd-é-je nàng-é tà kàrà tà ngem ndìrà*
 CON child REL 1P-left-3s-PLU ground-LOC PUR CAU COM guard root
kake-nn lé dá à gèrè àáng
 tree-DEM DET FOC FUT know NEG
 ‘And the child whom we left on the ground to guard the root of that tree will not know’

6. /àrà/ and /à/ as discourse markers

The basic syntactic structure for theses constructions is: (clause 1) + *àrà*+ clause. Some Kabba like to use *kàrà* with the meaning of ‘so that’, ‘therefore’, ‘because’ or ‘that’s why’ at the beginning of an utterance or between two full clauses, as the next example shows. However, the initial /k-/ is always deleted in this position and thus appears to be an important criterion for a grammaticalization chain.

6.1 /àrà/ as a sequential ‘thus’, ‘so’

Example (44) contains a sequential marker. It comes from a hortative discourse by a pastor who admonishes the children to pay careful attention. Because they listened well on a previous occasion, he knows that they will listen well again. There is no causer or causee involved. *àrà* precedes two intransitive clauses:

- (44) *Àrà m-gère màjì kòò tàr lèsé lé à tò tokú là yè ké*
 SEQ 1s-know well listen talk your DET FUT be like GEN one REL
dòngòr lé bbáy
 before det again
 ‘Thus I know well that your listening is going to be like before’

Baro, who lost one arm and his eyesight from a dangerous snake bite, it a great talker and a store house of knowledge. He often uses *àrà* at the beginning of an utterance as discourse marker to make a pause or for a moment of reflection. *àrà* is followed by a subject pronoun and a relative clause. It functions as a discourse marker preceding a word of advice, which was given to him when he had to assist with the burial of his grandfather as a young man:

- (45) *àrà éyi ké njè dōbb-é lé Ø-ande mè buwá-té lé kété*
 RES you REL person bury-3s DET 2s-enter inside hole-LOC DET first
pán
 INT
 ‘That’s why you, who are the person to bury him, you enter into the hole first’

The next example has hortative implications when the young man is being instructed how to prepare to bury his grandfather’s body. It must be done in the correct manner. Certain rules have to be observed and a series of actions must be performed. *àrà* functions as a connective between a series of transitive clauses, which it introduces. It also has subjunctive meaning of necessity ‘that’s why you must’:

- (46) *àrà Ø-úla éningá, e-pà tàr, e-dòn gùmán*
 SEQ 2s-wear bracelet 2s-tell story 2s-bite corpse
 ‘That’s why you wear bracelets, you tell a story, you bite the corpse’

Clément, the young intellectual, combines the discourse function of *kàrà* as a sequential marker linking a series of propositions with resultative connotations. The story is the ‘The Hunter and the Crocodile’. To cause the hunter to have mercy on him, the crocodile pretends to be dying of thirst, because he is too old to go to the river for water. Each *àrà* is followed by a complex clause starting with an abstract noun: ‘strength’, ‘means’ and ‘thirst’. The sequence of events are: an old man > who has no strength > cannot walk to the river for water > he will die of thirst. Thus *àrà* also functions both as a connective and a sequential marker:

- (47) *à màr lé pàrà, ‘Ema dá m-bbúka nya ngàa, àrà*
 and crocodile DET said 1s.EM FOC 1s-old INT now RES
síngá-m gòtó, àrà ddéw ké tà nja tèe né tà-màann-té
 strength-1s absent RES means REL PUR walk arrive INS bord-river-LOC
gòtó, àrà kùnda njàa à tà tɔlu-m
 absent RES thirst even FUT PUR kill-1s
 ‘The crocodile said, “Me, I am very old now, so that I have no strength, and I can no longer walk to the river, therefore I will die of thirst”’

In example (48) Deboura talks about the death of her husband while he was doing military service for the French. She uses *àrà* as a sequential discourse marker with resultative implications. The husband died as a result of having been taken into military service. *àrà* occurs after a transitive serial verb construction:

- (48) *Dìngàw.lé ùnnu.m j.àw mbámábá.té àrə ngà-rh.le gòtó mè*
 man.DET take.1s 1p.go military.LOC RES man.1s.DET die in
mbámábá.té.lé
 military.LOC.DET
 'That man took me to go with him to military service, as a result my husband died'

6.2 /à/ as a connective 'and, 'then'

The very common connective discourse marker could be considered a phonological reduction of the verb *kàrə* 'to give'. It occurs at the beginning of a clause and as a connective between clauses or phrases. In example (49) it takes the meaning of 'then' and occurs between a clause and verb phrase:

- (49) *aw úla dèw-je e-pàrà: bbo dèw á wòy à ndó bbo*
 go transmit people-s 2s-say if person FUT die then wake up if
neyn á wòy à wòy to dɔ̀bɔ̀n.
 moon FUT die then die COM for.good
 'Go tell the people that if the people die, they will wake up; but if the moon dies it will die for good.'

In example (50) a pastor uses *à* frequently with the meaning of 'and' at the beginning of a new clause in his hortative discourse. However, it also takes the meaning of 'but' when introducing an anti-climatic intransitive clause:

- (50) *ń-pà nà n-àw à nàw pà tàr lé n-àre éndòrí-je lé, à*
 3s-say COM 3s-go and 3-IMP tell story DET 3s-DAT termite-s DET but
éndòrí-je tàr né táji á ùnn pá
 termite-s story thing insult FOC start talk
 'He said that he went and told that message to those termites, but the termites uttered insults'

7. Conclusion

The Kabba people, both young and old, male and female, educated and uneducated are using grammaticalized forms of the verb 'to give'. It occurs in all types of discourse, but most frequently in narratives and conversations.

Kàrə as a dative or recipient marker is obligatory. It occurs with ditransitive verbs like 'to give', 'to say', 'to send', 'to tell', 'to ask' and 'to name' and functions like a preposition. Both the agent and the recipient are +human. The dative marker retains the /k-/ with future tense and irrealis/potential aspect. The benefactive occurs with verbs like 'to find', 'to sift' 'to crush' where the dative/ benefactive marker forms part of a serial

verb construction. With the benefactive both the subject and indirect object pronouns are affixes. Morphological reduction of pronominal arguments takes place between the benefactive and the recipient.

Kàrə has a basically causative function with the meaning of 'to cause', 'to make', 'to bestow', 'to allow', 'to enable', 'to force' and 'to kill'. It is often used as a threat of a potential action in the present or in the future. Both intransitive and transitive clauses may follow the causative construction, which occurs at the beginning, in the middle or between main clauses, but also in embedded relative clauses. It may also form part of an ellipsis and take pronominal inflections for the subject and for the object. It occurs in its infinitive form after the purposive markers *tà* or *mba* and after the future tense marker. Without the initial /k-/ it occurs following a pronoun, a noun, a noun phrase or a clause. It takes inflections for subject and direct object pronouns.

Kàrə frequently combines with the purposive *tà* either before or after. It may precede and follow a full clause. Agent-verb-object is the preferred structure for the causer and (agent)-verb-object for the causee clause. The verbs associated with the causer are 'to come', 'to fire', 'to ask', 'to want', 'to like' and 'to stay'. Those occurring with the causee: 'to learn', 'to rear', 'to remove', 'to return', 'to wash', 'to teach' and 'to give'.

Tà kàrə tà is a complex of three morphemes which combines purposive, causative, resultative, and complementizer. It occurs between complex clauses like: INT+TRA, TRA+INT, INT+INT and TRA+TRA and expresses the meaning of 'to make', 'to have to' and 'to cause'. The intransitive verbs found before this construction are: 'to look for', 'to talk', 'to like', 'to want' and 'to decide'; and one transitive verb: 'to leave'/'to abandon'. The verbs following the causative construction are transitive verbs like: 'to kill' and 'to guard' and intransitive verbs like: 'to arrive', 'to go' and 'to stay'.

As a discourse marker, *àrə* functions as a sequential semantic link between propositions, with causative and resultative implications and an element of subjunctivity. It is given the meaning of 'therefore', 'so', and 'so that'. It may be followed by a simple pronoun, a noun phrase, a transitive clause or an embedded relative clause. The following verbs are involved: 'to wear', 'to talk', 'to bite' and 'to walk'. *àrə* occurs at the beginning of an utterance or a clause and may link a series of clauses. As *àrə* never occurs with the initial /k-/ and never takes inflections, it could be considered a reduced phonological form of *kàrə*. A further phonological reduction takes place with the connective *à* 'then' and 'and', 'thus', which occurs frequently in all types of discourse, joining clauses, phrases and nouns.

Thus, the verb 'to give' serves a multitude of purposes. In its grammaticalized forms it appears to be a syntactically stable but versatile morpheme. An increase in abstractness and a change in meaning takes place. Although grammaticalization occurs in most types of discourse, it is not present in any of the 375 proverbs collected. Nor is it used for insults or ideophones. Two chains of grammaticalization have been identified: a dative (receptive-benefactive) and a causative-purposive-resultative-complementiser-sequential-connective track. Unidirectionality may be implied. Phonological reduction occurs with discourse markers and a decreasing usage of

pronominal affixes. Carefully constructed cross-sociolinguistic or longitudinal studies may reveal more about the diachronic development of grammaticalization in Kabba.

Grammaticalization chains have been observed and attested cross-linguistically (e.g. Heine et al. 1991a; Craig 1991; Aikhenvald 2000 and Chappell 2001b) with idiosyncratic variations. I hope that this contribution from the Kabba language will further contribute to the understanding and appreciation of this interesting phenomenon of grammaticalization.

Abbreviations

ANA	Anaphora	INS	Instrumental
BEN	Benefactive	INT	Intensifier
CAU	Causative	LOC	Locative
COM	Complementizer	NEG	Negation
CON	Connective	P(L)	Plural
DEM	Demonstrative	PRO	Pronoun
DET	Determiner	PROG	Progressive
DM	Development Marker	PUR	Purposive
EM	Emphatic	QUE	Question
FOC	Focus	REA	Reason
FUT	Future	REL	Relative Pronoun
GEN	Genitive	RES	Resultative
HON	Honorific	SEQ	Sequential
INT	Intransitive	TRA	Transitive
IMP	Imperfective	VA	Verbal Adjunct
		VOC	Vocative

Notes

1. This article forms part of a reference grammar on Kabba in progress for a doctoral dissertation. For valuable comments I am indebted to Sasha Aikhenvald, Hilary Chappell, Talmy Givon and Regina Pustet. My gratitude also goes to my Kabba teacher and mentor Jean-Pierre Dingatoloum.

2. Hopper and Traugott (1993:6) talk about 'clines' involving a gradual shift or a series of transitions from one category to another. Bisang (1998:16) points out that 'grammaticalization starts from a semantic change of a sign and the cognitive strategies by which it is caused'. This strategy leads to language change, which, to be successful, must undergo a series of sociolinguistic processes. According to Heine, Claudi and Hünemeyer (1991a:164) grammaticalization consists of a discrete component which is 'metaphoric in nature and largely free from discourse pragmatic constraints, and of a continuous component, which 'appears to be metonymic and depends strongly on the linguistic and extra-linguistic context'.

3. Greenberg (1991) observed that in Nilo-Saharan an initial /k-/ was randomly found “within and across related languages” (309), and that it was a noun marker which has lost its function as a definite and indefinite articles. It originally had a demonstrative meaning and in “some languages was prefixed to the verb to indicate third person.” (313). This does not apply to Kabba, although its lexicon contains 25% of words starting with /k-/. The question we are faced with is: Did Kabba originally have verb classes and gradually lose them, or did it acquire verb class elements from surrounding languages? It is beyond the scope of this paper to answer this question, but it is interesting to speculate, especially considering the fact that this initial /k-/ has introduced a lot of complications into the Kabba language as far as its phonology and morphology is concerned.

References

- Aikhenvald, A. Y. (2000). Areal typology and grammaticalization. In S. Gildea (Ed.), *Reconstructing Grammar* (pp. 1–37). Amsterdam: John Benjamins.
- Bisang, W. (1998). Grammaticalization and language contact, construction and positions. In R. A. Giacalone & P. J. Hopper (Eds.), *The Limits of Grammaticalization* (pp. 13–58). Amsterdam: John Benjamins.
- Chappell, H. (2001a). Dialect grammar in two early modern Southern Min texts: A comparative study of dative, comitative and diminutive. *Journal of Chinese Linguistics*, 28 (2), 247–302.
- . (2001b). The grammaticalization of the verb kóng ‘say’ in Taiwanese Southern Min: A case of reanalysis. Unpublished paper. Melbourne: La Trobe University.
- Comrie, B. (1976). The grammar of causative constructions. In M. Shibatani. (Ed.), *Syntax and Semantics*, Vol. 6: *The grammar of causative constructions* (pp. 261–312). New York, NY: Academic Press.
- . 1981. *Language Universals and Linguistic Typology*. Oxford: Blackwell.
- Craig C. G. (1991). Ways to go in Rama: A case study in polygrammaticalization. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*, Vol. 2: *Focus on types of grammatical markers* (pp. 455–491). Amsterdam: John Benjamins.
- Dixon, R. M. W. (1982). *Where Have all the Adjectives Gone?* Berlin: Mouton de Gruyter.
- & Aikhenvald, A. Y. (2000a.) Introduction. In R. M. W. Dixon & A. Aikhenvald (Eds.), *Changing Valency Case Studies in Transitivity* (pp. 1–29). Cambridge: Cambridge University Press.
- . (2000b). A typology of causatives: Form, syntax and meaning. In R. M. W. Dixon & A. Y. Aikhenvald (Eds.), *Changing Valency Case Studies in Transitivity*. Cambridge: Cambridge University Press.
- Eilfort, W. H., Kroeber, P. D., & Peterseon, K. L. (Eds.). (1985). *Papers from the Parasession on Causatives and Agentivity*. Chicago, IL : Chicago Linguistic Society [CLS 21: Part 2].
- Giacalone, R. A., & Hopper, P. J. (Eds.). (1998). *The Limits of Grammaticalization*. Amsterdam: John Benjamins.
- Greenberg, J. H. (1991). The last stages of grammatical elements: Contractive and expansive desemanticization. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*, Vol. 1: *Focus on theoretical and methodological issues* (pp. 301–314). Amsterdam: John Benjamins.
- Heine, B., Claudi, U., & Hünnemeyer, F. (1991a). *Grammaticalization: A conceptual framework*. Chicago, IL: University of Chicago Press.

- . (1991b). From cognition to grammar. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*, Vol. 1: *Focus on theoretical and methodological issues* (pp. 149–187). Amsterdam: John Benjamins.
- & M. Reh (1984). *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske Verlag.
- et al. (1993). *Conceptual Shift. A lexicon of grammaticalization processes in African languages*. Cologne: University of Cologne [Afrikanische Arbeitspapiere 34/35].
- Hopper, P. (1991). On some principles of grammaticalization. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*, Vol. 1: *Focus on theoretical and methodological issues* (pp. 19–35). Amsterdam: John Benjamins.
- & E. C. Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Newman, J. (1996). *Give: A cognitive linguistic study*. Berlin: Mouton de Gruyter.
- Shibatani, M. (Ed.). (1976). *Syntax and Semantics*, Vol. 6: *The grammar of causative constructions*. New York, NY: Academic Press.
- Talmy, L. (1976). Semantic types. In M. Shibatani (Ed.), *Syntax and Semantics*, Vol. 6: *The grammar of causative constructions* (pp. 43–116). New York NY: Academic Press.
- Traugott, E. C., & König, E. (1991b). The semantics-pragmatics of grammaticalization revisited. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*, Vol. 1: *Focus on theoretical and methodological issues* (pp. 189–218). Amsterdam: John Benjamins.
- Yap, Foong-Ha (2000). The Polysemy of 'Give' Constructions in Malay and Other Languages: A grammaticalization perspective, PhD Dissertation, University of California, Los Angeles.

Selectors in Cushitic*

Maarten Mous

Leiden University

Cushitic languages are verb final, but many among them have an additional inflectional element in the sentence that is separate from the verb for which I use the term selector here. The purpose of the article is to provide an overview of these elements, of their functions and of which categories are expressed on them. What most of these so-called selectors have in common is the marking of sentence type and/or focus, and of subject. The comparison results in three types of selectors: (i) those that define the left border of a syntactic unit such as the verbal piece in Somali; (ii) those that indicate focus as a pro-clitic to the verb, and (iii) those that indicate focus by their position in the sentence.

1. Introduction

Cushitic languages are verb final, but many among them have an additional inflectional element in sentence that is separate from the verb and which has been termed in various ways: selector in Southern Cushitic languages, indicator particle in Somali, focus marker in Oromo. The purpose of this article is to provide an overview of these elements, of their functions and of the categories that are expressed on them. For this overview, I take into account anything that resembles the Southern Cushitic selector in some respect. What most of these so-called selectors have in common is the marking of sentence type and/or focus, and of subject. The comparison is typological and no attempt is made at reconstruction. In fact the comparison results in three types of selectors: (i) those that define the left boundary of a syntactic unit such as the verbal piece in Somali; (ii) those that indicate focus as a pro-clitic to the verb, and (iii) those that indicate focus by their position in the sentence.

We will first present an overview of selectors by giving examples from each of the Cushitic languages that have such elements. This list of languages and an overview of the categories that are expressed in and on selectors are given in Table 1. In this table the column *Sub* indicates whether the subject is indicated in the selector, *SType* whether sentence type is indicated, *Focus* whether the selector has focus meaning, *ImpS* whether the selector may contain an impersonal sub-

Table 1. Overview of properties of selectors

	SEN TYPE	MOOD	FOC	SUB	IPS SUB	OBJECT	OBJ PRO	CASE	DEIXIS	TENSE/ ASPECT
Arbore	y	y	n	y	y	y	n	y	y	y
Dhaasanac	y	y	n	y	n	y	y	y	n	n
Elmolo	n	—	n	y	y	n?	y	y	y	y
Dullay	n	n	y	y	n	y	y	y	n	n
Konso	y	y	y	y	n	y	y	n	n	n
Oromo	n	y	y	(y)	n	n	n	n	y	n
Boni	n	y	y	n	y	n	y	y	y	n
Rendille	n	n	y	n	y	n	y	y	y	n
Somali	y	y	y	y	y	y	y	y	y	n
Dahalo	n	y	n	n	n	y	n	y	y	y
Alagwa	y	y	y	y	y	y	y	y	y	y
Burunge	y	y	y	y	y	y	y	y	y	y
Iraqw	y	y	n	y	y	y	y	y	y	y

ject, *Object* whether the object can separate selector and verb, *Opro* whether the language has an object pronoun series different from the independent pronouns, *Case* whether adverbial “case” markers occur on the selector; *Deixis* indicates direction marking and *Tense/Aspect* whether tense/aspect is expressed on the selector (in addition to the verb). These categories are discussed one by one in Sections 2–4 followed by some conclusions in Section 5.

The Cushitic languages that have selectors are, in alphabetical order, the following: Alagwa, Arbore, Boni, Burunge, Dahalo, Dhaasanac, Dullay, Elmolo, Gidole (Dirayta), Iraqw, Konso, Oromo, Rendille, Somali. I am excluding Gidole (Dirayta), Gorwa, and Yaaku from the overview in this article because I do not have sufficient data for these languages.

The languages that have no selector are the following (in alphabetical order): Afar, Agaw, Baiso, Beja, Burji, Haddiyya, Kambaata, Sidamo.

In the following resumé of the subclassification of Cushitic we argue that those languages that do not have a selector fall outside the core group of Cushitic languages, i.e., the Southern Lowland East Cushitic branch. The subclassification of the Cushitic languages in Figure 1 is taken from Tosco (2000). This subclassification takes Southern Cushitic as part of Eastern Cushitic which is also proposed by Hetzron (1980), Ehret (1995), and Kiessling & Mous (2003). With Southern Cushitic part of Eastern Cushitic the majority of Eastern Cushitic languages¹ have selectors of one kind or another and those outside of it, Beja and Agaw, do not. If one follows suggestions that Dahalo and the Iraqw group of South Cushitic fit lower in the tree, roughly in Transversal-Southern-Lowland East Cushitic, as suggested in Tosco (2000) and Kiessling and Mous (2003), then the group of

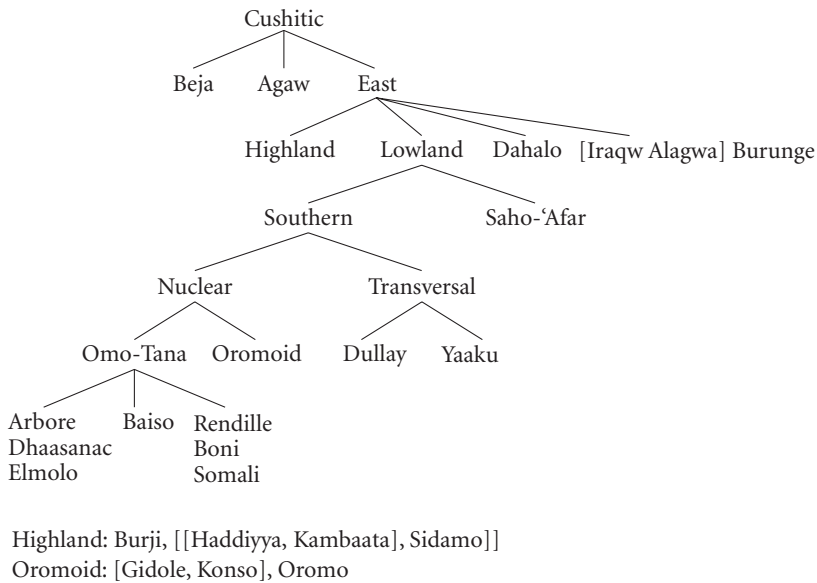


Figure 1. The Classification of the Cushitic languages

Southern-Lowland East Cushitic languages coincides with the languages that have selectors.

Relatively little attention has been paid to a comparison of the Southern and Eastern Cushitic selectors so far. Now that we have new data at our disposal on languages that were lesser known before and with the inclusion of Southern Cushitic in Eastern or Core Cushitic, such a comparison becomes possible and is needed for a more complete picture of the Cushitic verbal system and its syntax.

In Arbore the selector indicates both sentence type, which is (definite) indicative in (1), and subject, which is third person singular in (1).² Other sentence types that are indicated in the selector in Arbore are indefinite indicative future, indefinite indicative present, jussive and negative.³ The subject marking is either suffixed or prefixed to these sentence-identifying selectors. Not every sentence, however, has a sentence identifier (Hayward 1984).

- (1) *mo ʔi-y kor kúure* (Arbore, Hayward 1984:110)⁴
 man DEF.IND-3s tree cut:3SG.M:PERF
 'The man cut the tree.'

In Dhaasanac too there are these two categories of sentence type (indicative versus non-indicative) and of subject. Interestingly, the two elements are sometimes separated as is the case in (2) where the verbal pronoun, here the third person marker, and the sentence marker, here the sentence-initial focus particle, indicate subject and sentence type. The occurrence of verbal pronouns is linked to the presence

of a sentence type marker (Tosco 2001). Sentence type marking is also present in Alagwa, Burunge, Iraqw which have different sets of selectors according to sentence type. Indication of subject in the selector is very common in the Cushitic languages, see Table 1.

- (2) *h_a mí h_i ʔeggeðe* (Dhaasanac, Tosco 2001:260)
 FOC man 3.VERB take.RED.IMPF.A
 'He'll take it.'

In Elmolo the person aspect prefixes to the verb indicate in separate prefixes: aspect, subject, object, and a case clitic *ka*, as shown in (3), which indicates that somewhere in the sentence there is a beneficiary complement, here *hele* 'children'. Heine (1980) calls this case clitic aptly an applicative since this marker is cliticized to the verb rather than to the nominal phrase it refers to. In Somali studies such a clitic is called a case clitic or an adpositional clitic. A similar clitic is the *s-* in Alagwa in example (4). Apart from beneficiary objects such clitics can indicate other (adverbial) case relations. For example, the clitic *ri* in (5), Burunge, licenses a comitative complement in the sentence. Indication of direction is indicated in *ka* 'from' and *soó* 'to' in example (6) from Rendille.

- (3) *yesé mín hele aṇ-ká-dis-a* (Elmolo, Heine 1980:197)
 I house children s1SG-APPL-build-IMPFV
 'I build a house for my children.'
- (4) *kúu lo-s-o hhab-it Juma* (Alagwa, Mous in prep.)
 2SG.M OPT-BEN-O.M tell-2SG Juma
 'You should tell Juma'
- (5) *'ana fu'umay-hhank-ⁱ ha-gi-ni-ri fa/a /agim^a*
 1SG Fleish-N-DEM1 s1/2-O3PL-O.FOC-COM Brei essen.1SG.IPF
'ilibaa-goo-ba (Burunge, Kiessling 1994:163)
 Milch-PRED-NEG
 'Ich esse den Brei zusammen mit *diesem Fleisch*, nicht etwa mit der Milch.'
- (6) *á-í-ká-soó-weyne* (Rendille, Pillinger & Galboran 2000:30)
 FOC-IO-from-to-drove.animals:we
 'We drove the animals from [there] to [here] for him.'

Example (7) from Dullay shows a case in which an object noun separates the selector from the verb. Sasse (1984) has discussed several of such cases of incorporation of the object "into" the verb. In Iraqw there is a distinction between object placement after the selector for out of focus and noun incorporation proper, see Kooij & Mous (2002). In a number of languages full nouns can occur between the selector and the verb while at the same time there are syntactic restrictions to this position of the (object) noun. In these languages the selector functions as the left hand edge of a syntactic unit. In Somali this unit is termed the verbal piece (Saeed 1999). The

same applies to Iraqw where this unit has been termed core of the sentence (Mous 1993). Somali is the only language that allows *subjects* in this position, see *wiilkii* 'the boy' in (8).

- (7) *šampo-nu kán-talte=teeḥi* (Dullay, Amborn et al. 1980:84)
 boy-BEN SEL-goat=give:PST
 'Ich gab dem Jungen eine Ziege.'
- (8) *moos baa/b-uu wiilkii cunayaa* (Somali, Svolacchia et al. 1995:68)
 banana FM/FM-S.PRO boy:the eating
 'The boy is eating A BANANA.'

One of the prime functions of selectors is to order the sentence in terms of information structure. The languages in which the selector marks the left hand edge of a syntactic unit only allow out of focus elements to appear to the right of the selector, or inside the verbal piece. In other languages the position of the focus marker itself indicates the scope of focus. Thus in example (9) from Konso, *i* indicates that the verb receives selective focus; the sentence without *i* would have been neutral, 'she cleaned the house' and with *i* cliticized to the preceding noun, 'the house', would receive selective focus. In Oromo, in (10), *ni* is a focus marker, as is *á* in Rendille in (11) and *baa* in Somali in (12).

- (9) *tika (kara) i saha-t-i* (Konso, Mous 2001)
 house (inside) s3 clean-F-PF
 'Among the various actions she did, one was cleaning house.'
- (10) *ní-n-gab-bu'a ~ gan-ní-n-bu'a* (Harar Oromo, Owens 1985:60)
 FC-I-down-descend
 'I am going down.'
- (11) *inam yábar á-khaate* (Rendille, Pillinger & Galboran 2000:23)
 the.boy the.rope FOC-took
 'The boy took the rope.'
- (12) *moos baa/b-uu wiilkii cunayaa* (Somali, Svolacchia et al. 1995:68)
 banana FM/FM-S.PRO boy:the eating
 'The boy is eating A BANANA.'

In the Boni example in (13) the verb focus marker *a* is followed by a reflexive object pronoun. In the Rendille example (11) above there is an indirect object pronoun which is from a different set than those of the direct object pronouns. Somali too has such a second set of (indirect) object pronouns.

- (13) *an wa-'i-kíd'ifa-'a* (Kilii-dialect of Boni, Heine 1982:68)
 1SG VF-REFL-beat-T/A
 'I beat myself.'

Selectors with all the inflectional categories that can be expressed on them can develop into quite extensive inflectional complexes, specifically in Dahalo, Alagwa, Burunge, Iraqw as can be seen for Dahalo in (14) and for Iraqw in (15).

- (14) *b'á-ka-vá-ji* *łággwa* (Dahalo, Tosco 1991:71)
 NEG-IRR-PAST-HAB love-EL-3M
 'He didn't love him.'
- (15) *mu-s-tu-nd-a-y* *haníis* (Iraqw, Mous 1993:123)
 QUES-REAS-IMPS-O.2.PL-PERF-DIR give:PAST
 'Why were you (plural) favored?'

Some of the issues to be studied in more detail that arise from the preceding examples are the following:

- What kind of sentences require / allow a selector or does the selector include a sentence type marker?
- Is the selector a focus marker?
- Is subject marked in the selector?
- Is the selector the left hand boundary of a verbal complex?
- Can the object separate the selector from the verb?
- What other markers form a complex with the selector, such as impersonal subject pronoun, object pronoun, "case" particles, questioning or prohibitive mood, tense/ aspect?

2. Sentence type and focus

In some languages the choice of the selector is determined by the type of sentence and thus in these languages the selector itself is a sentence type marker; Hayward (1984) uses the term "sentence identifier". Sentence type (Sentype in Table 1) refers here to a distinction comparable to main/subordinate or independent/dependent clause. In Alagwa and Iraqw there are different sets of selectors for main clauses as opposed to consecutive clauses, and yet another set for object relative clauses. In example (16) from Alagwa the second sentence repeats the preceding sentence as a subordinate sentence to the next sentence; this next sentence contains new information in the story. This spaghetti style is a typical strategy for marking continuity in the storyline in Alagwa stories. Sentence (17) presents another strategy to mark such continuity, namely the use of a consecutive selector. For Iraqw, where the presence of a selector is obligatory in every sentence, it can be argued that the selector is an element which renders a string of words a sentence. It is natural for such a sentence-defining element to characterize the main syntactic distinctions among sentences, i.e. to mark sentence type.

- (16) *iyoo-r-oo'in iy-aa /alalaa tleehh-eehh-it.* (Alagwa, Mous to appear)
 mother-F-their 3-PST pots make-FRQ-3F
iyoo-r-oo'in k-y-aa /alalee tleehh-eehh-it, ...
 mother-F-their SUB-3-PST pots make-FRQ-3F
 'Her mother was making pots. Her mother being a pot maker, ...'
- (17) *maa /alu nongoo bats*
 and pot.M CSEC:S3:O.M put:3.M
 'and he removed the pot.'

In addition to the main distinctions in sentence type expressed in the choice of selector, some languages also have mood prefixes that mark negative and prohibitive sentences, as well as sentences questioning *what*. Other ways of expressing questions exist as well, e.g., in the form of question words; and for negation, most languages have additional negation marking on the verb. These prefixes – none of the languages has mood suffixes – are summarized in Table 3. The column *neg* gives the negation marker on the selector; *proh* the negative imperative marker and *ques* the recurrent *m(a)*- on the selector to make the sentence a (what-) question. Additional mood distinctions are made in Alagwa, Burunge and Iraqw. Some of these originate in grammaticalized adverbs.

Apart from jussive/optative marking, the typical sentence type marking in Table 2 above plays a crucial role in the marking of a specific element of information structure, namely that of marking continuity in the storyline in the form of a consecutive marking or of subordinate marking in repeated sequences. In fact, the marking of information structure is the central function of the selectors. In some languages the selector itself marks focus, most often verb focus or sentence focus. For example, in Dhasaanac topic and focus are the main organizational factors in

Table 2. Sentence type markers

	SEN'TYPE	distinctions
Arbore	Y	indicative (definite/indefinite)/jussive
Dhaasanac	Y ⁵	independent/dependent
Elmolo	N	
Dullay	N	
Konso	Y	independent/dependent/jussive/negative/prohibitive
Oromo	N	
Boni	N	
Rendille	N	
Somali	N	
Dahalo	N	
Alagwa	Y	main/consecutive
Burunge	Y	main/consecutive/optative
Iraqw	Y	main/subordinate (consecutive)

Table 3. Mood distinctions in selectors

	NEG	PROH	QUES	OTHER MOOD PFXS
Arbore	<i>ma</i>	<i>ma</i>	–	–
Dhaasanac	<i>ma</i>			
Elmolo				
Dullay	–	–	–	–
Konso			–	(cond)
Oromo	–	–	–	–
Boni	<i>hú, m</i>			
Rendille	–	–	–	–
Somali	<i>ma</i>	<i>ha</i>	<i>ma</i>	
Dahalo	<i>b’a</i>			
Alagwa	–	–	–	opt, subord, consec
Burunge	–	–	–	consec, cond,
Iraqw	–	<i>m-</i>	<i>m-</i>	cond, conces

syntax. Neutral sentences have subject case marking on the subject and no focus selector, (18a). The presence of a selector indicates verbal focus, i.e., the verbal focus marker ^h*a* in (18b); there is a possibility of an additional verbal subject pronoun as in (18b) or a full (subject) pronoun as in (18c). When the subject is topicalized a subject pronoun is used, as in (18d). Subject focus is expressed by the addition of a nominal focus marker cliticized to the subject NP, (18e), as shown in (Tosco 2001: 261–273). Several languages have constituent focus markers that are separate from the selectors.

- (18) a. *ʔar kufi*
 bull:s die:PF.A
 ‘The/A bull died’ (neutral)
- b. *ʔár ^ha ^hí díyyime*
 bull FOC 3.VERB make.noise:IMPF.A
 ‘the bull *is making noise*’ (verbal focus)
- c. *só ^ha yú muura*
 meat FOC I cut
 ‘I’ll cut the meat.’
- d. *ʔár ^hé kufi*
 bull 3s die:PF.A
 ‘The bull died’ (as answer to ‘What happened to the bull?’)
- e. *ʔár=ru kufi*
 bull=FOC die:PF.A
 ‘The bull died’ (subject focus)

In Konso the position of the selector is meaningful. A comparison of (19) and (20) shows that preposing the selector to the verb adds the meaning of selective focus to the action. Comparing (21a and b) shows that if the selector follows the object,

as in (21b) this selective focus is on the object. Thus the position of the selector is crucial in the interpretation of the domain of the focus. Apart from selective focus, the person index also indicates assertion, as shown in (22), since the addition of the selector before the verb adds the sense that the action has been witnessed by the speaker and that the speaker can assert its truth. In addition to the selector there is an NP focus marker *-n*, (23) and (24). Konso examples are from Mous (2001).

- (19) *tika (kara) saha-t-i*
house (inside) clean-F-PF
'She cleaned house.'
- (20) *tika i saha-t-i*
house 3 clean-F-PF
'Among the various actions, one was cleaning house.'
- (21) a. *in íshin tooy-é*
SEL.S1 2PL look-PF
'I looked at you.'
- b. *íshin in tooyé*
2PL SEL.S1 look-PF
'I looked at you while there were others.'
- (22) *isheetá qoyra-nne unt-i tum-t-i*
she stick-with millet-3 thresh-F-PF
'She threshed the millet with a stick.' (eyewitness account)
- (23) *parri-n olli-n anna*
tomorrow-FOC together-1 go:1PL:IMPFV
'Tomorrow we go together.'
- (24) *parri olli-n-in anna*
tomorrow together-FOC-1 go:1PL:IMPFV
'Tomorrow we go *together*.'

In Boni too the position of the selector (verbal focus marker) determines the scope of focus, compare the examples in (25) and (26).

- (25) *áddigée á-ji'iki áñali* Sasse (1981:280)
father:my FM-drunk honey.wine
'My father has drunk the honey-wine (while I have not)'
- (26) *šali á-jiidi weešee* Sasse (1981:280)
yday FM-came child.my
'As for *yesterday*, he *did* come, my child (although today he did not)'

Focus and topic are central to the syntax of Somali. Every declarative clause has obligatory focus. Focus is expressed by focus markers on the selectors (called indicators) with only one focus marker per clause; embedded clauses and clauses without tense (imperative, jussive, potential) have no focus marker. In the sen-

tences (27)–(30), all taken from Svolacchia et al. (1995), the focus markers are obligatory. The selector or focus marker (FM) marks the beginning of the verbal piece and arguments within the verbal piece are de-focused.

- (27) *naag baa libaax aratay*
 woman FM lion saw:she
 ‘A woman has seen a lion.’
- (28) *Cali moos w-uu cunay*
 Ali banana FM-he ate:he
 ‘Ali has eaten a banana.’
- (29) *Cali wax-uu cunay moos*
 Ali FM-he ate:he banana
 ‘Ali has eaten a banana.’
- (30) *moos baa/b-uu wiilkii cunayaa*
 banana FM/FM-S.PRO boy:the eating
 ‘The boy is eating a banana.’

The focus markers in selectors as well as the noun phrase focus markers are summarized in Table 4. The column *Focus* indicates whether the selector is a focus marker; *indirect* in this column means that the selector is not a focus marker but its presence or its position has relevance for focus, e.g., by indicating that the sentence is neutral or that certain phrases are background rather than focus. The

Table 4. Focus marking of the the selector

	FOCUS	VERB/SENTENCE FOCUS	NP FOCUS MARKERS	SUBJECT FOCUS
Arbore	indirect	selector	no AGR	no AGR, no SEL
Dhaasanac	indirect	^h a + verbal pronoun	- ^h é (topic)	-ú (contrast)
Elmolo	n			
Dullay	y	selector	NP- <i>kka</i> or <i>k</i> -SEL non-subj constituent, NP- <i>ta</i> or <i>t</i> -SEL ‘also’	no selector+V3m
Konso	y	no selector	position of SEL, - <i>n</i> EMPH	subject followed by SEL
Oromo	y 1SG.EMPH	<i>ní</i>	<i>llée</i> EMPH, - <i>nís</i> ‘also’, and others.	- <i>n</i>
Boni	y	selector	-é	=
Rendille	y	selector	-é	=
Somali	y	<i>waa</i>	<i>baa</i> , <i>ayaa</i> , <i>wáxaa</i>	=
Dahalo	n		’ <i>ini</i> , - <i>ni</i>	=
Alagwa	indirect	- <i>n</i>	–	selector <i>na</i>
Burunge	indirect	SEL-O-V	<i>ní</i> OBJ.FOC	selector <i>na</i>
Iraqw	indirect			

column *Verb/sentence focus* indicates roughly how neutral or verb/sentence focus is expressed; the column *NPfocus markers* gives the noun phrase focus markers for those languages where they are different from the selector. The *subject focus* column indicates how subject focus is manifested provided it is manifested differently from NP-focus (and if it is not, this is indicated by the = sign in the column). Additional remarks: The Oromo negative marker *(hi)n-* is mutually exclusive with VF *ní*; Arbore topicalized sentences have no sentence identifier in the selector.

3. Subject and object marking

Most selectors indicate the subject in one way or another, and most of them differentiate in the selector at least between speech act participants and third persons. Cushitic languages do not distinguish second singular and third feminine in the subject agreement on the verb; and many of them also do not distinguish between first singular and third masculine. The selectors resolve this homonymy of person marking on the verb. However, this cannot be taken to be the function of the selector since for many languages the homonymy is still present in many cases, among others in those sentences where no selector is present, or when the subject marking is neutralized or replaced by object marking. Whenever there is a choice of using a selector, this never seems to be determined by considerations of resolving the homonymy in subject marking on the verb.

The properties of subject marking in selectors for those languages that have subject marking in their selectors are summarized in Table 5. The column “Person” gives the distinctions that are made in the subject marking on the selector. Those languages that only distinguish between speech act participants and third persons are indicated by “P/3” in Table 5. The column “Position” indicates whether this subject marking is a suffix (clitic) to another element of the selector, such as a sentence type marker, a tense/aspect marker or a focus marker, or whether the

Table 5. Subject and Selector

	PERSON	POSITION
Arbore	1/2/3 sg/pl	sfx to sentence type marker
Dhaasanac	1/2/3 sg/pl	after focus marker
Elmolo	1/2/3 sg/pl	sfx to t/a marker
Dullay	1/2/3	sfx to focus marker
Konso	1/2/3 sg/pl	=
Somali	1/2/3 sg/pl f/m	sfx to focus marker
Alagwa	P/3	=
Burunge	P/3	=
Iraqw	P/3	=

subject marker is basically the selector (indicated by =). In none of the Cushitic languages is the subject marker prefixed to a selector-like element. For several languages one could argue that the selector complex including the subject marker is pro-cliticized to the main verb. For Alagwa, Burunge, Iraqw one could argue that their subject markers are suffixed to sentence type markers but in these languages the subject markers themselves are analyzed as stems and the sentence type markers as prefixes.

The noun subject, when explicit, always precedes the selector. Only in Somali is a full noun or pronoun subject possible after the subject marking, see (8). Independent full subject pronouns can, but need not, be used in addition to the selector. This is true for all the languages, even for those in which there is no or minimal subject marking in the selector, such as Oromo.⁶

The distinctions made in the subject marking are basically in person only, 1/2/3, and not in gender while independent pronouns and subject agreement on the verb distinguishes gender in the third and often also in the second person singular. The subject markers of Somali are deviant in this respect. Those of Arbore too make all the distinctions that the language recognizes except for the gender difference in the third person. For the first and second person singular in Dhaasanac the subject pronouns are used instead of the verbal pronouns (Tosco 2001:214) and the 1PL.EXCL is identical to the equivalent in the object pronoun series while 1PL.INCL is from the subject pronoun series. The first person and second person plural marking in Elmolo are taken over from the object pronoun paradigm. These forms from other paradigms are put in square brackets. If we leave out Arbore and Somali and these forms from other paradigms, the picture is quite uniform in displaying a distinction between 1/2/3 person whereby third person is indicated by *i*, and first and second person by *a* plus a consonant; this consonant is a nasal in the case of the first person, and a glottal stop (in Konso) for the second person.⁷ Alagwa, Burunge, Iraqw do not distinguish (anymore?) between first and second person. Table 6 gives the forms of subject marking in the relevant languages.

Table 6. Subject Pronouns in Selectors

	1SG	2SG	3SG/F	1PL/INCL	2PL	3PL
Arbore	<i>n</i>	<i>0</i>	<i>y</i>	<i>na</i>	<i>n</i>	<i>so</i>
Dhaasanac	[<i>yáa</i>]	[<i>kúo</i>]	^h <i>t</i>	[<i>ni</i>] / [(^h <i>é</i>) <i>ké</i>]	^h <i>t</i>	^h <i>t</i>
Elmolo	<i>aŋ</i>	<i>a</i>	<i>i</i>	[<i>ino</i>]	[<i>iŋ</i>]	<i>i</i>
Dullay	<i>an</i>	<i>aC</i>	<i>a, u / a, i</i>	<i>an</i>	<i>aC</i>	<i>i</i>
Konso	<i>in</i>	<i>i'</i>	<i>i</i>	<i>in</i>	<i>i'</i>	<i>i</i>
Somali	<i>-aan</i>	<i>-aad</i>	<i>-uu/-ay</i>	<i>-aan(nu)/-aynu</i>	<i>-aydin ~ -aad</i>	<i>-ay</i>
Alagwa	<i>a</i>	<i>a</i>	<i>i</i>	<i>a</i>	<i>a</i>	<i>i</i>
Burunge	<i>ha</i>	<i>ha</i>	<i>hi</i>	<i>ha</i>	<i>ha</i>	<i>hi</i>
Iraqw	<i>a</i>	<i>a</i>	<i>i</i>	<i>a</i>	<i>a</i>	<i>i</i>

Many of the Cushitic languages have a separate impersonal subject (IPS) pronoun that is used in passive-like sentences. In some of the languages this option is in addition to the possibility of a passive extension on the verb. The verb form used in the impersonal construction is always the third person singular masculine form (3M), except for Arbore where it is 3PL (Hayward 1984:305). One could also argue that the verb is simply not conjugated for person. For Iraqw, the impersonal subject marker can also be used to refer to a specific collective group, yet there is no plural marking on the verb. There are some indications that the impersonal subject marker is of a different order than the subject pronouns. One such indication is that its structural position is different from the subject pronouns in Somali, according to Svolacchia et al. (1995).⁸ The properties of the impersonal pronouns are summarized in Table 7. The column *ImpSubject* gives the form of the impersonal subject marker in the selector. The column *verb* gives the conjugation of the verb used with such an impersonal subject selector. The column *passive* indicates whether the language has a productive derivational passive in addition to the impersonal subject construction. The column *obj-pro* indicates whether the (foregrounded) patient is referred to by an object pronoun in the selector. These object (patient) pronouns – when present – cliticized to the impersonal marker. The forms of impersonal subject pronoun are very uniform across the Cushitic languages having an alveolar consonant followed by an *a*.

In Alagwa, Burunge, and Iraqw the same markers, Impersonal subject plus Object Pronoun, are used for predicative adjective constructions, (31).

- (31) *thuway ku hhéer* (Iraqw, Mous 1993:203)
 rain(m) o.3:IMPS:O.M insufficient:M
 ‘Rain is insufficient.’

As is evident from the remarks in the last column, the impersonal subject marker has diverging properties in several languages despite their uniformity in form. In Burunge the impersonal subject pronoun is so un-specific that no independent personal pronoun can be used in connection with it. In Iraqw the same imper-

Table 7. Impersonal (non-specific) subject pronoun

	IMP SUBJECT	VERB	PASSIVE	OBJ-PRO	REMARKS
Arbore	<i>na</i>	3PL	no	no	= 1PL, intransitive, middle
Elmolo	<i>(a)na</i>	3M	no	yes	intransitive prefix
Boni	<i>l(i)</i>	3M	no	yse	
Rendille	<i>la</i>	3M	yes	yes	
Somali	<i>la</i>	3M	yes	yes	
Alagwa	<i>ta/k-</i>	3M	no	yes	
Burunge	<i>da</i>	3M	no	yes	no indep pronoun possible
Iraqw	<i>ta</i>	3M	no	yes	collective, human

sonal subject is used to indicate collective agents and in Arbore the impersonal subject is identical in form to the first person plural marker. In Iraqw the impersonal subject marker can only refer to human agents, while in the closely related language Burunge it can be used even with weather verbs, i.e., verbs in which reference to anything remotely related to a controlling agent is absent. The Elmolo equivalent of the impersonal is termed “intransitive” prefix by Heine, since it suppresses the possibility to express two complements; it too can be used with agent-less transitive verbs, such as ‘to have diarrhea’ in (33). The Arbore impersonal subject construction is also used for middle situations,⁹ as is evidenced by example (32).

- (32) *ina kare* (Arbore, Hayward 1984:308)
 1PS shave
 ‘He shaved himself / he was shaved’
- (33) *kesé ené-ke-(e)ld-e* (Elmolo, Heine 1980:198)
 2SG 1PS-O.2SG-have.diarrhea-PF
 ‘You have diarrhea.’

Object marking in the selector is different from subject marking in several respects. First the marking of object is not always compulsory but is dependent on the position of the full noun object and it is also related to information structure. Second, the distinctions in object markers are not reduced as is the norm in subject markers. Those languages that allow an object noun after the selector also allow an object pronoun in that position.¹⁰ In addition, a language like Rendille that does not allow a noun object after the selector, does allow an object pronoun in that position. Oromo, Konso and Dhaasanac have object pronouns but not as enclitics to the selector. Those that have object pronouns in the selector complex (Elmolo, Dullay, Rendille, Somali, Dahalo, Alagwa, Burunge, Iraqw) have them as enclitics to the selector, often replacing the fronted full object noun, see (34), (35). For many languages the third person object pronoun is zero, or, the object pronouns exist only for speech act participants. It will be clear that in these languages the syntax of object pronouns is less relevant. Among the Somali dialects some have an overt third person object marker while the standard dialect does not. Elmolo, Boni, Rendille and Somali, all members of the Omo-Tana branch of East Cushitic, have a separate set of indirect object nouns. All the languages that have object pronouns also have a reflexive/reciprocal pronoun. Oromo is the only language to have a distinct reciprocal pronoun. These properties of object pronouns are summarized in Table 8. The column *Object* indicates whether the language allows objects between the selector and the verb. The column *Opro* indicates whether the language has object pronouns that are different from the independent pronouns. The column *Opro3=0* indicates whether the most common type of object, the third person object pronoun, is zero. The column *sel+Opro* indicates whether these object pro-

Table 8. Properties of object pronouns

	OBJECT	OPRO	OPRO3=0	SEL+OPRO	IOPRO	REFL/REC
Arbore	<i>y</i>	<i>no</i>	—	—	—	PRO- <i>tta wal-V</i>
Dhaasanac	<i>y</i>	<i>y</i>	<i>y</i> ¹¹	<i>no</i>	<i>no</i>	<i>y</i>
Elmolo	<i>y</i>	<i>y</i>	<i>y</i>	<i>y</i>	<i>y</i>	<i>y</i>
Dullay	<i>y</i>	<i>y</i>	<i>no</i>	<i>y</i>	(<i>y</i>)	?
Konso	<i>y</i>	<i>y</i>	<i>y</i>	<i>no</i>	—	<i>y</i>
Oromo	<i>no</i>	<i>y</i>	<i>no</i>	<i>no</i>	—	<i>y</i>
Boni	<i>no</i>	<i>y</i>	<i>y</i>	?	<i>y</i>	<i>y</i>
Rendille	<i>no</i>	<i>y</i>	<i>y</i>	<i>y</i>	<i>y</i>	<i>y</i>
Somali	<i>y</i>	<i>y</i>	<i>y</i>	<i>if</i> O<SEL	<i>y</i>	<i>y</i>
Dahalo	<i>y</i>	<i>y</i>	<i>no</i>	<i>y</i>	<i>no</i>	?
Alagwa	<i>y</i>	<i>y</i>	<i>no</i>	<i>y</i>	<i>no</i>	<i>y</i>
Burunge	<i>y</i>	<i>y</i>	<i>no</i>	<i>if</i> O<SEL	<i>no</i>	<i>y</i>
Iraqw ¹²	<i>y</i>	<i>y</i>	<i>no</i>	<i>if</i> O<SEL	<i>no</i>	<i>y</i>

nouns form a complex with the selector. The column *IOpro* indicates whether there is a second series of (indirect) object pronouns in the selector. The column *refl/rec* indicates whether the language has a reflexive/reciprocal object pronoun.

- (34) *mólú kesé e-ké-arg-a* Elmolo (Heine1980:187)
 he you SEL-you-see-IMP
 ‘He sees you.’

- (35) *yesé núúm anán-ka héle koon-e* Elmolo (Heine 1980:196)
 I food I-PERF-APPL children brought
 ‘I brought the children food.’

The Arbore reflexive is not a separate object pronoun but consists of the object pronoun followed by *tta* and *wal-* prefixed to the verb, (36). In Rendille the reflexive precedes the impersonal clitic. The reflexive/reciprocal pronouns have either a root similar to the Oromo distinct reciprocal root *wali* as is the case in Dhaasanac, Elmolo, and Arbore, or to the Oromo reflexive *ifi* as is the case in Dullay, Konso, Boni, Rendille, and Somali, or one that is related to the 1PL pronoun as is the case in Alagwa, Burunge, Iraqw.

- (36) *yé-tta ([éhe]) wal-síibe* Arbore (Hayward 1984:227)
 I-SFX léhe¹³ REFL-wash
 ‘It was I that answered myself.’

The forms of the pronouns are presented in Table 9–11. These tables build heavily on earlier comparative work and they have been adapted and enlarged from Appleyard (1990) and Biber (1984:53–54). The forms of the object pronouns in Somali and closely related languages are presented in Table 10, and those of Alagwa, Burunge, Iraqw in Table 11. While those of the remaining languages are in Table 9.

Table 9. Object pronouns

	Dhaasanac	Elmolo	idem IO	Dullay	Konso	Oromo
1SG	<i>ye</i>	<i>é</i>	<i>é</i>	<i>ye</i>	<i>in</i>	<i>na</i>
2SG	<i>ko</i>	<i>ké</i>	<i>éké</i>	<i>ho,he,ke</i>	<i>ce</i>	<i>sí</i>
3SG.M	<i>0</i>	<i>0</i>	<i>í</i>	<i>na</i> ¹⁴	<i>0</i>	<i>isá</i>
3SG.F	<i>0</i>	<i>0</i>	<i>í</i>	<i>=</i>	<i>0</i>	<i>istí</i>
1PL	<i>ʃí</i>	<i>ino</i>	<i>ínó</i>	<i>in-na, ino,ine</i>	<i>inu</i>	<i>nu</i>
2PL	<i>ʔitíní</i> ¹⁵	<i>íŋ</i>	<i>íŋ</i>	<i>hun-na, huno,hune</i>	<i>ishin</i>	<i>isini</i>
3PL	<i>0</i>	<i>0</i>	<i>î</i>	<i>na</i>	<i>0</i>	<i>ísaaní</i>
REC/REFL	<i>^húol</i>	<i>wal</i>	<i>–</i>	<i>hi</i>	<i>issi</i>	<i>ífi</i>

Table 10. Object pronouns in the SAM languages

	Boni ¹⁶	Rendille	idem IO	Somali	idem IO	Ashraaf Somali	Karre Somali	Tunni Somali
1SG	<i>i</i>	<i>i</i>	<i>í</i>	<i>i</i>	<i>káy</i>	<i>ii</i>	<i>i</i>	<i>i</i>
2SG	<i>ku</i>	<i>ki</i>	<i>kí</i>	<i>ku</i>	<i>káa</i>	<i>ku(u)</i>	<i>ku</i>	<i>ki</i>
3SG.M	<i>0</i>	<i>0</i>	<i>í</i>	<i>0</i>	<i>0</i>	<i>su</i>	<i>0</i>	<i>0</i>
3SG.F	<i>0</i>	<i>0</i>	<i>í</i>	<i>0</i>	<i>0</i>	<i>sa</i>	<i>0</i>	<i>0</i>
1PL.EXCL	<i>nu</i>	<i>nah</i>	<i>í</i>	<i>na</i>	<i>kayó</i>	<i>noo</i>	<i>nu</i>	<i>nu</i>
1PL.INCL		<i>inno</i>	<i>í</i>	<i>ina</i>	<i>kéen</i>			
2PL	<i>da</i>	<i>atin</i>	<i>í</i>	<i>idín</i>	<i>ktín</i>	<i>?</i>	<i>ad</i>	<i>iski</i>
3PL	<i>0</i>	<i>0</i>	<i>í</i>	<i>0</i>	<i>0</i>	<i>so(o)</i>	<i>0</i>	<i>0</i>
REC/REFL	<i>í</i>	<i>is</i>	<i>–</i>	<i>is</i>	<i>–</i>	<i>?</i>	<i>?</i>	<i>?</i>

Table 11. Object pronouns in the Southern Cushitic languages

	Dahalo	Iraqw	Alagwa	Burunge
1SG	<i>ʔi</i>	<i>i</i>	<i>i</i>	<i>ní</i>
2SG.M	<i>ku</i>	<i>u</i>	<i>oo</i>	<i>gu</i>
2.SG.F	<i>ki</i>	<i>i</i>	<i>i</i>	<i>gi</i>
3SG.M	<i>du</i>	<i>u</i>	<i>ii</i>	<i>gu</i>
3SG.F	<i>dí</i>	<i>a</i>	<i>a</i>	<i>ga</i>
REC/REFL	<i>?</i>	<i>tí</i>	<i>kunu</i>	<i>tí ~ ngí</i>
1PL	<i>ní</i>	<i>tí</i>	<i>kunu</i>	<i>ndí</i>
2PL	<i>kunná</i>	<i>nu</i>	<i>kunu</i>	<i>ngu</i>
2PL.F	<i>kinná</i>	<i>=</i>	<i>=</i>	<i>=</i>
3PL	<i>ʔiŋa</i>	<i>i</i>	<i>i</i>	<i>gi</i>

This division over three tables also allows for an economic display of the differences of distinctions in object pronouns. The languages in Table 10 (Somali and closest relatives) distinguish between inclusive and exclusive first person plural pronouns which the other languages do not. The Southern Cushitic languages in Table 11 are unique in that they distinguish between masculine and feminine

in the second person, as well as in the third person. Gender distinction in third person object pronouns is otherwise limited to Oromo and Ashraaf Somali. In general object pronouns expressed in or on the selectors tend to make the gender distinctions that are made in the language, while the subject marking in selectors is reduced to person marking and typically void of gender specification.

4. “Case” clitics, direction marking and tense/aspect marking

Many of the languages with selectors also have clitics that indicate a semantic function of one of the non-core argument in the sentence. Various terms have been proposed for these clitics: case or adpositional clitics, but also applicative; such clitics appear in the earlier examples (4), (5), (6), and (15). These clitics precede the verb and follow the selector. The semantic functions that they indicate range from beneficent or dative, allative and ablative, instrumental, and comitative to locative. Sometimes these clitics cliticize to the right (and then usually to the verb); in other languages, or in the same language as another option, to the left and then usually cliticize to an internal noun or pronoun argument. In such cases this noun or pronoun need not be the referent of the semantic function that is expressed in the clitic, in other words, the clitics may appear on the “wrong” referent (anti-iconicity) as in (37).

- (37) *'um ye kí šuún* (Dhaasanac, Tosco 2001)
 children me with gather.IMPF.A
 ‘Bring me the children.’

Table 12 summarizes the properties of these “case” clitics. The column *Sel-Case-Verb* indicates whether the language has the possibility of such a marker between the selector (including object pronouns). The column *stacking* indicates whether several such markers may be combined in that position as in (15). A hyphen indicates that the question is not relevant and a question mark that the question could not be answered yet. The column *pre-/postpos* indicates whether there are pre- or postpositions to an NP in the language that are related to these case clitics. The column *anti-iconic clitics* indicates whether these case clitics may end up on or next to the “wrong” NP or pronoun, as in (37). The column *clitic to pro* indicates whether these case clitics cliticize to an object pronoun. And the last column gives the direction of cliticization, to the right or to the left.

The actual forms of the case clitics are given in Table 13 and Table 14. These tables are based on Appleyard (1990:27) and Biber (1984:51–52).

In addition to allative and ablative “case” clitics, there are also hither (towards the deictic center) / thither (away from the deictic center) markers in this position in a number of languages and these too can become part of the selector complex.

Table 12. Syntax of case clitics

	SEL-CASE- VERB	STACKING	PRE-/POST- POSITION	ANTI-ICONIC CLITICS	CLITIC TO PRO	DIRECTION OF CLITICIZATION
Arbore	y	n	different	y	n	>
Dhaasanac	y	n	post	y	n	<
Elmolo	y	n	–	?	y	>
Dullay	y	?	n		n	>
Konso	n	n	post	–	–	<
Oromo	n	–	post	–	–	<
Boni	y	?	–			
Rendille	y	y	post	?	?	>
Somali	y	y	–	y	y	<
Dahalo	y	?	post			
Alagwa	y (BEN pfx)	n	diff	n	y	>
Burunge	y	n	prep	y	y	<
Iraqw	y	n	pre-&post	y	only iconic	<

Table 13. Case markers

	Arbore	Dhaasanac	Elmolo	Dullay	Oromo	Boni	Rendille
DATIVE	<i>ka</i>	<i>kí</i>	<i>ka</i>		(<i>qaa</i>) <i>f</i>	<i>’ú</i>	<i>í</i>
ALLATIVE		<i>ká</i>		<i>hilá</i>	(<i>i</i>) <i>tti</i>	<i>’ú</i>	
LOCATIVE	<i>ka</i>	<i>gaa</i>		<i>lá</i>	(<i>i</i>) <i>rraa</i>	<i>kí</i>	<i>ká</i>
INSTRUMENTAL	<i>karɿ</i>	<i>kí</i>	<i>ka</i>	(<i>k</i>) <i>ká</i> ¹⁷	(<i>qaa</i>) <i>n</i>	<i>kí</i>	<i>ká</i>
ABLATIVE	<i>ka</i>	<i>gaa</i>	<i>ka</i>		(<i>qaa</i>)	<i>kə</i>	<i>ká</i>
COMITATIVE	–	<i>kí</i>		<i>illé</i>		<i>lí~ ní</i>	<i>leé</i>

Table 14. Case markers second half

	Somali	Dahalo	Alagwa	Burunge	Iraqw
DATIVE	<i>ú</i>		<i>s-</i>	<i>sa-</i>	<i>as</i>
ALLATIVE	<i>ú</i>				<i>i</i>
LOCATIVE	<i>kú</i>	<i>da</i>			<i>wa</i>
INSTRUMENTAL	<i>kú</i>	<i>’í</i>		<i>-ri</i>	<i>ar</i>
ABLATIVE	<i>ká</i>		<i>-aa</i>		<i>wa</i>
COMITATIVE	<i>lá</i>	(<i>’in</i>) <i>nto</i>		<i>-ri</i>	(<i>al</i>)

The fine semantic details of the hither/thither marking in Somali are discussed in Bourdin (this volume). The *hith/thith* column in Table 15 gives the form for indication of direction for those languages that have such markers. Those languages with only one form in the column and no form after the slash have only a hither marker. In Oromo the markers indicate ‘down’ versus ‘up’ rather than hither/thither, see Owens (1985:60).

Table 15. Direction marking

	HITHER/THITHER
Arbore	<i>ar / 'ug</i>
Dhaasanac	
Elmolo	<i>ar /</i>
Dullay	—
Konso	—
Oromo	<i>gadi / oli</i>
Boni	<i>háa /</i>
Rendille	<i>soó /</i>
Somali	<i>soó / sii</i>
Dahalo	<i>'a / ji</i>
Alagwa	<i>n- /</i>
Burunge	<i>ni / ti</i>
Iraqw	<i>ni /</i>

Table 16. Tense/Aspect marking

	TENSE/ASPECT
Arbore	SEL
Dhaasanac	<i>n</i>
Elmolo	SEL
Dullay	<i>n</i>
Konso	<i>n</i>
Oromo	<i>n</i>
Boni	<i>n</i>
Rendille	<i>n</i>
Somali	<i>n</i>
Dahalo	<i>y</i>
Alagwa	<i>y</i>
Burunge	<i>y</i>
Iraqw	<i>y</i>

Several languages also indicate tense/aspect on the selector, either by the choice of the selector (SEL) or by an affix to the selector. This tense/aspect marking is in addition to tense/aspect marking on the verb. The tense/aspect marking on selectors is summarized in Table 16.

5. Summary and wider perspective

The prime function of selectors is to express elements of information structure. In all languages the selector has either a function as sentence type marker and these sentence types are at least partly related to information structure (backgrounding),

or the selector expresses focus in one way or another. The only exceptions are Elmolo and Dahalo. Elmolo does have prohibitive mood marked in the selector but does not seem to mark sentence type or focus in the selector. It is however possible that this exceptional status is due to the fact that the language was already in decay when it was described.¹⁸ Nearly all languages also mark subject in the selector with the exceptions of Boni and Rendille. These languages use independent subject pronouns where others use subject suffixes or clitics. Rendille and Boni do however have an impersonal subject marker that is integrated into the selector. They also have object pronoun clitics that are part of the selector. Inflectional subject marking is typically a characteristic of the verb. Thus we can say that selectors take up part of typically verbal functions. In this sense the sentence defining properties are divided over verb and selector. While in some languages the position of the selector determines the scope of focus, in a number of other languages the position of the selector is more fixed and the position of the object vis-à-vis the selector determines the information value of the object. Once the selector has a fixed position then it also has a stronger syntactic function in the semantics of placement of complements in relation to the selector; thus, in addition to subject marking these selectors also have the verbal quality of valency. By developing more verbal functions the selector attracts other inflectional marking such as tense/aspect marking. There seems to emerge a division of marking of grammatical roles in several of these languages: subject on verb, object in selector, and others in the “case” clitic which has a fixed position between selector and verb.

If the selector defines the beginning of a syntactic unit (verbal piece) then the function of placing the object within this unit is either neutral or backgrounding. Despite the fact that Cushitic languages are aptly characterized as verb final, some languages do allow for the placement of a complement after the verb. Somali allows for a complement to be after the verb for contrast;¹⁹ Dahalo has a position after the verb in particular for pronouns and for complex sentences; Burunge has it for object focus and for backgrounded subject; and Iraqw has it as a very marked option for objects in separate sentences that form a list that leads to a climax. Some languages allow for a split object: head noun preceding the selector and modifying noun after the selector, or even after the verb, specifically for external possession as in (38) or one in which the argument of the verbal noun has syntactic status in the main clause, (39).

- (38) *s-i laqaru kaw-t-ee iyoo-roo-kin* (Alagwa, Mous in prep.)
 BEN-O1SG showing go-2-OPT mother-F-POSS2PL
 ‘Go and show me your mother.’
- (39) *maamay-wóo-ín nongo quutl diintsa* (Alagwa, Mous in prep.)
 uncle-M-POSS3PL CSEC:S2/3:O.M cut:3M finger
 ‘He cut his uncle’s finger.’

Our knowledge of the syntax of several of the Cushitic languages that were surveyed in this article is insufficient to study subtle variation in object properties. Further study in this area could prove very fruitful for a better understanding of the correlation of the various syntactic factors and particularly within groups of closely related languages such the Iraqw-Alagwa-Burunge cluster, the Konso cluster, the Dullay languages, or the Somali dialects.

On the African continent the group of languages that is best known for a syntax in which the verbal functions are divided over two separate elements in the sentence is the Mande family. Mande languages have a strict S TAM O V X word order in which the verbal inflection on the verb is often limited to imperfective marking. But there are several other languages and groups of languages which have such a split inflection basic syntax: Kru and Senufo (plus some Gur languages in the same area) as well as the Bantu language Tunen have S TAM (O) O V (O) X as their word order. It would be interesting to compare the Cushitic languages with a “verbal piece” syntax with these West-African languages with a similar syntax. There are some similarities, for example, in an out-of-focus position immediately before the verb and the development of a focus position after the verb. But at the same time there are many differences too: in the West-African languages objects do not precede the TAM marker whereas in most of the Cushitic languages they do and Cushitic languages allow for adverbial elements within the verbal piece which is ruled out in a typical Mande syntax. In Heine’s typology of African languages (Heine 1976) these West-African S Aux O V X received their own type, type B, while the Cushitic languages were taken to be different, type D but this difference in typological classification is partly due to considerations of noun phrase syntax.

Notes

* This article was written during a research period at the Max Planck Institute for Evolutionary Anthropology in Leipzig. I am very grateful to all my colleagues there for fruitful discussions on this and various other topics.

1. The Highland East Cushitic languages being an important exception.
2. In this and following examples in the introduction (Section 1) what I consider to be the selector is in bold face.
3. The distinction definite/indefinite in Arbore refers to aspect marking, independent of the aspectual suffix to the main verb.
4. The following abbreviations are used (terminology, glosses and abbreviations are taken from original source): AGR for agreement, APPL for applicative, BEN for beneficent, COM for comitative, CSEC for consecutive, DEF.IND for definite indicative, DEM for demonstrative, DIR for directional, EXCL for exclusive, EMPH for emphasis, F for feminine, FOC and FC for focus, FM for focus marker, FRQ for frequentative, HAB for habitual, INCL for inclusive, IMPS for impersonal, IO for indirect object, IPF, IMPFV, or IMPF for imperfective, IRR for irrealis, M for masculine, N

for neuter, NEG for negative, O for object, OPT for optative, P for speech act participant, PERF or PF for perfect, PRED for predicative, POSS for possessive, PST for past, QUES for question, REAS for reason, REC for reciprocal, RED for reduplication, REFL for reflexive, S for subject, S.PRO for subject pronoun, SEL for selector, SFX for suffix, SG for singular, T/A for tense/aspect, VERB for verbal pronoun, VF for verb focus, 1 for first person, 2 for second person, 3 for third person.

5. In the sense that dependent clauses have no focus marker and no verbal pronoun.
6. Oromo, that is Harar-Oromo, has subject marking on the selector in the form of the first person singular only, namely *-n*, as shown in example (10).
7. In actual fact the glottal stop in the second person for Konso is very similar to the copy of the initial consonant of the verb root, indicated by C in the table, in Dullay, because the glottal stop in Konso assimilates to the following consonant and looks just like a copy of the initial consonant of the verb root.
8. But not for Saeed (1993:216).
9. A middle situation is not agentless but does not have the clear distinction between agent and patient as is the case in a standard active transitive situation.
10. Arbore does not have object pronouns.
11. But the verbal pronoun *hí* replaces a left-dislocated object.
12. Iraqw has transitive markers in addition to object pronouns.
13. The optional element *léhe* does not seem to have a semantic contribution and is possibly etymologically related to a verb 'to have' (Hayward 1984:226).
14. The indirect object series has *nu*, or *si*.
15. Identical to the absolute pronoun
16. The second series (indirect pronouns) of Boni have a high tone as only difference.
17. also post verbal
18. Another possibility is that the description is not full enough to show all the functions in the language.
19. And the selector is *wáxaa* originating in a "the thing, that... is X" construction.

References

- Amborn, H., Minker, G. and Sasse, H.-J. (1980). *Das Dullay: Materialien zu einer ostkuschitischen Sprachgruppe*. Berlin: Dietrich Reimer [Kölner Beiträge zur Afrikanistik 6].
- Appleyard, D. L. (1990). Prepositional particles in Somali and their cognates in other Cushitic languages. *African Languages and Cultures*, 3 (1), 15–31.
- Biber, D. (1984). The diachronic development of preverbal case markers in Somali. *Journal of African Languages and Linguistics*, 6, 47–61.
- Ehret, C. (1995). *Reconstructing Proto-Afroasiatic (Proto-Afrasian)*. Berkeley, CA: University of California Press.
- Hayward, R. J. (1984). *The Arbore Language: A first investigation including a vocabulary*. Hamburg: Helmut Buske [Kuschitische Sprachstudien 2].

- Hayward, R. J., & Saeed, J. I. (1984). NP focus in Somali and Dirayta: A comparison of *baa* and *pa*. In T. Labahn (Ed.), *Proceedings of the Second International Congress of Somali Studies, University of Hamburg, August 1–6, 1983*. Hamburg: Helmut Buske.
- Heine, B. (1976). *A Typology of African Languages. Based on the order of meaningful elements*. Berlin: Dietrich Reimer [Kölner Beiträge zur Afrikanistik 4].
- . (1980). *The Non-Bantu Languages of Kenya*. Berlin: Dietrich Reimer [Language and Dialect Atlas of Kenya 2].
- . (1982). *Boni Dialects*. Berlin: Dietrich Reimer [Language and Dialect Atlas of Kenya 10].
- Hetzron, R. (1980). The limits of Cushitic. *Sprache und Geschichte in Afrika*, 2, 7–126.
- Kiessling, R. (1994). *Eine Grammatik des Burunge*. Hamburg: Research and Progress Verlag [Afrikanistische Forschungen 13].
- Kiessling, R., & Mous, M. (2003). *The lexical reconstruction of West Rift* (Southern Cushitic). Cologne: Rüdiger Köppe Verlag.
- Kooij, J., & Mous, M. (2002). Incorporation: A comparison between Iraqw and Dutch. *Linguistics*, 40 (3), 629–645.
- Mous, M. (1993). *A Grammar of Iraqw*. Hamburg: Helmut Buske.
- . (2001). *Fieldnotes on Konso*.
- . (In preparation). *Alagwa: Grammar, texts, lexicon*.
- Owens, J. (1985). *A Grammar of Harar Oromo (Northeastern Ethiopia). Including a text and a glossary*. Hamburg: Helmut Buske [Kuschitische Sprachstudien 4].
- Pillinger, S., & Galboran, L. (2000). *A Rendille Dictionary Including a Grammatical Outline and an English-Rendille Index*. Cologne: Rüdiger Köppe [Kuschitische Sprachstudien 14].
- Saeed, J. I. (1999). *Somali*. Amsterdam: John Benjamins [London Oriental and African Language Library 10].
- Sasse, H.-J. (1981). Basic word order and functional sentence perspective in Boni. *Folia Linguistica*, 15 (3–4), 253–290.
- . (1984). The pragmatics of noun incorporation in Eastern Cushitic languages. In F. Plank (Ed.), *Objects: Towards a theory of grammatical relations* (pp. 243–68). London: Academic Press.
- Svolacchia, M. et al. (1995). Aspects of discourse configurationality in Somali. In K. Kiss (Ed.), *Discourse Configurational Languages* (pp. 65–98). New York, NY: Oxford University Press.
- Tosco, M. (1991). *A Grammatical Sketch of Dahalo: Including texts and a glossary*. Hamburg: Helmut Buske [Kuschitische Sprachstudien 8].
- . (2000). Cushitic overview. *Journal of Ethiopian Studies*, 23 (2), 87–121.
- . (2001). *Grammar of Dhaasanac*. Cologne: Rüdiger Köppe [Kuschitische Sprachstudien 17].

How Bantu is Kiyansi?*

A re-examination of its verbal inflections

Salikoko S. Mufwene

University of Chicago

(Kikongo-)Kituba, a contact language variety of central Africa, remains genetically unclassified. It is structurally so different from the languages of the ethnic Kikongo cluster that one must ask why, unlike Lingala (which is also considered a creole by most creolists), it has been completely overlooked by genetic linguists. Kituba's morphosyntax is also so peculiar that it prompted me to check whether any of the languages that came in contact with Kikongo-Kimanyanga, its "lexifier," may account for its partial isolating morphosyntax. My study of Kiyansi has revealed more than I had expected, for instance, the complete absence of Subject-Verb agreement – while there is evidence of object prefixation to the verb stem – and a significant proportion of sentence-final verbs of the OSV type.

This paper was prompted by my research on the development and morphosyntax of Kituba, a hitherto genetically unclassified language of Central Africa spoken primarily in the Western part of the Democratic Republic of Congo, in the Bandundu and Bas-Congo regions, and in the Southern part of the Republic of the Congo. It is a by-product of the exploitation colonization of Central Africa since the late 19th century, having developed out of the contact of Kikongo-Kimanyanga (its lexifier, according to Fehderau 1966) with primarily other Bantu languages (Mufwene 1994) and apparently with some West African languages (according to Samarin 1990). I was addressing the question why, unlike Lingala, which has similar origins (though its lexifier is Bobangi, according to Hulstaert 1989), Kituba has no Subject-Verb agreement markers, as below:

(1) Verb conjugation in Kituba:

- | | | |
|----|---------------------|---------------------|
| a. | <i>mono kee dia</i> | <i>mono dia-aka</i> |
| | me be eat | me eat-PAST |
| | 'I am eating' | 'I ate' |
| b. | <i>nge kee dia</i> | <i>nge diaaka</i> |
| | 'you-SG are eating' | 'you-SG ate' |

- | | | |
|----|----------------------|---------------------|
| c. | <i>yandi kee dia</i> | <i>yandi diaaka</i> |
| | 'he/she is eating' | 'he/she ate' |
| d. | <i>beto kee dia</i> | <i>beto diaaka</i> |
| | 'we are eating' | 'we ate' |
| e. | <i>beno kee dia</i> | <i>beno diaaka</i> |
| | 'you-PL are eating' | 'you-PL ate' |
| f. | <i>bo kee dia</i> | <i>bo diaaka</i> |
| | 'they are eating' | 'they ate' |

The question was: 'How could such a morphosyntactic peculiarity have developed out of the contact of primarily Bantu languages, which are agglutinating?' The question is made more relevant by the fact that one of the most salient peculiarities of Bantu languages is Subject-Verb agreement. I checked some Kimanyanga texts (Matuka 1990) and found out that although a handful of utterances did not display any obvious Subject-Verb agreement, as below in (2), in the vast majority of cases this Kikongo variety was consistent with the canon claimed in the Bantu linguistic literature:

- (2) Occasional absence of Subject-Verb agreement in Kimanyanga (Mufwene 1994):

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

I was then prompted to see whether some other languages of the area, which must have played an important role in the development of Kituba, had more cases of absence of Subject-Verb agreement. This led me to my own Kibwal dialect of Kiyansi (B85 in Guthrie's 1953 classification), for which I introspected the following sample of verb conjugation:

- (3) Sample Verb Conjugation in Kiyansi (Mufwene 1994):

<i>mε mā 'dia</i>	<i>nzé mā 'dia</i>	<i>ndí mā 'dia</i>	<i>bí mā 'dia</i>
me NPERF eat	you.SG NPERF eat	s/he NPERF eat	we NPERF eat
'I have eaten.'	'You have eaten.'	'S/he has eaten.'	'We have eaten.'
<i>mε ε dí</i>	<i>nzé ε dí</i>	<i>ndí ε dí</i>	<i>bí ε dí</i>
I RPERF eat	you.SG RPERF eat	s/he RPERF eat	we RPERF eat
'I (already) ate.'	'You (already) ate.'	'S/he (already) ate'	'We (already) ate.'
<i>mε āyí 'dia</i>	<i>nzé āāyí 'dia</i>	<i>ndí āyí 'dia</i>	<i>bí āyī 'dia</i>
me FUT/GO eat	you.SG FUT eat	s/he FUT eat	we FUT eat
'I will eat.'	'You will eat.'	'S/he will eat.'	'We will eat.'
<i>'mε a di+áná</i>	<i>nzé á di+áná</i>	<i>ndí á di+áná</i>	<i>bí a di+áná</i>
I NAR eat.HAB	you.SG NAR eat.HAB	s/he NAR eat.HAB	we NAR eat.HAB
'I usually eat.'	'You usually eat.'	'S/he usually eats.'	'We usually eat.'

There is no doubt that the beginning of the verb *'dia* 'eat' changes in these conjugations, but the changes have nothing to do with PERSON and NUMBER. They are determined by time reference. There are two particular patterns: Sometimes the verb starts without a prefix at all, since the time reference marker (aspectual in this case) is a separate word (a topic to which I return below). At times it starts with what seems to be a prefix. Whichever marker it is, it remains the same for all persons and numbers. There is, however, evidence from some conjugated verb forms that Kiyansi-Kibwal is still agglutinating, though it does not agglutinate everything and lacks Subject-Verb agreement. The following constructions show indeed that it is still similar to Kimanyanga in the way pronominal elements are incorporated:

(4) Agglutinated/Incorporated object pronouns in Kiyansi:

- a. *mε m+̂+sámí.* 'I have told him/you.'
I NPERF+you/him+tell
- a'. *mε nzé m+̂+sámí.* 'I have told you.'
I you-SG NPERF+you+tell
- a''. *mε (ndí) m+̂+sámí.* 'I have told him/her.'
I him/her NPERF+him/her+tell
- b. *mε (beá:g)² m+é+sámí.* 'I have told them.'
I (them) NPERF+them+tell
- c. *mε (bé) m+é+sámí.* 'I have told you.'
I (you-PL) NPERF+you+tell
- d. *bé (mé) ma+n+sámí.* 'You have told me.'
you-PL (me) NPERF+me+tell
- e. *nzé (bí) m+é+sámí.* 'You have told us.'
you-SG (we) NPERF+we+tell
- f. *bí (ndí) ma+k+̂+sámí.* [*k+ < k'ε* 'go']
We (he/she) NPERF+go+him+tell
We are leaving to tell him/her/We are going to tell him.
- g. *ndí (nzé) ma+y+̂+sámí.* [*y+ < ya* 'come']
he/she (you) NPERF-come-you/him-tell
He/She has come to tell you/him.

(5) Agglutinated/Incorporated object pronouns in Kimanyanga:

- a. *bu n+tel+ele mono.* 'This is what I personally said.'
this I+tell+NPERF me-EMPH
- b. *mo+mo u+n+tel+ele mono.* 'What you told me'
PL+what you+me+tell+NPERF me-EMPH
- c. *u+m+bak+idi.* 'You got me/him/her.'
you+me/him/her+get+NPERF

- d. *mono mpe n+kaanda Ø+ye+taanga wo.* 'The letter, I too read it.'
 me too letter I+PAST+read it
- e. *mono bu Ø+yi+zol+anga.* 'I like it this way.'
 me this-way I+it+like+HABIT
- f. *tu+mwiin+i yo.* 'We have seen it.'
 we+see+NPERF it
- g. *mono mpe n+kutu we oa.* 'I too was about to say it.'
 me too AGR+about it say

Kiyansi is like Kimanyanga in that not all object pronouns are incorporated in the verb complex. Examples (5f) is a particularly good one for Kimanyanga. (5g) is debatable, although the preverbal position of *we* suggests that it be treated as a prefix. In both languages the incorporated object pronoun precedes the verbal base morpheme. One thing is also obvious, viz., that the subject prefix in Kimanyanga is sometimes null, which may amount to saying that Subject-Verb agreement is not so regular in this language. When the object pronoun is a free morpheme it occurs post-verbally, consistent with the SVO pattern of major/free constituents of the basic sentence structure. However, the word order often diverges from this basic, unmarked one, as the OVS order in (5a), the OVO order in (5b), and the SOVX order in (5d).

Quite obvious about Kiyansi is the fact that there are very few object pronominal prefixes. With the exception of the first person singular, there are only two forms, viz., \mathfrak{c} for the singular and ϵ for the plural. The paucity of object pronominal prefixes is confirmed by the absence of a reflexive prefix in Kiyansi-Kibwal, as illustrated below:

- (6) Reflexive construction in Kiyansi:
- a. *ndí á+tyē:n* *I* *ndí ngáàkwēāg.*
 he/she NARR+talk CONN he/she EMPH.SELF
 'He/She is talking to him-/herself.'
- b. *ndí ngáàkwēāg y+í.*
 he/she EMPH-SELF COME+NPERF
 'He/She him-/herself came.'
- c. *ndí á+tyē:n* *I* *ndí ntú.*
 he/she NARR+talk CONN he/she alone
 'He/She is talking alone.'
- d. *ndí y+í:* *I* *ndí ntú.*
 he/she COME+NPERF CONN he/she alone
 'He/She came alone'
- (7) Reflexive construction in Kimanyanga:
- a. (*Yandi*) *ka+ki+zol+ele.*
 he/she AGR+REFL+like+NPERF
 'He/She likes him-/herself.'

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

In this respect, Kiyansi-Kibwal is quite different from the Bantu canon. Not only does it not have a reflexive affix; one can even argue that it does not have a pronoun that has specialized for the reflexive function. Like English and other languages, such as Hebrew, it uses a phrase whose basic function seems to be emphatic to express reflexivization as strictly coreference of Subject and Object in this case, i.e., the subject is the recipient of its own action or feelings. As the reader may also have guessed from some of the above examples, Kiyansi-Kibwal is not a Pro-Drop language. In this respect, it both diverges from the Bantu canon and differs from Kimanyanga. The subject must always be specified overtly, either as a noun or as a free pronoun.

One of the things to learn from some of the examples is the extensive way in which Kiyansi-Kibwal uses free time reference markers. Here are some examples of the present progressive, which consists of a combination of a demonstrative locative (proximal or distal) with a conjugated verb. It literally means 'Subject BE HERE/THERE Verb-ing' (although there is no language-internal evidence for calling the verbal form a gerund or present participle):

(8) Present progressive in Kiyansi-Kibwal:

<i>mε éwu a+diā</i>	<i>nzé éwú a+diā</i>	<i>nzé éwàá a+diā</i>
I here NARR-eat	you.SG here NARR-eat	you there NARR-eat
'I am eating'	'you are eating'	'you are eating'
	<i>ndí éwú ā+diā</i>	<i>ndí éwàá ā+diā</i>
	he/she here NARR-eat	he/she there NARR-eat
	'he/she is eating'	'he/she is eating'
<i>bí ébá ā+diā</i>	<i>bé ébá ā+diā</i>	<i>bé ébàá ā+diā</i>
we here NARR-eat	you.PL here NARR-eat	you.PL there NARR-eat
'we are eating'	'you are eating'	'you are eating'
<i>béá:g ébá ā+diā</i>	<i>béá:g ébàá ā+diā</i>	
they here NARR-eat	they there NARR-eat	
'they are eating'	'they are eating'	

This is a paradigm quite different from what is attested in Kimanyanga, where the progressive is formed with the copula verb *keti* (which has become *kele* in Kituba) or *kena*, as in:

(9) Progressive construction in Kimanyanga:

- a. *W+eta ku+n+tuma*. 'He is sending me.'
 he+be INFIN+me+send
- b. *K+eti Ø+pelesa*. 'He is rushing.'
 he+be rush

- c. *tw+ena*Ø+ *zenga*. 'We are deciding.'
we+be decide
- d. *me-eso m-ena (twa) taanga*. 'The eyes are (busy) reading.'
6-eye AGR6-be (?) read
- e. *bi-mbungu by-ena (twa) simwa*. 'The hyenas are suffering.'
4-hyena AGR4-be (?) suffer

Although Kimanyanga too resorts to a periphrastic construction for the progressive, Kiyansi-Kibwal exhibits a locative construction pattern that varies according to whether the subject is first person or otherwise. When the subject is first person, the marker must be the proximal demonstrative *éwú* for SINGULAR or *ébá* for PLURAL. In the other cases, the progressive marker is either one of the same proximal demonstratives, when the referent is near the speaker, or, when the referent is far from the speaker, one of the distal demonstratives *éwàá* when the subject is SINGULAR and *ébàá* when the subject is PLURAL. This information should interest students of grammaticalization, and the limited data presented here contains quite a few interesting cases. For instance, the FUTURE construction uses the verb +*y* 'come', apparently inflected in the NEAR PERFECT, AS IN:

- (10) *ndí ā+y-śsāmī*
he/she NARR-come-him/her-tell
'he/she will tell him/her'

The PERFECT marker *ma* is a cognate of the verb *mene* in Kituba and *meni* in Kimanyanga, though there is no evidence of such a verb in Kiyansi. The verbs for 'finish' in Kibwal is *ɔwa* for intransitive uses *ɔmaayi* for transitive ones. The latter takes only nominal and pronominal objects, no clausal ones.

The question of what is the grammatical status of these time reference markers is a digression I prefer to stay away from for now. I would also like to point out that the verb inflectional paucity of Kiyansi-Kibwal does not reflect any overall paucity in all aspects of its system. For instance, it has more than 10 vowels by my count (Mufwene 1973), including some centering diphthongs (as in *mbēāg* 'fire' and *k^wēg* 'hearth stone', and as the reader may have suspected from the form of the third person plural pronoun. However, the verb inflection paucity is consistent with lack of agreement with the head within the noun phrase, as can be observed in the following examples:

- (11) Lack of Head-Modifier agreement in Kiyansi-Kibwal's NP:
- | | | | | |
|-------------|----------------------|------------------|--------------------|--------------|
| <i>m̄</i> | <i>mw+áán m̄</i> | 'one child' | <i>d+íy ī m̄</i> | 'one eye' |
| one | 1-child one | | 5-eye EPENTH one | |
| <i>wé.l</i> | <i>b+áán bwé.l</i> | 'two children' | <i>m+íy mwé.l</i> | 'two eyes' |
| two | 2-child two | | 6-eye two | |
| <i>tār</i> | <i>b+áán ā tār</i> | 'three children' | <i>m+íy ā tār</i> | 'three eyes' |
| three | 2-child EPENTH three | | 6-eye EPENTH three | |

<i>mw+áán ī mē</i>	‘my child’	<i>d+íy ī mē</i>	‘my eye’
1-child EPENTH me		5-eye EPENTH me	
<i>b+áán ī mē</i>	‘my children’	<i>m+íy ī mē</i>	‘my eyes’
2-child EPENTH me		6-eye EPENTH me	

The variation observable here seems to be triggered by something other than a principle of Head-Modifier agreement, although I have not figured out yet what is going on in this case.

What prompted me to write this paper is that Kibwal is impoverished in a morphosyntactic respect that seems critical to identifying a language as Bantu. Based on one of the incidental examples adduced above, it differs from the Bantu vocalic canon of five or seven vowels. I would not be surprised if it diverged in any other significant way. Kiyansi was classified by Guthrie (1953) as a Bantu language, group B85. I have no reason to doubt that it is Bantu, especially because its speakers are Bantu people and Kiyansi shares a large proportion of its vocabulary with other Bantu languages, by which it is geographically surrounded. It also has noun classes and is agglutinating in some ways, as it still incorporates an object pronoun and can “extend” a verb by suffixation to the stem – something that I have not discussed in this paper but can be illustrated with the following examples:

- (12) a. *ndí mā k’ír*
 he/she PERFECT do
 ‘he/she did/has done [it]’
 a’. *ndí mā n+k’ír+i*
 he/she PERFECT me-do-APPL
 ‘he/she did/has done [it] for me’
 b. *ndí mā m+búl*
 he/she PERFECT me hit
 ‘he/she did/has hit me’
 b’. *ndí mā m+búúl+i*
 he/se PERFECT me+hit+APPL
 ‘he/she did/has hit [it] for me’
 c. *mē ma sāl*
 I PERFECT work
 ‘I have worked’
 c’. *mē m+ś+sāāl+i*
 I PERFECT you+work+APPL
 ‘I have worked for you’

Examining these structural facts from Kimanyanga and Kibwal, one is somewhat reminded of the way English and German are both identified as Germanic languages although they do not share identical etyma and differ with regard to word order and case marking. One can hardly help asking what the role of structural features is in determining whether a language is genetically Bantu, Germanic, of something else. To the

extent that structural features play an important role in the classification of languages, which aspects of the morphosyntax or of the lexicon or phonology are more critical? Regarding Bantu, Williamson and Blench (2000) discuss mostly the noun class system, with Subject-Verb agreement mentioned only in a table. Does this mean that the noun class system is more important? Or did they mean to suggest that the noun class system is primary and the Subject-Verb agreement is a consequence of it and therefore secondary?

Then I cannot help asking a few other questions, such as: Does the absence of Subject-Verb agreement markers reflect impoverishment from Proto-Bantu or does it reflect variation in Proto-Bantu? If it reflects impoverishment, how consistent is this with its complex vocalic system? Or is Kiyansi-Kibwal one of those odd languages with peculiarities that set it apart from a more general pattern, as is the case of Ijo in Nigeria, which differs in respect to, for instance, word order and adpositions from the Eastern Kwa languages?

I have no answers to these questions, especially because it now seems necessary to investigate more closely other morphosyntactic aspects of Kiyansi – which I could not do within the limited time available to me before this meeting – and other Bantu languages of group B and of the region. But I thought these facts were worthwhile reporting. At least I know we need not invoke a *deus ex machina*, in the form of children's innovations *ex nihilo* (or any unmarked values of parameters of Universal Grammar with regard to morphosyntax) in order to account for the absence of Subject-Verb agreement in Kituba. Irregularities in this respect in Kimanyanga and lack of this Bantu peculiarity in Kiyansi and undoubtedly in other languages must have contributed to the development of the same peculiarity in Kituba. Assuming there is no particular reason for reclassifying Kiyansi-Kibwal within Bantu, the question of what is the most critical criterion, or what are the most critical criteria, for classifying a language as Bantu remains worth addressing.

This paper undoubtedly joins others before it, such as Heine (1980), in highlighting the structural heterogeneity of Narrow Bantu, ironically within what Heine (1980:337) identifies as “Congo Bantu,” which he claims to have the most established classification. Given all the missing elements of comparison from this paper, it may very well be that there is nothing wrong with the classification, but then the criteria for the classification still must be articulated more explicitly. I thought a statement of the kinds of questions Kiyansi-Kibwal raises may interest geneticists and Africanists. If Kiyansi fits its classification, despite its partial divergence from the Bantu structural canon, it is more evident that genetic kinship does not go hand in hand with typological kinship.

Notes

* This essay was written in 2001. Since then Derek Nurse and Gérard Philippson have edited *The Bantu languages*, a precious encyclopedic anthology in which one can now find a lot of useful

information about this language family. The survey chapters by especially Thilo Schadeberg (on derivation), Derek Nurse (on aspect and tense), and Thomas Bearth (on syntax) complement my discussion by providing more information about morphosyntactic variation, highlighting the family resemblance principle which obviously underlies the grouping of languages into large genetic families. I would perhaps have written a different kind of paper had I read Bearth's chapter before my scholarly exercise, for instance by adducing more pragmatic considerations to bear on word order variation. Two important methodological questions are nonetheless worth stating (over) here: 1) What particular (combinations of) structural features are more heavily weighted in genetic classifications? 2) Has genetic classification generally proceeded independent of non-linguistic information, such as about the ethnic and cultural ties claimed by the speakers?

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

References

- Fehderau, H. (1966). *The Origin and Development of Kituba* PhD. Dissertation, Cornell University.
- Guthrie, M. (1953). *The Bantu languages of Western Equatorial Africa*. Oxford: Oxford University Press.
- Heine, B. (1980). Some recent developments in the classification of Bantoid. The problem of defining Bantu. In L. Bouquiaux (Ed.), *L'expansion bantoue, Actes du Colloque International du CNRS, Viviers (France) 4–16 avril 1977* (pp. 333–340). Paris: SELAF.
- Hulstaert, G. (1989). L'origine du lingala. *Afrikanistische Arbeitspapiere*, 17, 81–114.
- Matuka, Y. M. (1990). Three Koongo palavers: Wedding, bereavement and reconciliation. Ms., Ball State University.
- Mufwene, S. S. (1973). *The Segmental Phonemes of Yansi and English*. "Licence" thesis, Université Nationale du Zaïre, Campus de Lubumbashi, Département de Langue et Littérature Anglaise.
- . (1994). Restructuring, feature selection, and markedness: From Kimanyanga to Kituba. In K. E. Moore, D. A. Peterson, & C. Wentum (Eds.), *Historical Issues in African Linguistics* (pp. 67–90). Berkeley, CA: Berkeley Linguistics Society.
- Nurse, D., & Philippson, G. (Eds.). (2003). *The Bantu languages*. London: Routledge.
- Rottland, F. (1970). *Die Verbform des Yanzi*. Leiden: Leiden University.
- Samarin, W. (1990). The origins of Kituba and Lingala. *Journal of African Languages and Linguistics*, 12, 47–77.
- Williamson, K., & Blench, R. (2000). Niger-Kongo. In B. Heine & D. Nurse (Eds.), *African Languages: An introduction* (pp. 11–42). Cambridge: Cambridge University Press.

Diathesis alternation in some Gur languages

Brigitte Reineke and Gudrun Mieke

Humboldt-University at Berlin and University of Bayreuth

This paper aims at analysing the strategies used in Gur languages to express change of valency in transitive-active verbs. The analysis will show that such verbs may occur in both NP–V– NP and NP–V sentence patterns without morphological marking. This capacity allows classification of these verbs into two categories, i.e. ‘flexible’ vs ‘non-flexible’, regarding to their valency.

In contrast to other Niger-Congo languages, where voice diathesis alternations are marked by verbal derivational forms, Gur and Mande developed a different formative principle which does not necessitate morphological changes to the verbal form in question, but is based on the presence or the suppression of the second argument of a ‘transitive’ verb. It is this type of valency change which enables the languages to express change of voice.

1. Introduction

This paper deals with the valency of verbs in some Gur languages, aiming in particular at analysing the strategies used in these languages to express change of valency in transitive-active verbs. These changes will be seen to involve voice diathesis alternation¹ and we will illustrate how this capacity leads to a new parameter of verb classification. The first part of our paper deals with those changes of valency which are carried out without morphological marking. In spite of identical formational structures, each language has developed its independent set of constraints for the use of certain verbs in the interplay between the transitivity and detransitivization processes. The second part of the paper discusses the changes of valency expressed by the morphological means of verbal derivation.

The languages dealt with are classified as Central Gur, but spoken in opposite geographical regions. Ditammari and Byali are members of the Eastern Oti-Volta subgroup of Northern Central Gur, Kɛansa belongs to the Southern branch of Central Gur. These languages display a number of peculiarities absent or not found to such an extent in other Gur languages.

In all Central Gur languages, the unmarked transitive sentence is represented by the structure S P O. The degree of activity which is carried over from an agent to a patient varies according to certain semantic features, i.e. the cognitive concepts expressed

by the verb of a given predication. This becomes most obvious through the semantic roles taken over by the respective participants. Transitive verbs in these languages are linked to a high degree obligatorily with their objects,² which may nevertheless be deleted for contextual and pragmatic reasons.

In contrast to many Niger-Congo languages, where voice diathesis alternation is marked by verbal derivational forms (see Voeltz 1977), Gur and Mande have developed a different formative principle, one which does not necessitate morphological changes to the verbal form in question, but is based solely on the presence or the suppression of the second argument of a 'transitive' verb. It is this type of valency change which enables the languages to express change of voice.

2. Change of valence without morphological marking

The analysis of verbs which may occur in transitive constructions reveals variation in voice diathesis alternation. All the languages under consideration distinguish between verbs occurring exclusively in a transitive-active manner (see 2.1) and other verbs which can further be used with a passive-like sense (see 2.2). The latter group outnumbers the first group. But there are language specific constraints – motivated by the semantic properties of verbs – which have an impact on their syntactic behaviour.

It should again be emphasized that these languages do not apply any morphological strategies to mark this kind of diathesis alternation. Hence, the transitive-active and the intransitive-passive variants of the verb are not formally distinguished at all.

Whether one is dealing in a concrete case with the active or with the passive variant becomes clear only at the sentence level, when grammatical relations, semantic roles as well as the pragmatic functions of the verb arguments are recognisable. In intransitive-passive constructions, the agent participant of the verb is suppressed, and the patient participant moves to the subject position. It is generally accepted that both phenomena are largely pragmatically motivated. The suppression of the agent is commonly due to its low topicality. In such an intransitive-passive construction, the patient participant functioning as a non-prototypical grammatical subject represents the topic of the sentence about which a comment is made.

In dealing with event-views and transitivity in the Supyire verbal system, Carlson (1997:26; 2000:41) uses the term 'labile' for those verbs which – regarding their valency features – can neither be characterized unequivocally as bivalent – transitively used – nor as monovalent – intransitively used. He also emphasizes the interaction between the pragmatic structure, the semantic roles of the verb arguments and the grammatical syntactic structure.³ In using the term 'lability' he follows Haspelmath (1993), who refers to its use in Caucasian linguistics.⁴

However, for this special case of valency change without morphological strategy, we prefer the term 'valency flexibility'. Accordingly, we use the term 'flexible' for those verbs which can occur in both S P O as well as S P structures. As already indicated, they represent the majority of the verbs in question.

The semantic range of the verbs characterized by the feature ‘flexibility’ or ‘non-flexibility’ respectively varies from language to language. These differences – as well as the different constraints concerning the use of flexible verbs in passive-like sentences – are an indication of the wide scope of language specific concepts of transitivity and their relevance for expressing diathesis alternation.

The notion that verb behaviour and verb meaning are inherently connected (Levin 1993; Levin et al. 1995 and others) is widely accepted. Although alternations are manifested across languages by verbs of the same semantic types, language specific semantic analyses are still necessary, since even verbs in closely related languages may behave differently in this regard.

Regarding our data from Gur languages, some proof of the relevance of the semantic properties of verbs is being established. There is evidence that the possibility or non-possibility of transitive verbs being used in a passive sense is dependent – above all – on the semantic properties of the direct object.

2.1 Non-flexible verbs

Non-flexible verbs in this context are verbs that occur exclusively in a transitive-active manner. In both corpora, that of Eastern Oti Volta (Ditammmari/Byali) as well as that of Kɛansa, non-flexible verbs are in the minority. For the time being, their semantic features are best identified in Kɛansa (cf. 2.1.2).

2.1.1 Eastern Oti Volta languages

Since the verbal system in all Eastern Oti-Volta languages is characterized by an aspectual opposition between perfective and imperfective, which are overtly marked, every verb exists in two morphologically or tonologically differentiated aspectual forms, unless its semantic features are incompatible with one of these aspects. In a few cases, suppletive verbs express the imperfective aspect. Whereas the temporally unmarked imperfective verbal form serves to express the present tense, the unmarked perfective form is linked to the past due to its aspectual perfective character of completion. Since the aspectual distinction is basic to the verbal system of these languages, there are only two temporal markers that localize the state of events on the time axis, a past as well as a future marker. They can be linked with both aspectual forms, at least when the verbs are actively used. In the following, both aspectual forms of non-flexible Ditammari verbs are given:

(1)	PFV	IPFV		PFV	IPFV	
	kūú	kú	to vaccinate	tyé'	tīē'	to advise
	hūā	hūà	to wash (s.b.)	cōōtá	cōó	to get; to earn
	fītā	fítí	to sell	dōntá	dó'	to buy
	fū'	fúkú	to plant	tōtá	tōó	to tear

2.1.2 *Kaansa*

In contrast to the Eastern Oti-Volta languages, *Kaansa* does not have an overtly marked aspect system. The verbal system consists of a net of preverbal and postverbal morphemes interwoven by the existence of two different verbal stems⁵ which sometimes take on aspectual functions. In the following, the verbs are given in their (nominalized) citation form which always takes the suffix */-mV/*. Approximately one third of all transitive-active verbs have to be classified as non-flexible verbs regarding voice. In particular, verbs belonging to the cognitive dimension, which Halliday (1994) refers to as the 'world of consciousness', necessarily require [+animate] participants in both the subject as well the object position. Accordingly, almost all verbs of saying are found in this class, e.g.:

- | | | | | |
|-----|---------------------------|----------------------------|----------------------------|----------------------|
| (2) | <i>kha²-ma</i> | to speak, to criticize | <i>thoo-mo</i> | to relate, to repeat |
| | <i>kpir-ma</i> | to tell stories or riddles | <i>thia²-ma</i> | to greet |
| | <i>pɛ²-ma</i> | to ask | <i>yí²-ma</i> | to call |

This is also the case for verbs expressing mental as well as behavioural processes, e.g.:

- | | | | | |
|-----|--------------------------|------------------------|----------------------------|------------------------|
| (3) | <i>kɛ-mme</i> | to look at, to believe | <i>kusa²-ma</i> | to insult |
| | <i>pákí-nsa</i> | to compete, to deny | <i>thu-ma</i> | to judge |
| | <i>bɔ²-mɔ</i> | to advise, to educate | <i>fɪ-mma</i> | to deceive, to flatter |
| | <i>mamɔ-ma</i> | to learn | <i>sásáa-ma</i> | to disturb s.b. |
| | <i>i²-me</i> | to hate | <i>khɔɔ²-mɔ</i> | to love, to like |

Note that, in contrast, perception verbs like 'to see', 'to hear', 'to listen' are flexible.

Some verbs from Halliday's 'physical world', which include 'material' processes like 'happening (being created)', 'creating/changing' and 'doing (to)/acting', belong to the group of non-flexible verbs in *Kaansa*, but only under the condition that the object is specified as [+animate], e.g.:

- | | | | | |
|-----|----------------|---------------------|--------------------------|--------------------------|
| (4) | <i>fata-ma</i> | to slap in the face | <i>kɔ-mma</i> | to embrace |
| | <i>sɔ-mma</i> | to send s.b. | <i>nɔ-mma</i> | to bite, to inflict pain |
| | <i>sɛv-ma</i> | to sting | <i>du²-mo</i> | to bite, to cut |

This semantic group also includes actions effected on animals:

- | | | | | |
|-----|--------------------------|---------------------------|---------------------------|-------------|
| (5) | <i>be²-me</i> | to breed, to farm animals | <i>phɪ²-ma</i> | to castrate |
|-----|--------------------------|---------------------------|---------------------------|-------------|

Some verbs expressing ingestion as well as food preparation in general are also counted as non-flexible verbs, e.g.:

- | | | | | |
|-----|---------------|--------------------------|--------------------------|--------------------------------|
| (6) | <i>núu-ma</i> | to drink | <i>khɔɔ-mɔ</i> | to chew, to ruminate |
| | <i>dɛɛ-me</i> | to taste | <i>pɔɔ-mɔ</i> | to mix flour with water |
| | <i>sur-mo</i> | to sprinkle, to spew out | <i>tɪ²-ma</i> | to spit, to write ⁶ |

But note that, on the other hand, verbs like 'to eat' and some other verbs related to the preparation of food belong to the class of flexible verbs. The following group of verbs do not provide an indication of their semantic explanation at all:

- (7) *khɪ²-ma* to strike, to beat the drum *mɛ̃-mma* to kindle
púkɪ-mo to weed *súu-mo* to steal
*dori²-mo*⁷ to pay for several items

To summarize, the semantic definition of the group is not homogeneous. Most of the verbs belong to Halliday's 'sensing' domain (1994), the rest stem from the domain of 'doing'. As for the feature [+animate], which seems to be a common characteristic of the patient affected by non-flexible verbs, we have to concede that a small group of verbs requiring [-animate] patients are nevertheless members of this group.

2.2 Flexible verbs

Another group of verbs bears the feature of valency flexibility, i.e. one has to assign such verbs a transitive-active as well as a intransitive-passive meaning; however, in spite of this shared feature, language specific constraints need to be taken into consideration.

2.2.1 Eastern Oti-Volta languages

2.2.1.1 Aspect and tense in the intransitive passive use of verbs

In the following, attention will be focussed on the intransitive-passive use of verbs. First, the sentences (8)–(11) show the transitive-active use of the respective verbs in NP-V-NP constructions, in which the second participant has the semantic role of a patient.⁸

Transitive-active use

Ditamhari

- (8) PFV *ō* *kòòtā kũ- tōō-kũ*
 CLS sweep CL room-CL s/he has swept the room
 ō dō *kòòtā kũ- tōō-kũ*
 CLS PAST sweep CL room-CL s/he had swept the room
 ō bō *kòòtā kũ- tōō-kũ*
 CLS FUT DUR sweep CL room-CL s/he will sweep the room
- (9) IPFV *ō* *kòõ kũ tōō-kũ*
 CLS sweep CL room-CL s/he sweeps the room
 ō dō *kòõ kũ tōō-kũ*
 CLS sweep CL room-CL s/he swept the room
 ō bō m *kòõ kũ tōō-kũ*
 CLS FUT DUR sweep CL room-CL s/he sweeps the room

Byali

- (10) PFV *ū* *sāārə̌* *sə̌* *sáy-ī*
 CLS sweep PFV room-CL s/he has swept the room
ū ī *sāārə̌* *sə̌* *sáy-ī*
 CLS PAST sweep PFV room-CL s/he had swept the room
ū yí *sāārə̌* *sə̌* *sáy-ī*
 CLS FUT sweep PFV room-CL s/he will sweep the room
- (11) IPFV *ū* *sāārú* *sáy-ī*
 CLS sweep room-CL s/he sweeps the room
ī *sāārú* *sáy-ī*
 CLS sweep room-CL s/he swept the room
ū yí n *sāārú* *sáy-ī*
 CLS FUT DUR sweep room-CL s/he will sweep the room

There are instances of both perfective and imperfective verbal forms used intransitive-passively (cf. (12)–(15)). As for the perfective forms being used passively, there seems to be an ambiguity insofar as they can be interpreted dynamically as well as statively. A stative interpretation is given when the result of a (dynamic) activity is specified (= stative-resultative state of event). In contrast, the imperfective aspectual form allows only a dynamic view of the event.

The examples given verify these two kinds of interpretation, each being directly linked to the aspectual verbal forms.

Concerning the passive use of the imperfective aspect, there are differences between Ditammari and Byali. In Byali, the temporally unmarked imperfective form expresses an action which will be realized at a non-specific moment in the future, i.e. it does not allow a reading of ongoing action. So, being combined with the past tense marker, the passively used imperfective aspect expresses a non-ongoing action in the past. A combination with the future marker is not possible. In contrast, in Ditammari, the imperfective aspectual form does not occur in the passive sense. It is a grammaticalized progressive⁹ form that – in indicating an ongoing action in the passive, as in ‘the room is being swept’ – serves to express this meaning in the imperfective aspect.

*Intransitive-passive use**Ditammari*

- (12) PFV *kūtōokū* *kòtā* ‘the room has been swept/is swept’
kūtōokū qō *kòtā* ‘the room had been swept/was swept’
kūtōokū bō *kòtā* ‘the room will be swept’

Byali

- (13) PFV *sáyī* *sāārəsə̌* ‘the room has been swept/is swept’
tāpú ī *wāyəsə̌* ‘the house has been destroyed/was destroyed’
tāpú yí *máásə̌* ‘the house will be built’

*Ditammari (Progressive construction)*¹⁰

(14)

IPFV	<i>kū</i>	<i>tōo-kū</i>			<i>bō</i>	<i>kā</i>	<i>kə̌</i>	
	CL	room-CL			be.LOC	CONSEC	sweep	‘the room is being swept’
	<i>kū</i>	<i>tōo-kū</i>	<i>dō</i>		<i>bō</i>	<i>kā</i>	<i>kə̌</i>	
	CL	room-CL	PAST		be.LOC	CONSEC	sweep	‘the room was being swept’
	<i>kū</i>	<i>tōo-kū</i>	<i>bó</i>	<i>m</i>	<i>bō</i>	<i>kā</i>	<i>kə̌</i>	
	CL	room-CL	FUT	DUR	be.LOC	CONSEC	sweep	‘the room will be swept’

Byali

- (15) IPFV *sávi* *sāarú* ‘the room will be swept’ (sometimes in the future)
sávi *ī* *sāarú* ‘the room has been swept’ (at some point in the past)
tāpú *wāylsú* ‘the house will be destroyed (=cracked)’
tāpú *ī* *wāyàsú* ‘the house was destroyed (=cracked)’

The last Byali examples indicate the problems of semantic interpretation which may arise with regard to passively used imperfective forms. Whereas the verb 'to sweep' expresses unequivocally a passive perspective, the interpretation of 'to destroy' offers two interpretations, a clearly passive one and an intransitive-active reading which calls to mind middle verbs in other languages.

2.2.1.2 Constraints with regard to the semantic roles of the participants

As already indicated above, our research is providing evidence in all languages for the relevance of meaning in determining different syntactic behaviour patterns. The following sentences containing the verbs 'to beat' and 'to bite' make clear that whether they may occur in a passive-like sense is dependent on the semantic properties of the direct object.

Transitive-active use: NP–V–NP[+ANIM]

- (16) Ditammari *cī́tá pwṓtá qā́bírá*
 Byali *pwḗú pwóns(ə) bííyā*
 ‘The father has beaten the child’
- (17) Ditammari *támž́tá bṓtá qā́bírá*
 Byali *búú́yā hāsə́sə bííyā*
 ‘The dog has bitten the child’

The attempt to raise the animate patient participant of these sentences ('the father has beaten *the child*' and 'the dog has bitten *the child*') to the grammatical subject position shows that a feature (+animate) of the object is incompatible with the passive diathesis alternation. The resulting structure NP^(+ANIM) – V (in (18), (19)) can only be interpreted in an active sense: 'the child has beaten/has bitten'.

Intransitive-active use: NP– V – NP[+ANIM]

- (18) Ditammari dābírā pwōtá
 Byali bíivā pwóns(ā)
 ‘The child has beaten’
 not: *‘The child has been beaten’
- (19) Ditammari dābírā bōtá
 Byali bíivā hāsàsā
 ‘The child has bitten’
 not: *‘The child has been bitten’

In contrast, the alternation is possible when the object of the transitive active clause is inanimate, cf. (20) and (21). In such a case, regardless of the morphological identity of the verb forms any ambiguity is totally excluded, as the examples given under (22) and (23) show.

Transitive-active use: NP– V – NP[–ANIM]

- (20) Ditammari dābírā pwōtá kuyòòkù
 Byali bíivā pwóns(ā) yīā
 ‘The child has pounded the millet’
- (21) Ditammari dābírā bōtá kúmāā
 Byali bíivā hāsàsā kóndí
 ‘The child has bitten the meat’

Intransitive-active use: NP– V – NP[–ANIM]

- (22) Ditammari kuyòòkù pwōtá
 Byali yīā pwóns(ā)
the millet has been pounded
- (23) Ditammari kúmāā bōtá
 Byali kóndí hāsàsā
the meat has been bitten

The relevance of the semantic roles of a participant also becomes evident in the Ditammari verb pair,

- (24) PFV pītírā / PFV pítí ‘to tie’
 PFV pītā / PFV pìtīrī ‘to untie’

however in a somewhat different way. The pair PFV pītírā / PFV pítí ‘to tie’ belongs to the subgroup of verbs which can be used intransitive-passively, at least in a sentence non-final position, (see 2.2.1.3). Such a use is also possible when the patient as subject of the sentence is animate.

- (25) *dīkàrì pītīrā kūḍḍ yīnkà* ‘the chair has been tied up on the wood’
tāmṣṭà pītīrā káqūā ‘the dog has been tied up and sleeps’¹¹

But when used sentence-finally, this verb does not allow voice diathesis:

- (26) **tāmṣṭà pītīrā* ‘the dog has been tied up’

In contrast, such an occurrence is attested for PFV *pītā* / PFV *pītīrī* ‘to untie’, even if the patient has the feature [+animate]:

- (27) *tāmṣṭà pītā* ‘the dog has been unfastened (the dog has broken free)’¹²
ōnītī pītā ‘the man has been untied’

The fact that an animate subject is able to unfasten itself, may function as the condition for the passive use of the verb.

Byali follows the same pattern insofar as the verb ‘to untie, to unfasten’ PFV *hūbārāsā* / IPFV *hūbārū* can be used passively, when the patient subject is animate, but not the verb ‘to tie’ PFV *hūbāsā* / IPFV *hūbū*.

- (28) *búúyā hūbāsā* ‘the dog has been unfastened (the dog has broken free)’

but:

- **búúyā hūbāsā* ‘the dog has been tied up’

With regard to verbs which do not allow voice diathesis alternation in Ditammari, no uniformity in the semantic roles of their participants was determined in Byali. Verbs which can take both animate and inanimate participants, e.g. ‘to buy, sell, lift up, take out, grasp, guard’ belong to that group as well as verbs that are exclusively linked with animate participants, as ‘to wash s.o.’ or verbs that go exclusively with inanimate participants, like ‘to plant’.

The fact that, in Byali the syntactic behaviour of some of the verbs mentioned deviates from that of these verbs in Ditammari, points to the already mentioned phenomenon of language specific constraints; e.g. the verb ‘to lift’ occurs in Byali passively with both kinds of participants, whereas ‘to take out’ is used in a passive construction only with an inanimate participant.

Syntactic constraints for the intransitive-passive use of verbs in Ditammari

In Ditammari, verbs belonging to the valency-flexible group do not follow one uniform syntactic behavioural pattern in their intransitive-passive use, but are part of one of the following two subgroups:

1. verbs which do not undergo any syntactic restriction in their passive-like use; (they occur in sentence-final as well as in non-sentence-final position.)
2. verbs which can be used in a passive sense, only if they occur in non-sentence-final position.

This condition has been demonstrated for Ditammari; how far it holds true for the other languages of the subgroup, among them for the Byali, still has to be investigated. The intransitive-passively used Byali verbs (cf. (22), (23)) can be positioned sentence-finally, an occurrence which is excluded for Ditammari verbs of the second group, (cf. (30)).

Ad 1: In the database, verbs belonging to the first subgroup are numerically fewer than those of the second subgroup. e

(29)	PFV	IPFV		PFV	IPFV
	<i>nētá</i>	<i>nṣṣ</i>	‘to split’	<i>dí</i>	<i>yō</i> [?] ‘to eat’
	<i>tṣṣá</i>	<i>tṣṣ</i>	‘to tear’	<i>kòṣṣā</i>	<i>kòṣṣ</i> ‘to sweep’
	<i>máá</i>	<i>māá</i>	‘to build’	<i>pwōtá</i>	<i>pwōtì</i> ‘to beat’
	<i>pwàntā</i>	<i>pūà</i> [?]	‘to destroy’	<i>bōtá</i>	<i>bwōtì</i> ‘to bite’
	<i>dèè</i>	<i>dōó</i>	‘to finish’	<i>kpētīrā</i>	<i>kpétì</i> ‘to close’
	<i>yūīrā</i>	<i>yūīí</i>	‘to stop up’	<i>pūtīrā</i>	<i>pūtí</i> ‘to tie’

Ad 2: It is a common feature of the Ditammari verbs of the second subgroup that they can be used in an intransitive-passive sense only when not in sentence-final position, i.e. they have to be followed by any element, be it an adverb, a phrase or a whole clause, in order to allow voice diathesis alternation.

Examples for verbs of this syntactically constrained subgroup are:

(30)	PFV	IPFV		PFV	IPFV
	<i>háá</i>	<i>hṣṣ</i>	‘to sharpen’	<i>bwṣṣí</i>	<i>bwṣṣì</i> ‘to seed’
	<i>pāāí</i>	<i>pāāì</i>	‘to plait’	<i>dwṣká</i>	<i>dwṣkù</i> ‘to write’
	<i>dwṣkā</i>	<i>dwṣkù</i>	‘to weave’	<i>kàà</i> [?]	<i>kàà</i> [?] ‘to count’
	<i>hòṣṣā</i>	<i>hóbú</i>	‘to wash’	<i>cṣṣ</i> [?]	<i>cṣṣmmū</i> ‘to burn’
	<i>kṣṣ</i>	<i>kṣṣmmū</i>	‘to fold’	<i>yōōrà</i>	<i>yōōrí</i> ‘to shake’
	<i>béntā</i>	<i>byē</i> [?]	‘to beat (drum)’	<i>pūtá</i>	<i>pūtì</i> ‘to spread out’
	<i>tí</i>	<i>tī</i>	‘to pound’	<i>bí</i>	<i>biití</i> ‘to plaster’
	<i>yá</i> [?]	<i>yàntì</i>	‘to sew’	<i>yètū</i>	<i>yèètì</i> ‘to light’
	<i>yà</i>	<i>ywà</i> [?]	‘to drink’	<i>kwā</i>	<i>kòṣṣ</i> ‘to kill’
	<i>tūtá</i>	<i>tú</i>	‘to dig’	<i>pēētá</i>	<i>pékí</i> ‘to winnow’

According to this syntactic condition for the passive use of these verbs, sentences like those under (31), are judged as ungrammatical, but when the same verbs are in non-sentence-final position as in the NP – V-constructions under (32), they are used intransitive-passively.

(31)	<i>*yā.yābòr.à</i>	<i>hòṣṣā</i>	‘the shirts have been washed’
	<i>*dī.kāṣkà.ne</i>	<i>dō bēntā</i>	‘the drum has been beaten’
	<i>*kū.fṣṣ.kù</i>	<i>bō tūtá</i>	‘the hole will be dug’

- (32) *yā.yābòr.à* *hòòtā* *ká-dōrí*
 CL.shirt.CL wash.PFV CONSEC-be clean
 ‘the shirts have been/are washed and are clean’
- ḍī.kàṇkà.nì* *bō* *béntā* *kā-yèn* *ká-mṣṣ*
 CL.drum.CL be beat.PFV CL.night CL.whole
 ‘the drum will be beaten’
- kū.fṣṣ.kù* *ḍō* *tūtá* *yénkā*
 CL.hole.CL PAST dig.PFV yesterday
 ‘the hole has been/was dug yesterday’

A passive meaning as given in the English translation of (31): ‘the shirts have been washed’, ‘the drum has been beaten’ etc. can be expressed in Ditammari only by an ‘indefinite passive construction’, the functional passive equivalent that is a known feature of many other languages :

- (33) *bāhòòtā yāyābòrā* the shirts have been washed = ‘they have washed the shirts’
bābéntā ḍīkàṇkà.nì the drum has been beaten = ‘they have beaten the drum’
bātūtá kūfṣṣkù the hole has been dug = ‘they have dug the hole’

Here, the pronoun of the 3RD P.PL. is used for a non-referential agent, while the real agent remains unexpressed for pragmatic reasons, and the patient retains its function as grammatical object. As a result, the verb is used actively.

Transitive-active verbs used intransitively

The hitherto discussed intransitive use of transitive verbs realized without any morphological marking of the verb has been seen to be due to voice diathesis alternation; thus, it was pragmatically motivated. Pragmatic reasons are also responsible for the suppression of the objects of some transitive-active verbs, above all those of eating and drinking, whose second participant is non-referential, i.e. not referring to a specific referent in the discourse.¹³

- (34) *tē* *ḍī* *kē* *ḍèè* ‘he has finished eating’
 CLS eat.PFV CONSEC finish.PFV

In the concrete context, it is not important for the hearer to know what has been eaten. It is only the fact of eating food that is of interest.

Another kind of morphologically not marked intransitive use of transitive-active verbs is also conditioned by a pragmatically motivated suppression of objects. Since in (35) the objects of each of the second verbs, i.e. of *tó* ‘carry’ and of *bōnte* ‘throw’, are referentially identical with those of the first verbs and in this way known from the context, the repetition of *ō pókù* ‘his wife’ and *ō kpāānī* ‘his lance’ is not relevant for the hearer:

(35)

kē pín ō-póku kē tɔ́ kē náké ō-kótì
 CONSEC grasp.PFV POSS-wife CONSEC carry.PFV CONSEC speak.PFV CL-old man
 ‘... and he grasped his wife and carried her and said to the old man ...’
kò ò-twóté ò-kpāā-nī kē bɔ́nté í-kpāā myēē-kē kē múnkē kwɔ́
 CON CL.took POSS-lance CON threw guinea fowl inside CON also killed
 ‘... and he took his lance and threw it into the guinea fowl and killed it.’

Such an intransitive use of transitive-active verbs is very often found in narrative texts.

It has to be emphasized that in these languages, when the semantic roles of the participants are not taken into account and only the formal structure, i.e. the order subject – verb (= structure NP-V) is considered, passive constructions, constructions with intransitively used transitive-active verbs as well as those containing intransitive-active verbs cannot be distinguished at all.

2.2.2 *Kaansa*

As already indicated, about two thirds of the transitive-active verbs in Kaansa can be classified as flexible verbs, i.e. they can be used passively if the second participant is deleted. In general, all that has been said above on the use of flexible verbs in Eastern Oti-Volta is also valid for Kaansa. Only a few differences need to be mentioned. These are due to language specific semantic concepts and - as far as constraints on occurrence with certain tense forms are concerned- they probably derive from the different tense-aspect systems.

2.2.2.1 *Passive-like constructions*

In general, voice diathesis alternation is found either with the perfect tense – to denote the state resulting from the event expressed by the verb – or with the unmarked present tense to indicate that the event is going on. The former seems to be the most preferred construction.

The perfect formation deviates from other tenses in that the verbal stem is moved to the initial position of the predicate so that it precedes the subject pronoun, which is then followed by the perfect marker.¹⁴ In addition, the assertive suffix /-ma/ is generally attached to declarative predicates.

Transitive-active use; Perfect tense

- (36) *bará-ŋv-r-ma¹⁵ du-k-ira* ‘s/he locked the house’
 lock-CL.S-PF-ASS house-CL-DEF

Present tense

- (37) *ɔ́ bar-ma du-k-ira* ‘s/he locks the house’
 CL.S lock-ASS house-CL-DEF

Intransitive-passive use; Perfect tense

- (38) *du-ki* *bara-ŋgi-r-ma* ‘the house has been locked (up)’
 house-CL lock-CL.S-PF-ASS
- tá-ŋgi* *bvv-gi-r-ma* ‘the hole is plugged up’
 hole-CL plug-CL.S-PF-ASS

Present tense

- (39) *du-ki-ra* *bar-ma* ‘the house is being locked up’
 house-CL-DEF lock-ASS

It should be emphasized that a small group of verbs (about 1/5 of the flexible verbs in our corpus) is used only in the perfect tense, probably motivated by their semantic feature [+terminative]:

- (40) *baʔ-ma* ‘to fasten’ *kɔʔ-mɔ* ‘to hold, to seize’
 bɛʔ-mɛ ‘to bury’ *khvv-ma* ‘to kill’
 bɛʔrɛ-mɛ ‘to level down (of iron)’ *ɲáa-ma* ‘to see, to find, to receive’
 daa-ma ‘to set on fire’ *sɛɛʔ-mɛ* ‘to say’

Among these are derivations from intransitive verbs, all of them showing slight semantic variation:

- (41) *kanaʔ-ma* ‘to install a fetiche’ ← *ka-mma* ‘to consult the divine’
 bukɪ-mo ‘to make a mistake/deceive’ ← *bur-mo* ‘to be lost, to disappear’

Transitive-active use

- bukɪ-mi-n-u* ‘I deceived him’
 deceive-1sg-PF-CL.O

Intransitive-passive use

- bukɪ-mi-ni-ma* ‘I have made a mistake’
 deceive-1sg-PF-ASS

As for the use of the present tense we note in some cases a non-ongoing and therefore potential reading as well (in particular, if an adverb follows), e.g.:

- (42) *dɔŋgi-mo* ‘to sell’ ← *dɔ-mmo* ‘to buy, to pay’
 ki *dɔŋgi-ma* *tákɪrsa* ‘it sells well’
 CL.S sell-ASS well

Of course, terminative verbs are excluded from this usage. Such verbs are preceded by the respective verb of modality:

- (43) *du-ki-ra* *wɔni* *bara* ‘the house can be locked’
 house-CL-DEF can-PF lock

tɪ wɔ̌-nɪ bɛv-ma 'it can be folded up'
CL.S can-PF fold-ASS

As in Eastern Oti-Volta, voice diathesis alternation does not take place when the patient is specified as being [+human]. The following example illustrates this constraint:

- (44) *kpónóʔ-mo* 'to respect' [+HUMAN], 'to enlarge' [-HUMAN]

Transitive-active use

kpónóʔ-u-r-uu-ma 'S/he has respected him'
respect-CL.S-PF-CL.O-ASS

Intransitive-passive use

(*duki*) *kpónó-ki-r-ma* 'It (the house) has been enlarged'
(house-CL)-enlarge-CL.S-PF-ASS

Nevertheless, and in contrast to Eastern Oti-Volta, there is one verb which can be used passively with a [+animate, -human] patient:

- (45) *kpɛ-ɛ khv-ʔv-r-ma* 'the goat is killed'
goat-CL kill-3S-PF-ASS

Note that this verb belongs to the group of terminative verbs, which do not form the passive with the present tense. Moreover, a few examples of verbs with [+human] patients are attested where diathesis alternation implies a change of verb meaning. As in the following case, the former patient moves to the subject position and assumes the role of experiencer.

- (46) *khqɪ-ma* 'to beat'

Transitive-active use

ʊ ma bii wá khqɪ-m-ɔ 's/he is beating the child'
CL.S be child CN beat-SF-LOC

Intransitive use

bii-ra ma khqɪ-m-ɔ 'the child is suffering'
child-CL-DEF be beat-SF-LOC
ʊ khqɪ-ma 's/he suffers'
CL.S beat-ASS

It goes without saying that indefinite passive constructions are also to be found in Kaansa, they sometimes even appear to be the preferred or even the only possible means of expressing an event.

2.2.2.2 Anti-passive constructions

As already indicated for Eastern Oti Volta, patient suppression may occur without a change in voice, i.e. the process remains active. Besides the well-known pragmatic phenomena which trigger such constructions (see Givón 2001:168), we may add a semantic parameter. Verbs of ingestion and food preparation are attested as preferable candidates for object suppression. This is also true for Eastern Oti Volta (see 2.2.1.4). Moreover, as Carlson pointed out for Supyire, the choice of tense is also an important factor in that “there is an affinity between imperfective aspect and subject suppression” (2000:51). Similar evidence is attested for Kɛansa where the progressive form is used in these cases.

This may be demonstrated with the verb ‘to eat’, which is classified as valency-flexible.

In an unmarked intransitive sentence in the present tense it may express either a passive or an active meaning:

- | | | | | | |
|------|-----------|--------------|----------------|-----------|------------------------------------|
| (47) | <i>si</i> | <i>di-ma</i> | | | ‘this is being eaten up/is edible’ |
| | CL.S | eat-ASS | | | |
| | <i>u</i> | <i>ma</i> | <i>dii-m-o</i> | | ‘s/he is eating’ |
| | CL.S | be | eat-SF-LOC | | |
| | <i>mi</i> | <i>ma</i> | <i>su-si</i> | <i>wá</i> | <i>di-m-o</i> |
| | 1sg | be | tô-CL | CN | eat-SF-LOC |

3. Change of valence by morphological marking

3.1 Eastern Oti-Volta languages

In the cases discussed so far change of valency was linked with an alternation between the transitive-active and the intransitive-passive meanings of the construction as well as with the deletion of the object for contextual and pragmatic reasons without any morphological marking. In all languages of the Eastern subgroup of Oti-Volta, however, derivation is another means by which the valency of a verb can be altered. Verbs which are assigned to the semantically defined verb pair inchoative/causative, are differentiated by being linked to either one or to two participants. The causative verb meaning includes an agent or a force participant who causes the situation, i.e. the event is externally caused, whereas the inchoative verb meaning excludes a causing agent.

A most interesting fact in this connection is the following bipartition: in some cases, it is the causative-transitive verb that represents the basic form from which the inchoative-intransitive is morphologically derived, e.g. in Ditammari:

(48)		PFV	IPFV	
	CAUSE-TRANS	<i>cɔɔʔ</i>	<i>cɔɔmmū</i>	'to burn'
	INCHO-INTRANS	<i>cɔɔtā¹⁶</i>	<i>cɔɔʔ</i>	'to burn'
	CAUSE-TRANS	<i>ɣiāʔ</i>	<i>ɣiāmmū</i>	'to melt'
	INCHO-INTRANS	<i>ɣiātā</i>	<i>ɣiārī</i>	'to melt'

In other cases however, the inchoative-intransitive variant is basic and the causative-transitive is morphologically marked:

(49)		PFV	IPFV	
	CAUSE-TRANS	<i>kwáʔ</i>	<i>kūāʔ</i>	'to weep'
	INCHO-INTRANS	<i>kwáná</i>	<i>kwánímū</i>	'to weep'
	CAUSE-TRANS	<i>kpáʔ</i>	<i>kpàʔ</i>	'to curse'
	INCHO-INTRANS	<i>kpānnà</i>	<i>kpānnà</i>	'to tell s.o. off'

From the morphological point of view, we are faced with different directions of derivation representing a process of detransitivization or of transitivity respectively. Haspelmath (1993) and Levin/Rappaport Hovav (1995) consider spontaneous occurrence of events as well as their external causation as factors that determine transitivity. Spontaneous occurrence of events is seen to be the opposite of external causation, so that non-spontaneous occurrences, i.e. externally caused events, are transitive.

The examples given under (50) seem to confirm the opinion that external causation by an agent in semantically and morphologically linked causative-inchoative verb pairs goes with a basic transitive verb; the intransitive verb is then morphologically marked, i.e. it is derived from the transitive form. The derivational suffix in the intransitive verb in (50) is an example of this markedness.

(50)	CAUSE-TRANSITIVE	<i>ōmáátà ɣiā n tīmátì</i>	'the smith has melted the iron'
	INCHOA.-INTRANS	<i>tīmátì ɣiātā</i>	'the iron is melted'

In the latter sentence – 'the iron is melted' – the same eventuality 'to melt' occurs without the direct intervention of an agent, so that the agent is not expressed in the syntax. In contrast to this, in spontaneously occurring eventualities it is the transitive verb that is marked and the intransitive occurs as the basic form. The following two sentences under (51) with the inchoative-intransitive verb pair 'to curse' versus causative-transitive 'to tell s.b. off' confirm this for Ditammari, too. The causative-transitive verb contains the derivational suffix.

(51)	INCHOAT-INTRANS	<i>qābírā kpáʔ</i>	'the child has cursed'
	CAUSE-TRANSITIVE	<i>cīta kpānnà qābírā</i>	'the father has told the child off'

The morphological shape of the causative or inchoative verb respectively seems to be determined by "the status of eventuality as externally or internally caused" (Levin/Rappaport Hovav 1995: 106). Accordingly, the different derivational directions

of causative/inchoative verb alternation illustrated for Ditammari are semantically conditioned.

3.2 Kɛansa

As to verbal derivation, Kɛansa has two operative extension suffixes which both have an impact on the valency of the verb.¹⁷ The first to be mentioned is the causative morpheme /-ʔra-/. This derivational function is attested in many Gur languages, and for this reason will be not dealt with in this paper. It changes intransitive verbs into transitive ones.

3.2.1 The verbal extension /-GEE-, -REE-/

Of special interest is the extension /-GEE-, -REE-/ which seems to be unique within Gur. It denotes patient suppression with transitive verbs. The agent remains in its topic position.

(52) *baʔ-ma* → *ba-kɛɛ-sɛ* ‘to fasten’

Transitive-active

ʊ *ba-kɪ-ma* ‘s/he fastens it’
CL.S fasten-CL.O-ASS

Intransitive-active

ʊ *ba-kɛɛ-ma* ‘s/he fastens’
CL.S fasten-GEE-ASS

cf. the intransitive-passive use of the underived form:

ba-kɪ-r-ma ‘it has been fastened’
fasten-CL.S-PF-ASS

The verbal extension /-GEE-, -REE-/ occurs only with the basic, i.e. the underived form of the verb. Apart from its syntactic function, on the semantic level we notice a plurality function in a wider sense, i.e. it renders habitual or repetitive meaning.¹⁸ Its etymological origin is far from being clear. The long vowel /-EE-/ brings the object pronoun of the plural class /-yV/ (singular -RV) to mind. But since the consonantal realization of the extension in /-GEE-/ or in /-REE-/ is not predictable, we are not able to refer it unambiguously to the extension /-Gɪ-/ which expresses plurality.

Despite the fact that the main function of the extension is to mark patient suppression with transitive-active verbs, it is also found with a few intransitive verbs. The meaning of these extended forms confirms the supposed underlying function to express habitual or repetitive actions:

- (53) *fji-me* 'to sprout' → *fji-nee-se* 'to grow (of hair)'
kpáa-ma 'to be silent' → *kpáa-rée-se* 'to take a rest'
kpá-mma 'to tremble' → *kpá-née-se* 'to show joy'

As for the transitive verbs, the syntactic function of patient suppression seems to be more dominant than the above mentioned semantic feature, e.g.:

- (54) *beʔ-me* → *beʔ-ree-se* 'to farm animals'
u béebeʔma súmbo 's/he wants to farm poultry'
u tu beʔreema 'she knows how to farm (animals)'

In a few cases we notice semantic change:

- (55) *kpaʔ-ma* 'to cut' → *kpaʔ-rée-se* 'to harvest yam'
seeʔ-me 'to say' → *see-rée-se* 'to speak'
sv-mma 'to send, to cultivate' → *sv-née-se* 'to cultivate, to work'
suʔ-mo 'to pound (yam)' → *suʔ-ree-me* 'to pound (millet)'

Regarding the nominalizing suffix */-sV/*, we should note that if this suffix occurs in nominalized form, it never takes a complement.¹⁹ In such cases it expresses an intransitive inchoative meaning, whereas the progressive construction with the nominalizing suffix */-mV/* obligatorily takes an object representing its transitive-causative counterpart:

- (56) *tɪ ma dáʔɪnsɔ* this is melting
yírera ma ú khɪgɪ wá dáʔɪmɔ the blacksmith is melting his iron

If the speaker wishes to express the patient, i.e. if he adds the notionally obligatory object to the respective verb, he is obliged to use the underived simple verb form. This interplay between underived and derived forms of the verb is shown in the following examples:

- (57) *sv-mma* 'to send, to cultivate' → *sv-née-se* 'to cultivate, to work'

The underived verbal form occurs with its obligatory direct object:

<i>sv-mbɪ-r-ma</i>	<i>biira</i>	'they have sent the child/'
send-CL.S-PF-ASS	child-CL-DEF	'the child has been sent by them'
<i>sv-mbɪ-r-ma</i>	<i>susv-ga</i>	'they have cultivated maize'
cultivate-CL.S-PF-ASS	maize-CL	

The derived verbal form without an object expresses only 'to cultivate' or generally 'to work':

<i>u sv-née-ma</i>	<i>tákɪrsa</i>	'he cultivates well'
CL.S cultiv.-REE-ASS	well	
<i>u sv-née-ma</i>	<i>khangɪ-rá</i>	'he cultivates on the field'
CL.S cultiv.-REE-ASS	field-LOC	

In contrast, the extended form /-REE-/ can be used if the second participant is benefactive:

sɔ-nɛʔv-r-vv-ma 'he has cultivated for him'
cultiv.-REE-CL.S-PF-CL.O-ASS

The case of 'to pound yam/millet' (see last example in the table above) is exceptional insofar as the derived form with /-REE-/ may occur either with the patient participant – then denoting 'to pound millet' or – as usual – without a patient. In the latter case its meaning is 'to pound yam'.

With the patient 'millet':

- (58) *u suʔ-ree-ma svv-ya* 'she pounds millet (habitually)'
CL.S pound-REE-ASS millet-CL
- v ma sv-yɪ wá suʔ-ree-m-o* 'she is pounding millet'
CL.S be mil-CL CN pound-REE-SF-LOC
- v ma sv suʔ-ree-g-o* 'she is millet pounding'
CL.S be mil pound-REE-SF-LOC

This exceptional use can only be explained by the semantic specification expressed by the object (pounding yam vs. pounding millet). In all other cases, may it be used actively or in passive-like constructions, the underived verb 'to pound' always refers to 'yam':

- (59) *thaa-ɲgɪ su-ki-ri sód* 'the yam are already pounded'
yams-CL pound-CL.S-PF finish
- mɪ ma suʔ-m-o* 'I am pounding (yam)'
1sg be pound-SF-LOC

3.2.2 Object incorporation

This very common type of anti-passivization (see Givón 2001: 169) occurs in Kaansa with only one nominalized verbal form, which is marked by the suffix /-KV/. The most attested use is that in progressive constructions. In contrast to the progressive form characterized by the suffix /-mV/, where the obligatory patient precedes the nominalized verbal form in its full specified form,²⁰ in this type of progressive constructions the nominal object is not referential, i.e. it occurs in its radical form without class markers, e.g.:

Object incorporation

- (60) *v ma khɪ dǎfɪŋgɔ* 'he is iron melting, i.e. he does it habitually'
v ma súm beko 'he is poultry farming, i.e. he is a poultry farmer'

Progressive

- (61) *v ma khɪŋɪ wá dáʔɪmɔ* ‘he is melting the iron’
v ma súmbi wá beʔmo ‘he is farming the poultry’

In summing up our findings, we can state that in Kàansa four different types of forms of transitive-active verbs with the feature [flexible] may occur in one and the same intransitive sentence pattern NP–V. Each indicating a modification of the perspective on the event expressed in the verb, they are distinguished by the parameter ‘orientation’, by the morphological marking of the verb and/or by the semantic specification of the respective participant in topic position. The first type with this NP–V pattern, indicating a passive-like perspective, is patient-oriented whereby the form of the verb is not affected. This type is also attested for Eastern Oti-Volta and other Gur languages.

In contrast, the second type remains agent-oriented and the extension suffix /-REE-, -GEE-/ is added to the verb. It denotes anti-passive-like constructions and is also used with non-flexible verbs. This type is unknown in other Gur languages.

In addition, Kàansa forms anti-passive constructions either by suppression of the patient participant or by incorporation of the object. Both these strategies are also to be found with non-flexible verbs. The former type of construction appears to be restricted to the semantic field of ingestion and preparing food. It is also attested in other Gur languages. The latter, which only occurs in progressive constructions, has to our knowledge not yet been attested elsewhere within Gur.

The following table schematizes the four types:

Subject	Verb	Object	Perspective
PATIENT [-ANIM] ²¹	[flexible] unchanged with regard to transitive-active use	suppressed	passive
AGENT [+HUMAN]	+ extension [-REE-/GEE]	suppressed	anti-passive
AGENT [+HUMAN]	unchanged (only with verbs of ingestion)	suppressed	anti-passive
AGENT [+HUMAN]	nominalized	non-referential	anti-passive

4. Conclusion

In this paper, we have tried to show the different strategies developed in quite different Gur languages to express voice diathesis alternation. Our attention has been focussed on the behaviour of transitive-active (or bivalent) verbs in intransitive sentence patterns. The analysis has clearly shown that in each language most such verbs may occur in both NP–V–NP and NP–V sentence patterns without further morphological marking. This capacity appears to be a salient feature dividing the transitive-active verbs into two different classes which we refer to as ‘flexible’ or ‘non-flexible’ with regard to

their valency. We are of the opinion that this feature would serve a valuable purpose in descriptions of this language type, and should at least be marked in the dictionary.

Data from other Gur languages substantiate the evidence presented in our paper. The same phenomenon is also documented in descriptions of other Gur languages, as in Nicole (2000) for Nawdm (a branch of Oti-Volta) but not treated as a special issue. Moreover, in Nateni and Waama (Eastern Oti-Volta) as well as J̄aneand Lobiri (South-Central Gur)²² we note the same ‘valency-flexibility’ of transitive-active verbs as is reported in our paper, but of course with different constraints concerning their realisations in intransitive patterns. These different constraints – apart from the fact that there are above all pragmatic reasons for using passive-like constructions – are due to the very individual shaping of verb properties in the respective languages, in particular the semantics of the verb itself as well those of the required participants. The same is true for Moore, a western Oti-Volta language.²³ Similar evidence is reported by Carlson (1997, 2000) for Supyire, and two other Senufo languages.²⁴ He drew our attention to this phenomenon, which he calls ‘verbal lability’ (see 2) and which he considers the ‘Good Trick’ against the background of relatively poor morphology. He writes: “Actually, the Good Trick can be reconstructed for Proto-Senufo, and moreover seems to be such a Good Trick that it is in fact an areal feature, being present in varying degrees in other languages in the area, both Gur and Mande” (2000:57).

All the languages require, more or less obligatorily, the second argument in transitive clauses, – in our view, one of the prerequisites for this kind of construction. This condition provided the space for using the existence or non-existence of the second argument as a functional tool to express modification in the perspective of the event rendered in the sentence, in other words, to provide a tool for expressing variation in voice. By this ‘new’ property, the number of verbs which can be used in intransitive patterns has been increased enormously. Carlson (2000:49) speaks in the case of Supyire of the, ‘vast majority of verbs’ which may occur in intransitive clauses. This leads us to the assumption that the passive-like perspective linked with this clause-type enjoys some attraction among the speakers of the languages dealt with. We do not mean to suggest that this is the preferred perspective, as one may say of ergative languages, but there does seem to be a strong tendency in that direction.

Nevertheless, it should be noted that according to our data apparently none of the languages treated in our paper has developed a specific construction to express the agent or the causer of such passive-like constructions as it is possible for example in Kabiye²⁵ and in general in the Mande languages.

Concerning the wider regional context we find that in three language families – Mande, Gur, and Kwa – the same constraint is employed: the second argument in transitive clauses is obligatory. However, only Gur and Mande make functional use of this.

Abbreviations

ANIM	animate	FUT	future
ASS	assertive	IPFV	imperfective
CAUS	causative	LOC	locative
CL	class marker (nominal)	NP	noun phrase
CL.O	class marker (object pronoun)	P	person
CL.S	class marker (subject pronoun)	PF	perfective
CN	connective marker in genitive construction	PL	plural
CONSEC	consecutive	POSS	possessive
DEF	definite	REE	-GEE/-REE verbal extensions
DUR	durative	SF	suffix (verbal noun)
		SG	singular

Notes

1. In the sense of Bußmann (1990:182), who defines the term 'diathesis' as follows: "Genus verbi ... used as well as for other regular changes of valency like applicative, accusativization, dativization etc. [Genus verbi ... und für andere reguläre Valenzrahmenwechsel wie Applikativ, Akkusativierung, Dativierung u.a.]".
2. The term 'object' is used always in the sense of 'direct object'.
3. Lébikaza (1998) treating similar phenomena in Kabiye (Eastern Gurunsi) argues in almost the same way. In his opinion, it is the relation between the semantic roles of the arguments as well as the interaction between the semantic roles and their syntactic functions that determine the behaviour of verbs. Concerning this interrelationship he emphasizes the 'prééminence, équivalence ou symétrie des arguments sémantiques' (1998:75).
4. Haspelmath describes labiality conditions for inchoative/causative verb pairs in many languages, for example for verbs like 'to break, to burn, to open, to begin, to finish' etc., where the inchoative verbs represent the intransitive use and the causative ones their transitive counterparts.
5. One has to distinguish between so-called 'long' and 'short' forms, each of them being linked to special functions. The long form occurs in only three conjugational paradigms: the imperative, the aorist, and the negation of the perfect. Besides this, it is attested in each conjugational paradigm in those cases where an object pronoun consisting only of a vowel immediately follows the verb. Moreover, a kind of participle is based on the long form of the verb.
6. A possible etymological explanation may be that when writing was introduced during colonial times, ink and pen-holder were the usual requisites.
7. Derived from: *dq-mmo* 'to pay', 'to buy' cp. also: *dɔŋgi-mo* 'to sell'. Both belong to the group of flexible verbs as in Byali, but not in Ditammari.
8. The examples given are elicited. But the transitive-active or intransitive-passive use of the verbs mentioned is confirmed by their respective occurrence in texts.
9. The progressive form corresponds to the widespread locational pattern of progressive source constructions insofar as it is formed with the locative verb of being '*bo*' + a non-finite verbal form.

10. As in its active use, the progressive construction of Ditammari can be combined with both tense markers indicating the past (*do*) and the future (*bo*).
11. If *pītīrā* 'tie' was linked to an object, *tāmšīà* 'dog' would bear the semantic role of an agent.
12. This event can also be expressed with the aid of the third person plural pronoun, i.e. by the indefinite passive construction, which seems to be the preferred variant in such cases.
13. Van Valin/La Polla (1997:123) call such a non-referential argument *inherent argument*, i.e. "an argument which expresses an intrinsic facet of the meaning of the verb..."
14. In general, the anaphoric subject pronoun is omitted, if a nominal subject precedes the predicate. In contrast, in perfect tense constructions the anaphoric pronoun remains in its postverbal position, if introduced by a nominal subject, e.g. *khoo barā-ngv-r-ma du-k-ira* 'the man locked the house'.
15. In the perfect tense, this verb shows the extension /-a-/. Its function still has to be investigated.
16. The morpheme /t/, which serves here in deriving the inchoative from the causative verb occurs with the same function, when an inchoative meaning is derived from a stative one.

	PFV	IPFV	
STATIVE-INTRANS.		<i>čīnī</i>	to be weak
INCHOAT.-INTRANS.	<i>čītā</i>	<i>čītīrī</i>	to become weak
CAUS.-TRANSITIVE	<i>čīkūrā</i>	<i>čīkuri</i>	to weaken

But in a such a derivational series it is first of all the morpheme /k/ which takes the inchoative function:

STATIVE-INTRANS.	—	<i>pēè</i>	to be white
INCHOAT.-INTRANS.	<i>pēēkā</i>	<i>pēēkú</i>	to become white
CAUS.-TRANSITIVE	<i>pēēkūrā</i>	<i>pēēkūrī</i>	to make white

For the sake of completeness it should be mentioned that in all these languages a causative verb often derives from a stative verb taking the intermediate stage of an inchoative verb. In such a derivational series the causative nasal /n/ is realized as lateral [r] after the inchoative derivational morpheme /k/.

17. A third operative extension (-*G₁* -) is also to be found, which denotes plurality in a broader sense.
18. In accordance with that, its formation is based on the short form of the verb which renders basically imperfective meaning.
19. In general, this suffix is only found with derived verbal forms. Nevertheless, there are some underivable verbal stems which resemble the extended form but for which no simple verbal base is available.
20. Kąansa – like all other Gur languages – has the order modifier-modified in genitive constructions.
21. With the exception of a few [+animate, -human] nouns.
22. For all these languages we have collected respective field material.
23. Personal communication from Norbert Nikiema.
24. The Senufo languages form a distinct branch of Central Gur. In contrast to the other Gur languages, they have a S O P word order.

25. Cf. Lébikaza (1998:72/73).

References

- Bußmann, H. (1990). *Lexikon der Sprachwissenschaft*. Stuttgart: Alfred Kröner.
- Carlson, R. (1997). The Senufo languages. *Gur papers/Cahiers voltaïques*, 2, 23–42.
- — —. (2000). Event-views and transitivity in the Supyire verbal system. *Gur papers/Cahiers voltaïques*, 5, 39–58.
- Givón, Talmy (2001). *Syntax. An introduction*, Vol. I–II. Amsterdam: John Benjamins.
- Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Haspelmath, M. (1993). More on the typology of inchoative/causative verb alternations. In B. Comrie & M. Polinsky (Eds.), *Causatives and Transitivity* (pp. 87–120). Amsterdam: John Benjamins.
- Lébikaza, K. K. (1998). Les verbes à arguments prééminents et arguments symétriques, et la forme stimuli-passive en kabiyè. *Gur papers/Cahiers voltaïques*, 3, 63–76.
- Levin, B. (1993). *English Verb Classes and Alternations*. Chicago, IL: University of Chicago Press.
- Levin, B., & Rappaport Hovav, M. (1995). *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: The MIT Press.
- Nicole, J. (2000). La chèvre ne me mange pas bien. Syntaxe et discours dans la phrase simple en nawdm. *Gur papers/Cahiers voltaïques*, 5, 115–121.
- Voeltz, E. F. K. (1977). *Proto-Niger-Congo Verb Extensions*, Ph.D. Dissertation, University of California, Los Angeles.
- Van Valin R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.

Structure and function of incorporation processes in compounding

Claudia Maria Riehl and Christa Kilian-Hatz

Universität zu Köln

Noun-incorporation seems to be a very common strategy used in compounding. It will be shown in this paper that we are dealing here with a process that is not only common in polysynthetic languages, but in languages in general. The function of this process is not only a semantic one, i.e. narrowing the scope of the predication, but also – and perhaps mainly – a syntactic one, i.e. backgrounding the given information. Whereas polysynthetic languages use noun incorporation as a strategy to background nouns in otherwise foregrounded sentences, rather analytic languages use it to background the whole predication by nominalization of the verb-noun complex.

1. Introduction

Noun incorporation is a process that is mainly discussed as a typical feature of polysynthetic languages. Although parallels in word formation patterns of other language types have already been pointed out (cf. Baker 1988:78ff.; Whaley 1997:131; Spencer 2000:314ff.), these observations rather stress the formal difference than the common features of the respective constructions. This is due to the fact that they mainly concentrate on the structural properties of noun incorporation processes and thus neglect the functional properties of these constructions.

We wish to claim, however, that noun incorporation is a rather universal phenomenon used in many languages for *discourse-pragmatic* reasons. Noun incorporation is a syntactic device that can – like all type of syntactic constructions – undergo lexicalization. By doing this the verb-noun complex becomes a so-called ‘sentence word’, e.g. *forget-me-not*. Constructions like this in which “syntactic expressions are reanalyzed as words”, represent, as Spencer (2000:316) remarks, “a particularly interesting, if poorly studied, type of interaction between morphology and syntax.”

In order to gain further insight into this kind of interaction of syntactic and morphological processes and to understand the functional motivation of incorporation this paper aims to show differences and similarities of formation patterns, correlations between formal pattern and content, and the possibilities to express syntactic and

morphological relations within syntactic words. Starting from these structural and semantic properties we will then examine the discourse function of noun incorporation within compounds.

To demonstrate these phenomena we will mainly take examples from two unrelated African languages, from the Khoisan (Kxoe) and Niger-Congo language family (Baka), as well as from Romance (French, Italian, Spanish) and Germanic languages (German and English).¹

The question what kind of words or phrases are to be subsumed by the term ‘sentence word’ is not easy to answer. In a classical definition this term covers expressions like German *Hans-Guck-in-die-Luft*, English *Forget-me-not* or French *rendezvous* (Fleischer & Barz 1995; Bauer 1996, et al.). Besides those lexicalised items there are a number of ad-hoc-constructions used mostly in newspaper language and poetry, e.g. *Wir-engagieren-uns-für-Köln-Initiative* (‘we-engage-ourselves-for-Cologne-initiative’; cf. Lawrenz 1997:121). Similar patterns of this kind of ‘sentence-words’ can be found in the lexicalised Baka expression for ‘West’:

- (1) *gàje kè bàkò á làti mò nè*
 side DEM sun 3.SG.PFV lie down where REL
 ‘West’ (lit. ‘side where the sun lies down’)

Or in Kxoe, where sentence-like words denote numerals over ‘five’ such as:

- (2) *cèú- hè è #tǒ- na-xao- kò cí- lúí- (xa)*
 hand 3.F.SG o put in II COMP CONV reach one GER
 ‘six’ (lit. ‘put in a full hand and come to one’)

In all these cases we observe a process where a syntactic entity (sentence, phrase) is transferred into a morphological entity, i.e. a noun. This can be regarded as a typical process of *incorporation* (cf. Eichinger 2000:31ff.), i.e. a noun stem is combined with a verb to form a new complex entity. In a classical definition the result of the incorporation process is a verb (Mithun 2000:916 citing Sapir 1911), whereas in our cases the complex entity is a noun.² It is, however, difficult to say, where the process of incorporation in word formation starts: Strictly speaking incorporation is already involved, when a simple syntagma of a noun and a genitive attribute become unified as in *car window* or when an agentive noun is derived from a verbal stem and the agent is expressed by an affix like *bak(e)-er* or Fr. *visit(er)-eur*.

In this paper we will concentrate on instances of incorporation that express complex syntactic relations, i.e. agent, verb, and an additional complement, i.e. expressions that have a phrase-like internal structure as the so-called *verb-complement-compounds* (Lloyd 1968) or *constructional compounds*, as for instance *hairdryer* or *cigarette-smoker*, where a complex predication (‘something dries hair’, ‘someone smokes cigarettes’) is expressed in a single word.

The paper is divided into five parts: In the first part we will analyze the general structure of the compounds in the different languages. In the second part we will show what kinds of semantic fields are expressed by these compounds and how they in-

terrelate with different types of word formation. In Section 3 we will examine what kind of semantic roles are expressed within the internal syntactic structure of the words and how these roles are marked on the surface structure. Section 4 outlines the functions of incorporation processes carried out by constructional compounds. In Section 5 finally we will summarize the results by placing the structures of the expressions concerned on a scale that illustrates the transition from pure syntactic to pure morphological features.

2. Structure of the compound

By comparing the structural principles of composition in different languages we will examine whether those patterns of word formation are consistent with the typological syntactic structure of the respective language.

The typical structure of the constructional compounds in Germanic languages is the following:

Comp + *V*_{stem} + *Suffix*
 Germ. *x macht Spaß* (lit. 'x makes fun') → *Spaßmacher* ('joker')
 Engl. *x makes trouble* → *troublemaker*

The same principle holds also for Slavic languages, i.e.

Comp + *INFIX* + *V*_{stem} + *Suffix*
 Russ. *x stroit machinie* (lit. 'x builds machines') →
*machinostroitel*³ ('machine-builder')

In Romance languages we have a different pattern:

*V*_{IMP/3.SG.} + *Comp*⁴
 Fr. *x garde les malades* (lit. 'x watches the sick') →
garde-malade ('watch-sick', i.e. 'nurse')
 It. *x cava i tappi* (lit. 'x pulls out the corks') →
cavatappi ('pullout corks', i.e. 'cork-screw')
 Span. *x corta los cigarros* (lit. 'x cuts the cigars') →
*cortacigarros*⁵ ('cut-cigars', i.e. 'cigar-cutter')

Here the respective complement is sometimes in the plural form (cf. It. *tapi*, Span. *cigarros*). This poses structural problems for the morphological system: The plural can no longer be marked by the inflectional suffix, but only by the definite article: *il cavatappi* – *i cavatappi*, *el cortacigarros* – *los cortacigarros*.

In Kxoe two constructions can be found, whereby the more productive type in (a) is completely identical with the pattern of Germanic and Slavic languages. In the less productive one (b) the suffix is lacking.

- (a) *Comp* + *V_{stem}* + *Suffix*
ávuru x wòó or *x ávuru wòó* ('x has/finds goods') →
ávuru-wòó-khòè ('rich man')

The suffixes *-khòè* (denoting an animate agent) and *-xù* (denoting an inanimate agent), however, unlike the corresponding Germanic *-er*, are noun-like entities that can be considered as classifiers. The difference between this classifier ('person', 'thing', 'being') and the corresponding full noun is marked by a difference in tone: Thus the full nouns each have a high tone *khóé* and *xó/xú* whereas the suffixes, *-khòè* and *-xù*, have a low tone.

- (b) *Comp* + *V_{stem}*
ápa x kxò or *x ápa kx'ó* (lit. 'x eats dogs') →
ápa-kx'ó ('dog-eater', i.e. 'Bantu')

In Baka, however, the Romance pattern can be observed. Three constructions are found:

- (a) *Prefix* + *V_{PART/INF}* + *Comp*
x kpù njèngè (lit. 'x catches fish') → *wà-kpù-njèngè* ('fisherman')
- (b) *V_{INF}* + *Comp*
x ja kəkə (lit. 'x catches chicken') → *jáà-kəkə* ('sparrowhawk')

From a synchronic perspective these compounds are nominalized forms of the verb, but from a diachronic point of view they are compounds: They are derived from complex verbs consisting of a verb and a complement (e.g. *na màà bo* 'to cure people', nominalized as in *wà-na-màà-bo* / *wà-máà-bo*). The verb-complement structure is still transparent in less productive constructions as in (c):

- (c) *V_{stem}* + *Comp*
x ye kà (lit. 'x likes wounds') → *ye-kà* ('s.b. who is always harmed')

In contrast to Romance languages, the complement is always in the singular, expressing a more generic concept, while the plural suffix *-o* – when attached at the end of the phrase, i.e. to the complement – always indicates the pluralization of the whole compound, e.g. in *ye-kào* ('the ones who are always harmed') or *wà-kpù-njèngèò* ('fishermen').

From a typological point of view it is interesting that Germanic languages – which are to a high extent VO-languages (thus typologically mixed languages, cf. Askedal 1996) – show a similar pattern as Kxoe, a typical OV-language. This can be explained by the assumption that constructional compounds behave like other compounds. In German, English and Kxoe the determinants (adjuncts) are placed in front of the determinant (head), whereas in Romance languages and in Baka the adjuncts follows the head. Finally, the complements in the constructional compounds behave syntactically like adjuncts and no longer like verbal complements. Thus they follow a nominal syntax rather than a verbal syntax. That is the reason why in some cases it is not

clear whether the compound is a constructional or a determinative compound, e.g. is the compound *car-driver* 's.b. who drives a car', i.e. in a constructional compound like *hairdryer*, or 's.b. who is the driver of a car', i.e. a determinative compound like *house-door*. In cases where the deverbal element also exists as a simplex (cf. *driver*) this question is very difficult to answer (cf. Bauer 1996:202). Before, however, we try to find an explanation for that question, we first take a look at what kind of semantic areas are covered by the compounds mentioned above.

3. Semantic fields

Almost exclusively, constructional compounds are found in the following five semantic areas.

3.1 Agentive nouns (*Nomina agentis*)

This is the most productive domain, and we find examples in all languages of our sample:

The noun can denote someone who is following

- (a) a profession as in Engl. *bar-keeper*, Germ. *Dachdecker* ('roofer'), It. *portalettere* ('postman'), Fr. *garde-malade* ('nurse'), Kxoe *ngú-n#óh-khòè* ('builder of a house'), Baka *wà-máà-bo* ('doctor')⁶
- (b) a habit as in Engl. *day sleeper*, Germ. *Frühaufsteher* ('early riser'), Fr. *gagne-petit* ('s.b. who earns few money'), Span. *cantaclaro* ('s.b. not mincing one's words'; lit. 'clear singer'), Baka *wà-tóò-bè* ('song-leader') or Kxoe *àvuru-woó-khò-mà* ('rich person')
- (c) an occasional activity as in Engl. *peace-maker*, Germ. *Spaßmacher* ('joker'), Fr. *monte-en-l'air* ('burglar') and Baka *wà-gbégbè-bo* ('host', 'enemy')

3.2 Instruments (*Nomina instrumenti*)

This kind of compound denotes an instrument with which one can do a certain activity: cf. Engl. *stain remover*, Germ. *Sockenhalter* ('sock suspender'), Fr. *porte-parapluies* ('umbrella stand'), It. *tergicristallo* ('windscreen-wiper'), Span. *reposapiés* ('foot stool') or Kxoe *máke-tc'éri-o-xò* ('ashtray'; lit. 'tobacco-extinguish-in-thing'), *#xéi-mfy-kà-xò* ('glasses'; lit. 'eye-see-with-thing'), or *ngú-xàù-kà-xò* ('strings for sewing a house-mat'; lit. 'house-sew-with-thing').⁷

3.3 Places (*Nomina loci*)

The compound denotes a place where a certain activity is carried out: This semantic implication is rather rare and not found either in Germanic languages nor in

Kxoe; in Romance languages and Baka only few examples such as Fr. *guarderobe*, Span. *guardarropa* ('cloak-room') or It. *marciapiede* ('pavement'), and Baka *ndándà-na-nǝlǝ-a-bo* ('pavement'; lit. 'place-INF-walk-people') or *nda-na-bonà-'èe* ('shop'; lit. 'house-INF-sell-thing') can be found.

3.4 Plant names and animal names

The naming of plants or animals by constructional compounds is already an old procedure in all languages as shown by the examples in Kxoe and Baka, where the respective formation pattern is not productive anymore: cf. Engl. *grasshopper*, Germ. *Strandläufer* ('sand piper'), Fr. *tournesol* ('sunflower'), It. *bocaneve* ('snowdrop'), Kxoe *dú-lxéú* (not determined plant species; lit. 'eland-ruminate'), Baka *jáà-kǝkǝ* ('sparrow hawk'; lit. take.PART/INF-chicken') and *gbǝ-ngǝmbi* ('mantis'; lit. 'beat-guitar').

3.5 Contemptuous names

Also very common in different languages is the expression of contemptuous names or joking expressions denoting other people.

(a) *ethnic names*

Pejorative expressions for people of another ethnic group are probably created in analogy to plant and animal names; cf. Germ. *Spaghettifresser* ('Italian'; lit. 'spaghetti eater'), It. *mangia-patate* ('German'; lit. 'potato eater'), Kxoe *ápà-kx'ó* ('Bantu'; lit. 'dog eater') or *lxúni-kx'ó* ('Bantu'; lit. 'crocodile eater'), and Baka *bée-lo* ('Bantu'; lit. 'carry wood').

(b) *contemptuous expressions*

Those are expressions like Engl. *arse-licker* and its translation equivalent in all the other languages analyzed here (Germ. *Arschkriecher*, Fr. *léche-cul*, It. *lecca-piedi*, Span. *lame-culos*). This kind of contemptuous expressions are very common in Spanish (cf. Lloyd 1968).

(c) *joking denotations*

In German we recently found a number of spontaneous pejorative constructional compounds that denote timid persons or weaklings: cf. *Cabriogeschlossenfahrer* ('s.b. driving a convertible with a roof closed'), *Leberkäskaltesser* ('s.b. who eats cold meat loaf'), *Schattenparker* ('s.b. who parks in the shade'), *Bei-Mama-Wohner* ('s.b. who still lives at his mother's house'), *Teletubbie-Winker* ('s.b. who greets the teletubbies [characters which appear in a popular children's program]'). These expressions are ad-hoc-constructions and only few of them are lexicalised already (like *Warmduscher* 's.b. who takes only hot showers'). But by analyzing the formation patterns of these

expressions we gain interesting insight into principles and constraints of this kind of word-formation processes and on their affinity to syntax.⁸

When considering the semantic fields covered by constructional compounds we find no significant difference between the languages of our sample: Only the domain of place names is not covered by all languages.⁹ This can be explained by a universal cognitive rule, i.e. a semantic shift from persons to animals and instruments by means of personification. A further step is the semantic shift to places, whereby places are perceived as containers keeping things (Lakoff/Johnson 1980).

The only difference between the European and the African languages analyzed here concerns the structure of the compound: In Germanic, Romance and Slavic languages exactly the same formation pattern holds for all semantic areas that can be expressed by the compound: the same suffix is used for all categories. The differentiation then is to be explained only by semantic shift.¹⁰ Kxoe also uses the same pattern for all semantic domains, but varies the 'suffix' (classifier): Agentive nouns are marked by the classifier *-khòè* 'person', *nomina instrumenti* by the classifier *-xò* 'thing'. As we can see from the differentiation in tone mentioned above, the 'classifier' is on the midway between full noun and suffix, but semantically still transparent.

In Baka there is even a difference between the formation of *nomina agentis*, where the use of the prefix *wà-* is obligatory, and plant-, animal- and pejorative names, which do not have a prefix. Moreover, compounds can be formed with the pure verbal stem without infinitival suffix. This phenomenon might be explained by the fact that words in these semantic fields are often part of a vocabulary of a secret language, which may differ from the normal lexicon not only on the semantic and phonological level but also in the formation of constructional compounds.

4. Semantic roles

In this section, the syntactic relations expressed by the constituents of the compound words will be examined in more detail. The following semantic roles can be encoded by the respective complements or other constituents (such as affixes).

4.1 Agent

In Germanic and Slavic languages as well as in Kxoe the agent usually is expressed by a suffix, i.e. *-er*, *-el'* etc., or by a classifier *-khòè* and *-xò*. In Romance languages and in some of the cases in Baka (examples without the prefix) it is not marked by any morphological device. In very few cases the agent is explicitly expressed by the complement: The agent is a word initial noun, e.g. in

Kxoe *ʔxéi-mfy-kà-xò* ('glasses'; lit. 'eye-see-with-thing'),
 It. *marciapiede* ('pavement'; lit. 'walk.foot') or *batticuore* ('heartbeat'; lit. 'beat.heart'), and Span. *reposapiés*¹¹ ('footstool'; lit. 'repose.foot').

4.2 Patient

In the vast majority of cases the complement encodes the patient role:

Germ. *Gepäckträger*: x carries the luggage (OV-order)

Engl. *cigarette-smoker*: x smokes cigarettes (OV-order)

Fr. *garde-malade*: x watches sick persons (VO-order)

It. *portalettere*: x carries the letters (VO-order)

Span. *cortacigarros*: x cuts cigarres (VO-order)

Kxoe *ápa-kx'ó*: x eats dogs (OV-order)

Baka *wà-mód-yà*: x hunts elephants (VO-order)

4.3 Benefactive (Dative)

The encoding of the semantic role of benefactive is very marginal in constructional compounds, according to Rivet (1999:308) it is not encoded at all. But we have found examples from the recent German joking constructions, cf.:

Teletubbie-Winker:

teletubbie(.DAT)-wave.AG

'teletubbie-waver' ('x waves the teletubbies').

Although the complement has a benefactive role, this is not encoded by any morphological device (Ø-case-marker).

Neither Baka nor Kxoe have compounds with transitive verbs where the verb only governs an indirect object; in cases where the verb (e.g. *give* or *say*) is ditransitive the complement has to be interpreted as direct object (s. Rivet 1999:317) as in German *Geldgeber* ('sponsor'; lit. 'money-giver') or *Jasager* ('s.b. committing himself to anything'; lit. '“yes” say-er').

4.4 Locative (time)

In a few cases the complement can also express a place as in the following examples:

Engl. *city dweller*: x dwells in a city, *sunbather*: x bathes in the sun

Germ. *Schattenparker*: x parks in the shade

Span. *trotacalles*: x walks on the streets (= 'streetwalker')

But there are cases where a locative extension appears instead of a pure noun:

Germ. *Bei-Mama-Wohner*: x lives at his mother's house¹²

Fr. *monte-en-l'air*: x climbs up into the air (= 'burglar')

It. *saltimbanco*: x jumps onto the bench (= 'ropedancer')

In contrast to the European languages analyzed here, the semantic role of place in Kxoe is expressed by the insertion of the verbal derivational affix *-o* which assigns a location

status to the classifier (but not to the complement) as in *máke-tc'érí-o-xò*: 'thing where x extinguishes tobacco in it' (= 'ashtray').

Time is expressed in very few cases, like in the Engl. formation *day sleeper*: 'x sleeps during the day' and in Engl./Germ. *Tagträumer* ('daydreamer') or Germ. *Nachtschwärmer*: 'x swarms out during the night' ('moth', also 'night owl').

4.5 Instrumental

The instrumental role is encoded in the complement in just a few cases in German and English:

Germ.: *Radfahrer*: x goes by bike ('biker', 'cyclist')

Engl.: *faith curer*: x cures with/by faith

This semantic role is not found in examples taken from Romance languages, but we can find it in Kxoe where the role is expressed by a verbal derivation suffix that assigns an instrument status to the classifier as in *ngú-xàú-kà-xò*: 'thing to sew a house with it' ('strings for sewing a house-mat').

In Baka place and instrument can only be expressed by determinative compounds where the determinatum is a verb-noun complex: *nda-na-bonà-'èè* 'house where goods are sold/house for selling goods' ('shop').

4.6 Manner

In cases where manner should be expressed the slot of the complement is filled by an adverb/adjective instead of a noun:

Germ.: *Warmduscher*: x showers hot; *Frühaufsteher*: x rises early

Engl.: *well-wisher*: x wishes well

Fr.: *lève-tard*: x gets up late; *gagne-petit*: x earns little

Span.: *cantaclaro*: x sings clearly (= 's.b. not mincing one's mind')

4.7 Summary

Summing up the findings presented above, we can make two general observations:

- (a) The overwhelming percentage of internal semantic roles expressed by the noun complement of the construction in all languages of the sample is the patient role. This observation can be explained by the fact that the main part of constructional compounds are based on prototypical transitive bivalent verbs like *to make* (s.b. makes sth.).¹³ The direct object is part of the verbal frame and is the default case for patient role. So in most cases the verb is even transitivized in the word formation process (e.g. by an affix like Germ. *be-*): *x bewohnt die Stadt* ('x inhabits the city') is nominalized as *Stadt-be-wohner* where *Stadt* ('city') is to be interpreted

as direct object. The alternative construction **Stadtwohner* underlies formation constraints and is not possible and would not be expected in this case.¹⁴

- (b) In most cases the respective semantic roles are expressed without the use of a grammatical marker: No case marking, no marking of definiteness or indefiniteness (article). In a number of cases prepositions are not used as grammatical markers: e.g. *faith curer* (x cures *by* faith), *Strandläufer* (x runs *on* the beach), *trotacalles* (x walks *in* the street). This corresponds to general observations on incorporation processes, where the nominal morphology (number, affixes, case markers, determiners, quantifiers, and other modifiers) is normally not incorporated with the noun (Mithun 2000:917).

Thus, in cases like *faith curer* only the semantic interpretation has to be taken into account: one can only cure animated entities, *faith* is not animated, and therefore its interpretation has to be different from the default one. As syntactic roles are not marked, because the constituents are loosing their relative flexibility within a sentence, there is a kind of implicational syntactic rule that has to be employed by the language user to find out the semantic role of the nominal complement:

- (i) interpretation as a patient
- (ii) if the verb does not open a patient role: interpretation as locative, instrumental etc.
- (iii) if the interpretation of (i) and (ii) is not possible: agent

There are, however, two exceptions in our data:

- (a) Examples, where number is expressed, e.g. in Romance languages (*cortacigarros*, *cavatappi*). The plural marking can only be explained by semantic motivation; it is much more common in Spanish than in other Romance languages (see Rainer 1993:271f.).
- (b) Examples, where an additional element (a derivational suffix as in Kxoe or a preposition as in German and Romance languages) is inserted – esp. for the semantic role ‘place’: *Bei-Mama-Wohner*, *monte-en-l’air* or *saltimbanco*. In this case there seems to be a differentiation within locative expressions: the locative preposition *in* (indicating the place) is not expressed, cf. *Strandläufer*, *city dweller*, whereas the preposition *into* (indicating the direction) in some languages has to be explicitly expressed (e.g. *saltimbanco*, ‘jump onto the bench’). In contrast, Kxoe language has to mark the change of the prototypical agent-role of the classifier into an instrument or place-role by an additional derivational suffix (-ò ‘in’ and -kà ‘with’).

More complex compounds, i.e. with more than one role or with an additional modifier such as Germ. *Cabrioletgeschlossenfahrer* are not lexicalised yet. They exist only in recent joking expressions of the German language, mentioned above:¹⁵

Handy-am-Gürtel-Träger (‘s.b. who carries his mobile on his belt’)
 dir. obj. – locativ – verb.AG

Cabrioletgeschlossenfahrer ('s.b. drives a convertible that is closed')
 dir. obj. – manner – verb.AG

Here, even the embedding of a sentence is possible as in *Wie-war-ich-Frager* ('s.b. who asks: how was I?'). The last expression can be considered as a transitional construction between constructional compounds and 'true' sentence-words as mentioned in the introduction of our paper.

5. Functions of incorporation processes

After considering the formal aspects of constructional compounds we now will look at the functional motivations of the incorporation processes involved. According to Mithun (2000:917) "the basic function of incorporated nouns is to modify the verb, narrowing the scope semantically". The result in polysynthetic languages is a new lexical, verbal item for this unitary concept. Constructional compounds can be interpreted in a similar way but the construction has undergone a further step: first the verb-noun complex forms a predication (x doing y) with a narrower scope than the simplex verb has: *cigarette-smoking* for instance has a more narrow scope than *smoking*, because it denotes the action of smoking just cigarettes and not only smoking in general. And second – in addition to noun incorporation processes in polysynthetic languages – the whole predication is nominalized.

Additionally, it should be taken into consideration that the complements have to be regarded as qualifiers (Mithun 2000:917) and therefore they are nonreferential and unindividuated items: i.e. in *cigarette-smoker* the noun *cigarette* cannot denote a special individual cigarette, but expresses the generic term 'cigarette'. The *cigarette-smoker* smokes cigarettes in general. If an individuated act such as 'smoking a special kind of cigarette' has to be expressed one has to use an explicit syntactic construction,¹⁶ for instance a relative clause *a person who smokes his last filter cigarette*. This applies also for Romance languages: *garde malade* vs. *une personne qui garde tous les malades* ('a person who watches all sick people'), for Kxoe and Baka: *wà-tšò-bè* ('songleader') vs. *bo kè 'é à tšò jókò bè kè nè* ('the person who strikes up this beautiful song') as well as for polysynthetic languages where a specific object is referred to by a noun outside of the verb (cf. Mithun, *ibid.*). This observation is contradictory to the assumptions of Baker (1988:78ff.) who distinguishes noun incorporation within constructional compounds from noun incorporation with complex verbs by the fact that in the latter case also referential nouns are employed. But the examples he gives are not convincing to his theory and can also be interpreted in a different way, as proposed by Mithun.¹⁷

As Mithun (2000:918) states for polysynthetic languages noun incorporation is also utilized in many other languages to manipulate case role. The incorporated noun is no more in a salient core case, but backgrounded. But in contrast to polysynthetic languages, in constructional compounds not only the noun but the whole predication (*doing y*) is backgrounded. Therefore, incorporation processes can be considered as

a strategy of backgrounding information like participle constructions or subordinate clauses do. According to the theory of grounding in discourse background information is information that doesn't move the reference time forward (Reinhardt 1984), i.e. all kind of things that are not part of the event-line: Something that happens at another time level or without any reference to time, like inherent characteristics of people. In discourse there is phased transition from maximal foregrounding to maximal backgrounding that can well be observed in the use of constructional compounds, too: Whereas the sentences in (a) each are part of the event-line, the information given in the relative clause in (b) is already background information but still individualizing the person. The person denoted by the constructional compound in (c), however, is completely nonreferential and its inherent predication is not part of the event-line at all:

- (a) *I saw a man at the town hall. He smoked a cigarette. Then he came down the stairs.* (smoking the cigarette is part of the action reported);
- (b) *I saw a man at the town hall, who smoked a cigarette [or smoking a cigarette]. He came down the stairs.*
(smoking the cigarette is not part of the action reported, but something that happens simultaneously and is a momentary characteristic of the person);
- (c) *I saw a cigarette smoker at the town hall. He came down the stairs.*
(smoking cigarettes is a permanent characteristic of the person. It is not necessary that the action happens at the moment of reporting).

This observation may explain the range of semantic fields covered by the lexicalised versions of constructional compounds: profession, habit or contemptuous expressions are ascribed to a person and are *per se* inherent (background) characteristics of the person: as well as a *teacher* teaches many times a *crocodile eater* doesn't only eat crocodiles one time, but in general.

6. Conclusions

As we have shown, incorporation seems to be a very common strategy used not only in polysynthetic languages, but in all types of languages to integrate lexical items into a larger lexical complex. The function of this process is not only a semantic one, i.e. narrowing the scope of the predication, but mainly a syntactic one, i.e. backgrounding the given information. Whereas polysynthetic languages use noun incorporation as a strategy to background nouns in otherwise foregrounded sentences, rather analytic languages use it to background the whole predication by nominalization of the verb-noun complex. This can be regarded as a further step in syntactic integration of propositions (cf. Raible 1992).

The transition of the different constructions of lexicalised items consisting of a complex predication can be arranged on a scale leading from full syntactic marking to full morphological marking (cf. Figure 1):

I	II	III	IV	V
<i>main clauses</i>	<i>sentence words</i>	<i>constructional compounds</i>	<i>determinative compounds</i>	<i>simplex</i>
	invariable fossilized word order + morphological markedness; defective word morphology	limited semantic roles; word order analogous to attributive adjuncts; defective noun phrase morphology; condensation of syntactic information	full noun phrase morphology; clear attributive relation	

- I. On the one end of the scale we have main clauses containing verb and complements: in a clause all constituents are marked by special morphological devices or word order; the order of the constituents is fixed or can be varied by focusing techniques.
- II. In 'sentence-words' in a classical definition like *Hans-guck-in-die-Luft*, the numerals in Kxoe and the Baka noun for 'West' the syntactic relations are still marked (by prepositions and case marking for instance), but the order of the constituents is invariable. It can be integrated as a constituent of a clause by adding an article for instance, but case marking and plural formation are not possible.
- III. Constructional compounds are defective insofar, as syntactic relations are no more marked within the internal structure of the compound and not all semantic roles can be expressed. In some languages (e.g. Germanic type) the order of the constituents does not correspond to the order in the free clause, but is analogous to attributive adjuncts (genitive, adjectives, determinants), so following not the verbal but yet the nominal syntax. In other languages the compounds can only be partly integrated by noun morphology (defective plural marking in Romance languages).
- IV. fully integrated compounds do not express any internal syntactic relations (type: *house-door*), but have full noun morphology.

Abbreviations

3	third person	IMP	imperative
AG	agentive	INF	infinitive
COMP	completive	INSTR	instrument
COMPL	complement	LOC	locative
CONV	converb	O	direct object
DEM	demonstrative	PART	participle
DU	dual	PFV	perfective
F	feminine	REL	relative
GER	gerund	SG	singular
II	juncture for past		

Notes

1. The examples for the respective European languages are taken from the monographs on word formation cited below as well as from the internet site www.ffh.de with a collection of joking constructions of German. The examples of Baka and Kxoe are taken from Brisson & Boursier (1984) and Kilian-Hatz (2003).
2. Baker (1988:78ff.), too, stresses the differences of incorporation processes in polysynthetic languages and English (representing the average standard European): the respective expressions belong to different verbal categories and show different referentiality. We, however, don't agree with the last assumption that will be discussed in Section 5 below.
3. The only difference between Slavic and Germanic languages is that there is a broader range of suffixes in the first group (cf. Horn-Helf 1997).
4. In this case it is not clear whether the verbal form is derived from a imperative form or the third person singular form; the diachronic analysis is in favor of the imperative hypotheses (Rohlf's 1969), but from a synchronic and semantic point of view it is likely to assume 3.Pers. forms (s. Rainer 1993; Thiele 1993). The imperative hypotheses would, however, be consistent with imperative 'sentence names' in German, like *Schauinsland* ('look into the country') as name of a mountain.
5. This typical Romance pattern is also found in English compounds dating back to the 14th century and borrowed from French: cf. *pick-pocket*, *spoil-sports*. They are, however, not productive anymore or limited to non-human denotata like trade names (*Xpel-air*, cf. Bauer 1996:205).
6. Here the variant with the infinitival form of the verb is also possible: *wà-na-máà-bo*.
7. Other semantic fields esp. found in Romance languages such as food or clothes (cf. Bierbach 1982:91ff.; Lloyd 1968) can be considered as a subtype of this field.
8. Lexicalisation of complex ('sentence-like') word-units is often used as a criterion of distinction between morphological and syntactic entities: the first are stored and retrieved from the mental lexicon as a whole, whereas the latter are repeatedly assembled anew (cf. Olsen 2000:899f.).

9. Dokulil (1964; 1981) hints at the fact that constructional compounds in languages like German and English can denote any of the mentioned semantic categories and are only semantically fixed by lexicalisation.
10. Here one also has to mention German constructions like *Hinkebein* ('limping person'; lit 'limp-leg'). Those, however, are to be considered as sentence words in a classical definition, where the verb is to be interpreted as a imperative form (cf. Fleischer & Barz 1995:213f.). See also Note 3.
11. As mentioned above these compounds are only joking expressions and ad-hoc-constructions denoting 'weaklings'.
12. This also holds for incorporation in polysynthetic languages (cf. Mithun 2000:917).
13. For formation constraints see Rivet (1999) and Toman (1987).
14. Similar constructions in English are marginal and occur almost exclusively in humorous prose, e.g. *lawn sprinkler runner througher* (Ch. M. Schulz, The Snoopy Festival). The double suffix is even a deviant pattern and never occurs in lexicalised compounds (cf. Hohenhaus 1996:86).
15. Here we again encounter the problem mentioned above: is *cigarette* a determinant of the noun *smoker* or is the whole complex nominalized? In our other examples we always dealt with compounds that don't have a counterpart derived from the simplex verb, this example is only taken here for illustration purposes.
16. For a discussion see Mithun (2000:925).

References

- Askedal, J. O. (1996). Überlegungen zum Deutschen als sprachtypologischem 'Mischtyp'. In E. Lang & G. Zifonun (Eds.), *Deutsch – typologisch* (pp. 369–383). Berlin: de Gruyter [IdS-Jahrbuch 1995].
- Baker, M. C. (1988). *Incorporation. A theory of grammatical function changing*. Chicago, IL: University of Chicago Press.
- Bauer, L. (1996). *English Word-formation*. Cambridge: Cambridge University Press.
- Bierbach, M. (1982). *Die Verbindung von Verbal- und Nominalelement im Französischen*. Tübingen: Narr.
- Brisson, R. (1984). *Lexique Français-Baka*. Douala: College Liberman.
- Dokulil, M. (1981). Zum wechselseitigen Verhältnis zwischen Wortbildung und Syntax (first published in 1964). In L. Lipka & H. Günther (Eds.), *Wortbildung* (pp. 82–93). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Eichinger, L. M. (2000). *Deutsche Wortbildung: Eine Einführung*. Tübingen: Narr.
- Fleischer, W., & Barz, I. (1995). *Wortbildung der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Hohenhaus, P. (1996). Ad-hoc-Wortbildung. Terminologie, Typologie und Theorie kreativer Wortbildung im Englischen. Frankfurt: Lang.
- Horn-Helf, B. (1997). *Kondensation als terminologisches Prinzip im Russischen*. Tübingen: Narr.
- Kilian-Hatz, C. (2003). *Kxoe-English Dictionary*. Köln: Rüdiger Köppe.
- Lang, M. F. (1990). *Spanish Word Formation: Productive derivational morphology in the modern lexis*. London: Routledge.

- Lakoff, G., & Johnson, M. (1980). *Metaphors we Live by*. Chicago, IL: University of Chicago Press.
- Lawrenz, B. (1997). Zu-spät-Kommer und Dumme-Fragen-Steller im Mann-von-Welt-Look: Phrasenkomposition und Phrasenderivation im Deutschunterricht. *Wirkendes Wort*, 47, 112–136.
- Lloyd, P. M. (1968). *Verb-Complement Compounds in Spanish*. Tübingen: Niemeyer.
- Marchand, H. (1983). *The Categories and Types of Present-day English Word-formation: A synchronic-diachronic approach*. München: Beck.
- Mithun, M. (2000). Incorporation. In G. Booij, C. Lehmann, & J. Mugdan (Eds.), *Morphology: An international handbook on inflection and word-formation*, Vol. 1 (pp. 916–928). Berlin: de Gruyter.
- Olsen, S. (2000). Composition. In G. Booij, C. Lehmann, & J. Mugdan (Eds.), *Morphology: An international handbook on inflection and word-formation*, Vol. 1 (pp. 897–915). Berlin: de Gruyter.
- Raible, W. (1992). *Junktion: Eine Dimension der Sprache und ihre Realisierungsformen zwischen Aggregation und Integration*. Heidelberg: Winter.
- Rainer, F. (1993). *Spanische Wortbildungslehre*. Tübingen: Niemeyer.
- Reinhardt, Ta. (1984). Principles of Gestalt perception in the temporal organization of narrative texts. *Linguistics*, 22, 779–809.
- Rivet, A. (1999). Rektionskomposita und Inkorporationstheorie. *Linguistische Berichte*, 179, 307–342.
- Rohlf, G. (1969). *Grammatica storica della lingua Italiana e dei suoi dialetti*, Vol. 3: *Sintassi e formazione delle parole*. Torino: Einaudi.
- Spencer, A. (2000). Morphology and syntax. In G. Booij, C. Lehmann, & J. Mugdan (Eds.), *Morphology: An international handbook on inflection and word-formation*, Vol. 1 (pp. 312–334). Berlin: de Gruyter.
- Thiele, J. (1993). *Wortbildung der französischen Gegenwartssprache*. Leipzig: Langenscheidt.
- Toman, J. (1987). *Wortsyntax. Eine Diskussion ausgewählter Probleme deutscher Wortbildung*. Tübingen: Niemeyer.
- Whaley, L. J. (1997). *Introduction to Typology: The unity and diversity of language*. Thousand Oaks, CA: Sage.

Toward a typological perspective for Emai's BE constructions*

Ronald P. Schaefer and Francis O. Egbokhare

SIU Edwardsville / University of Ibadan

This paper assesses the typological character of five BE constructions in Emai, a Benue-Congo language of Nigeria's Edoid group. It builds on previous investigations of BE by Verhaar (1967–1973), Declerck (1988), Keizer (1991) and particularly Stassen (1997). We assume that Emai's BE types reflect a typological parameter where forms of BE within and across languages stand relative to one another. To explore the nature of this parameter, we attend to previously undescribed formal and functional properties of Emai's BE constructions. These are derived from on-going documentation incorporating oral narrative texts as well as dictionary and grammar descriptions (Schaefer & Egbokhare 1999a, b, in preparation). As a result, we postulate a multidimensional grammatical hierarchy integrating BE forms and their distributional behavior.

Introduction

An integral aspect of Stassen's (1997) cross-linguistic investigation is a distinction between predication and identification. Through identification, a speaker re-aligns an entity within the discourse knowledge of the hearer, either by naming it or by equating it with another entity. Through predication, a speaker assigns an entity to one of a limited number of types: class, property or event. Predication types manifest a hierarchical relation between their prototypical syntactic category composition and their phenomenological realization over time (Frawley 1992). That is, the syntactic categories through which the predicate concepts class, property and event are realized differ in their temporal stability. Nominal class predicates illustrated by the noun *witch* in *Harriet is a witch*, tend to be more stable regarding their temporal character than verbal event predicates shown by the verb *sit* in *Harriet is sitting in the house*. In essence, verbal predicates are less time stable than nominal predicates. Between these polar points of temporal stability lies the predicate type property, exemplified by the adjective *sleek* in *Harriet is sleek*. According to Givón (1984), the categories noun, verb, and adjective underlying these predication types reflect a Time Stability Scale. For this paper, we will assume the following representational shape for the Time Stability

Scale applicable to predication, with the leftmost position reflecting the least temporal stability, rightmost the greatest stability, and a median level of stability between these two.

EVENT < PROPERTY < CLASS

No comparable discussion of temporal stability has been advanced for identification. According to Stassen, identification is composed of specification and equation types. The former refers to specifying an entity (often accompanied by a pointing gesture) while the latter equates two entities, already known to the hearer, through a relationship not assumed to be known by the hearer. For identification forms, one might usefully postulate a relationship of time stability grounded to the deictic moment of utterance. That is, equation and specification varieties of identification appear asymmetrically related to deictic context. Specification is closely associated with a pointing gesture and the availability of a specified referent in the immediate physical or discourse context. This is not so for the referents linked in an equational identity construction, which is less dependent on the immediate deictic context. From this deictic perspective, constructions illustrating identification's two types manifest different degrees of stability over time. Below, the leftmost position occupied by specification identity is more highly dependent on deictic context, therefore manifesting less temporal stability, than the rightmost position held by equational identity.

SPECIFICATION < EQUATION

At this juncture, there is little linguistic motivation for such an identification scale or even for linking identification and predication scales in some fashion. The distributional behavior of Emai's BE forms suggests, nonetheless, that identification and predication are linked and that they and their respective forms are ranked relative to one another in the fashion indicated below.

EVENT < PROPERTY < CLASS < SPECIFICATION < EQUATION

As for Stassen, his review of intransitive predication includes an extensive cross-linguistic analysis of BE forms and their functions. He amplifies the distinction between predication and identification by identifying five functional types of BE: location (LOC), property (PRP), class membership (CMB), specification (SPE) and equational identity (EID). As predication types, the first three of these assign an entity to a location (LOC), a property (PRP) or a class (CMB). The latter two, concerned with the identity of a discourse participant, either specify that participant (SPE) or equate some known participant with the identity of another known participant (EID). Assuming these functional types, Stassen examines their formal realization in a large sample of languages from around the world. In the most transparent analytic case, one would expect an isomorphic, 1–1 relation between form and function. In more opaque cases, a single form might express more than one function or, equivalently, more than one function might be expressed by a single form. Stassen refers to the more opaque cases as "takeover." In general, "takeover" reflects an economy principle claiming that lan-

guages tend to minimize the number of surface forms encoding functions within a domain. His survey shows that only a minority of languages distinguishes between predication and identification functions through formal encoding. Emai's BE forms, however, make exactly this distinction.

Stassen's analysis of his cross-linguistic sample assumes simple or complete takeover. A BE form encoding one function, class membership for instance, also encodes another function, e.g. property. This leads him to arrive at several conclusions applicable not only to language in general but also to the Niger-Kordofanian phylum of Africa in particular. Among others, these include the following. Predicate locations exhibit a strong preference for coding by a fully verbal form of BE. Seldom does a predicate from class membership or property take over predicate location functions. Predicate nominals are less resistant to takeover, giving way to equational identity constructions in many languages. Property predicates are the most highly susceptible to takeover either by location predicates, nominal predicates, or both. No obvious takeover relationship between specification and equational identity appeared in his sample.

Stassen's conclusions about complete takeover resonate with statements from investigators of various West African languages (Welmers 1973). For instance, the DeBose and Faraclas (1993) assessment of Benue-Congo languages in southeastern Nigeria claims the presence of BE forms serving multiple functions. More specifically, Ellis and Boadi (1969) show that Twi, a Benue-Congo language of West Africa, exhibits three BE forms relative to the four functions of class membership, property, location and equational identity. The copula form *ye* serves as conceptual predicate for both class membership (CMB) and property (PRP), although the direction of takeover is not discernible from their data. A wider range of constructions seems required to shed light on takeover direction.

- | | | |
|--------|----------------------------------|-----|
| (1) a. | <i>obarima no ye osofoo</i> | CMB |
| | man the BE priest | |
| | 'The man is a priest.' | |
| b. | <i>dua no ye tia</i> | PRP |
| | tree the BE short | |
| | 'The tree is short.' | |
| c. | <i>nnipa no wo sukuu</i> | LOC |
| | people the BE school | |
| | 'The people are at school.' | |
| d. | <i>obarima no ne osofoo</i> | EID |
| | man the BE priest | |
| | 'It is the man who is a priest.' | |

1. Emai's BE forms

Stassen pays little attention to what might be called "partial takeover," e.g. cases where a BE form distributionally aligned with the majority of constructions of one function articulates a single construction or small set of constructions within another function. This phenomenon characterizes Emai and reveals a relative ranking among its BE forms.

We turn now to the analysis of Emai.¹ Its five BE forms and their corresponding functions, as indicated below, are equational identity (EID) *khi*, specification (SPE) *o*, class membership (CMB) *vbi*, property (PRP) *u*, and locative (LOC) *ri*.²

- | | | |
|--------|---|--------------------------------------|
| (2) a. | àlèkè lí í khi óì
Aleke PF SC BE thief
'It is Aleke who is a thief.' | equational identity (EID) <i>khi</i> |
| b. | óì <u>óò</u>
thief BE
'It's a thief.' | specification (SPE) <i>o</i> |
| c. | àlèkè í ì vbi óì
Aleke SC NEG BE thief
'Aleke is not a thief.' | class membership (CMB) <i>vbi</i> |
| d. | ólì èkpèn ú nwènénwèné
the leopard BE spotted
'The leopard is spotted.' | property (PRP) <i>u</i> |
| e. | àlèkè rù vbi <u>ékóà</u>
Aleke BE LCT room
'Aleke is in the room.' | locative (LOC) <i>ri</i> |

Each of these five forms exhibits some asymmetric distributional behavior with regard to its construction types. This asymmetry is most strongly evident in the interrogative mode, especially with polar (*yes/no*) and information (*wh-*) questions, somewhat less in the declarative with the marking of predicate and participant negation, but never in the imperative, where BE forms uniformly fail. In the following sections, we will explore the properties of Emai's BE forms in question frames as well as negative and focus constructions. As a prelude to our examination, we advance our postulated multi-dimensional hierarchy that attempts to integrate the behavior of these forms, from LOC *ri* to EID *khi*, with PRP *u*, CMB *vbi* and SPE *o* then ranked respectively from left to right.

LOC < PRP < CMB < SPE < EID

Overall ranking on this hierarchy respects the dimensions of predication (LOC, PRP, and CMB) and identification (SPE and EID). Specific ranking within each of these dimensions is motivated by the distributional stability of forms within and across construction types. Furthermore, we conclude that this ranking based on distributional stability aligns with the temporal stability ranking of BE forms within predication and

identification dimensions. As it happens, forms with the highest temporal stability on their respective dimensions (e.g. CMB and EID) exhibit the greatest amount of syntactic and semantic restrictiveness.

The distinction between predication and identification correlates with properties of initial noun phrases in Emai's BE constructions. An emphatic pronoun (*iyòìn*) is obligatory as initial noun phrase for basic constructions with identification forms, while ungrammatical for predication constructions. Both SPE and EID constructions assign their initial noun phrase to focus position, with an obligatory emphatic (EMP) pronoun (*iyòìn*) (3a) and (3b). They do not allow non-emphatic subject pronouns (*ò*).

- (3) a. *iyòìn* / **ò* *kí í khi òì* EID
 3S-EMP 3S NF SC BE thief
 'It isn't she who is a thief.'
- b. *iyòìn* / **ò* *kí óò* SPE
 3S-EMP 3S NF BE
 'It isn't she.'

In contrast, the predication forms CMB, PRP and LOC assign their initial noun phrase to subject position. They require a non-emphatic, subject pronoun in third person (*ò*) (4a–c). In subject position they do not accept an emphatic pronoun (*iyòìn*).

- (4) a. **iyòìn* / *ò* *ì vbi òì* CMB
 3S-EMP 3S NEG BE thief
 'She is not a thief.'
- b. **iyòìn* / *ò* *ú nwènénwèné* PRP
 3S-EMP 3S NEG spotted
 'It is spotted.'
- c. **iyòìn* / *ò* *rî vbi ékóà* LOC
 3S-EMP 3S BE LCT room
 'She is in the room.'

The identification forms SPE and EID also fail to accept predicate negation. They do not admit the third person negative particle (*ì* NEG) (5a) and (5b). As shown by (3a) and (3b) above, identification constructions only admit participant negation marked by the negative focus particle *kí*. As was true of their emphatic pronoun use, the identification forms EID and SPE exhibit comparable marking of negation.

- (5) a. **àlèkè lí í ì khi òì* EID
 Aleke PF SC NEG BE thief
 'It is Aleke who is not a thief.'
- b. **òì ì ò* SPE
 thief NEG BE
 'It is not a thief.'

Although identification constructions manifest comparable marking of negative focus, they do not uniformly tolerate overt marking of positive focus position. EID allows

overt marking by the positive focus (PF) particle *lí* (6a), while SPE does not (6b). Relative to the latter, we deduce the focus position of the initial noun phrase from its obligatory emphatic pronoun form.

- (6) a. *ìyò̀ìn lí í khi óì* EID
 3S-EMP PF SC BE thief
 ‘It is she who is a thief.’
 b. *ìyò̀ìn óò / *ìyò̀ìn lí óò* SPE
 3S-EMP BE 3S-EMP PF BE
 ‘It’s she.’

Despite this slight difference concerning overt focus marking, specification and equational identity exhibit grammatical properties that distinguish the dimension identification from predication on our postulated BE hierarchy. Two of these properties, as we have seen, are pronoun type and negation.

2. Polar interrogatives for identification

Within and between identification and predication dimensions on our BE parameter, correspondence relations across query and response units of polar interrogation for individual BE forms exhibit further distributional asymmetry. Relations of significance include polarity agreement (affirmative vs. negative) between query and response, acceptability of an explicit confirming (*hèè* ‘yes’) or disconfirming (*òghò* ‘no’) lexical response, and construction responses incorporating marking of polarity type both lexically and through either participant negation (*ki*) or predicate negation (*í í*). Most important, these constructions reveal the extent to which query and response units of an interrogative frame share or fail to share a BE form. In this regard, we find that all BE forms serve as queries for polar interrogation but not all serve as responses.

Within the identification dimension, EID is the more restricted form. It admits SPE responses but not EID responses. Affirmative EID queries about a non-definite referent in post-*khi* position do not accept confirming EID response constructions with *hèè* and participant affirmation (7b). Instead, they allow confirming or disconfirming lexical responses with *hèè* or *òghò*, respectively, (7c). They also accept SPE construction responses but only with *hèè* and participant affirmation (7d), not *òghò* and participant negation (7e).

- (7) a. *àlèkè lí í khi óì?* EID
 Aleke PF SC BE thief
 ‘Is it Aleke who is a thief?’
 b. *!hèè, àlèkè lí í khi óì* EID
 yes Aleke PF SC BE thief
 ‘Yes, it is Aleke who is a thief.’

- c. *hèè* / *òghò*
yes no
'Yes.' 'No.'
- d. *hèè*, *òì* *óò* SPE
yes thief BE
'Yes, she's a thief.'
- e. *!òghò*, *òì* *kí* *óò* SPE
no thief NF BE
'No, she is not a thief.'

Affirmative EID queries receive a slightly different response when overt definite reference (marked by the definite determiner *òlì*) characterizes post-*khi* position (8a). In such cases, EID queries allow EID or SPE responses. As a lexical response, confirming *hèè* and disconfirming *òghò* are acceptable (8b). As construction responses, only disconfirming responses are acceptable. A disconfirming EID *khi* response with participant negation is acceptable (8c), as is a disconfirming SPE response with participant negation (8d). Unacceptable is a confirming EID response with participant affirmation (8e).

- (8) a. *àlèkè lí í khi òlì òì?* EID
Aleke PF SC BE the thief
'Is it Aleke who is the thief?'
- b. *hèè* / *òghò*
yes no
'Yes.' 'No.'
- c. *òghò*, *ìyòìn* *kí í khi òì* EID
no 3S-EMP NF SC BE thief
'No, it isn't she who is a thief.'
- d. *òghò*, *ìyòìn* *kí óò* SPE
no 3S-EMP NF BE
'No, it is not she.'
- e. *!hèè*, *àlèkè lí í khi òlì òì* EID
yes Aleke PF SC BE the thief
'Yes, it is Aleke who is the thief.'

Continuing with the identification dimension, we find that SPE *ò* is less restricted. SPE queries, at least in the affirmative, admit SPE responses. In response to affirmative queries from SPE about a non-human referent, little restriction is evident. Affirmative SPE queries (9a) accept lexical confirmation (*hèè*) or disconfirmation (*òghò*) responses (9b). In addition, SPE admits as a construction response either confirming *hèè* and participant affirmation (9c) or disconfirming *òghò* and participant negation (9d).

- (9) a. *ùbèlè óó?* SPE
gourd BE
'Is it a gourd?'

- b. hèè / òghò
yes no
'Yes.' 'No.'
- c. hèè, ùbèlè òò SPE
yes gourd BE
'Yes, it is a gourd.'
- d. òghò, ùbèlè kí òò SPE
no gourd NF BE
'No, it is not a gourd.'

Affirmative SPE queries about a human referent lead to SPE and CMB responses. Affirmative SPE queries about a human referent (10a) accept confirming or disconfirming lexical responses (10b). As a construction response, SPE with participant affirmation and hèè is acceptable (10c). However, SPE does not allow a disconfirming response construction with participant negation and òghò (10d).

- (10) a. òì óó? SPE
thief BE
'Is she a thief?'
- b. hèè / òghò
yes no
'Yes.' 'No.'
- c. hèè, òì òò SPE
yes thief BE
'Yes, she is a thief.'
- d. !òghò, òì kí òò SPE
no thief NF BE
'No, she is not a thief.'

Instead, affirmative SPE queries about a human referent require a distinct BE form on the hierarchy. That is, SPE queries about a human referent demand a response construction in which the SPE form is taken over by the CMB form *vbi* from the predication dimension. For an affirmative SPE query about a human referent (11a), a disconfirming CMB response with òghò and predicate negation is required (11b). Although one might consider a disconfirming construction response with equational identity possible because of the common identification dimension, an affirmative SPE query never attracts the identification form EID *khi* with òghò and participant negation (11c).

- (11) a. òì óó? SPE
thief BE
'Is she a thief?'
- b. òghò, ó ì *vbi* òì CMB
no 3S NEG BE thief
'No, she is not a thief.'

- c. *!òghò, àlèkè kí í khi òì* EID
 no Aleke NF SC BE thief
 'No, it is not Aleke who is a thief.'

We turn now to participant negation queries for EID and SPE. Of the two, negative EID queries admit only EID responses. With regard to participant negation EID queries (12a), we find a lexical confirming *hèè* or disconfirming *òghò* response (12b). As a construction response, negative EID queries allow confirming EID constructions with participant affirmation (12c). They do not accept as response either confirming or disconfirming EID constructions with participant negation (12d); confirming or disconfirming SPE constructions with participant affirmation or participant negation (12e); or disconfirming CBM constructions with predicate negation (12f).

- (12) a. *àlèkè kí í khi òlí òì* EID
 Aleke NF SC BE the thief
 'Is it not Aleke who is the thief?'
- b. *hèè / òghò*
 yes no
 'Yes.' 'No.'
- c. *hèè, ìyòìn lí í khi òlí òì* EID
 yes 3S-EMP PF SC BE the thief
 'Yes, it is she who is the thief.'
- d. *!òghò, / !hèè, ìyòìn kí í khi òlí òì* EID
 no yes 3S-EMP NF SC BE the thief
 'No / Yes, it isn't she who is the thief.'
- e. *!hèè, ìyòìn óò / ìyòìn kí óò* SPE
 yes 3S-EMP BE 3S-EMP NF BE
 'Yes, it is she / it isn't she.'
- f. *!òghò, ó ì vbi òlí òì* CMB
 no 3S NEG BE the thief
 'No, she is not the thief.'

SPE participant negation queries about a human referent accept SPE and CMB responses. SPE queries with participant negation (13a) attract lexical disconfirming responses with *òghò* but not confirming responses with *hèè* (13b). Construction responses incorporating contrasting BE forms are available for participant negation SPE queries. As one option, we encounter confirming SPE *ò* response constructions with *hèè* and participant affirmation (13c). As another response, we find disconfirming *òghò* and predicate negation CMB *vbi* (13d). Construction responses that incorporate confirming or disconfirming SPE responses with participant negation are unacceptable (13e).

- (13) a. *òì kí óó?* SPE
 thief NF BE
 'Isn't she a thief?'

- b. òghò / !hèè
no yes
'No.' 'Yes.'
- c. hèè, ói òò SPE
yes thief BE
'Yes, she's a thief.'
- d. òghò, ó ì vbì ói CMB
no 3S NEG BE thief
'No, she is not a thief.'
- e. !hèè / !òghò, ói kí òò SPE
yes no thief NF BE
'Yes / No, she isn't a thief.'

SPE participant negation queries about a non-human referent accept only SPE responses. An SPE participant negation query about a non-human referent (14a) admits lexical disconfirming òghò responses but not confirming hèè responses (14b). It allows disconfirming but not confirming SPE construction responses with participant negation (14c). It does not admit confirming SPE responses with participant affirmation (14d) or disconfirming CMB responses with predicate negation (14e).

- (14) a. ùbèlè kí óó? SPE
gourd NF BE
'Is it not a gourd?'
- b. òghò / !hèè
no yes
'No.' 'Yes.'
- c. òghò, / !hèè, ùbèlè kí òò SPE
no yes gourd NF BE
'No, it is not a gourd.'
- d. !hèè, ùbèlè óò SPE
yes gourd BE
'Yes, it is a gourd.'
- e. !òghò, ó ì vbì ùbèlè CMB
no 3S NEG BE gourd
'No, it is not a gourd.'

As observed throughout this section, the identification functions SPE and EID in polar interrogation frames exhibit partial take over. Their distributional behavior supports several aspects of the proposed BE hierarchy. It reveals that EID is distributionally more restricted than SPE. The EID form occurs in response units for EID queries, although the query-response frame must contrast along the affirmative-negative dimension. EID never serves in a response unit for SPE queries; it does not take over in query or response units where one might expect SPE for instance. On the other hand, SPE is less distributionally restricted in polar interrogation frames. SPE serves as a response unit not only for SPE queries that manifest predicate affirmation but also for

affirmative EID queries. In the latter, the SPE form takes over for the EID form. In another case, that of a human referent and predicate negation, the SPE form itself is taken over in response units by the predication form CMB.

The behavior of identification forms in polar interrogation bears directly on their ranking on our postulated BE hierarchy. Of the two identification forms SPE and EID, SPE partially takes over for EID but EID never partially takes over for SPE. Furthermore, since SPE is partially taken over by CMB, SPE rather than EID is more closely ranked to the predication form CMB.

LOC < PRP < CMB < SPE < EID

3. Polar interrogatives for predication

Within the predication dimension, polar interrogation reveals asymmetry among constructions incorporating location, property and class forms. It also shows that, in contrast to information questions to be examined shortly, polar interrogation queries incorporating predication forms never admit as responses identification constructions.

Of all the predication BE forms in Emai, CMB *vbi* is the more highly restricted in polar interrogation frames. CMB queries accept CMB responses, but only with predicate negation. Relative to CMB query units with predicate negation (15a), only disconfirming lexical and CMB construction responses are acceptable. As their lexical responses, CMB predicate negation queries admit disconfirming *òghò* but not confirming *hèè* (15b). As their construction response, CMB queries allow disconfirming constructions with *òghò* 'no' and predicate negation (15c). Unacceptable are confirming CMB response constructions, regardless of predicate affirmation or negation (15d).

- (15) a. *àlèkè í ì vbi òì?* CMB
 Aleke SC NEG BE thief
 'Is Aleke not a thief?'
 b. *òghò / !hèè*
 no yes
 'No.' 'Yes.'
 c. *òghò, ó ì vbi òì* CMB
 no 3S NEG BE thief
 'No, she is not a thief.'
 d. *!hèè, ó vbi òì / ó ì vbi òì* CMB
 yes 3S BE thief 3S NEG BE thief
 'Yes, she is a thief / she is not a thief.'

Overall, CMB requires predicate negation except where scaffolded by extremely rich context. The limited distribution of the affirmative CMB construction is evident in its failure as query unit in a polar interrogative frame. CMB never occurs as query (16a)

relative to either a disconfirming CMB response construction with *òghò* and predicate negation (16b), a confirming CMB construction with *hèè* and predicate affirmation (16c), or any other response construction.

- (16) a. *!àlèkè vbi òì?* CMB
 Aleke BE thief
 'Is Aleke a thief?'
 b. *!òghò, ó ì vbi òì* CMB
 no 3S NEG BE thief
 'No, she is not a thief.'
 c. *!hèè, ó vbi òì* CMB
 yes 3S BE thief
 'Yes, she is a thief.'

The remaining predication forms PRP *u* and LOC *ri* exhibit fewer restrictions across query and response units of polar interrogation frames. For their respective affirmative queries, each serves as its own response in the confirming case. In the disconfirming case, this is also true except that LOC *ri* requires its suppletive form *e*.

Polar interrogation reveals that PRP queries accept PRP responses. An affirmative PRP construction (17a) serves as query for lexical responses of confirmation with *hèè* or disconfirmation with *òghò* (17b). It also attracts confirming PRP construction responses incorporating *hèè* and predicate affirmation (but not predicate negation) (17c), and disconfirming PRP constructions manifesting *òghò* and predicate negation (but not predicate affirmation) (17d).

- (17) a. *ólí ógédé ú kísín?* PRP
 the plantain BE tiny
 'Is the plantain tiny? / The plantain is tiny, isn't it?'
 b. *hèè / òghò*
 yes no
 'Yes.' 'No.'
 c. *hèè, ó ú kísín / !hèè, ó ì ù kísín* PRP
 yes 3S BE tiny yes 3S NEG BE tiny
 'Yes, it is tiny.'
 d. *òghò, ó ì ù kísín / !òghò, ó ú kísín* PRP
 no 3S NEG BE tiny no 3S BE tiny
 'No, it is not tiny.'

Similarly, LOC queries in polar interrogation receive LOC responses. An affirmative LOC construction as query exhibits a response range comparable to PRP. Affirmative LOC queries (18a) admit lexical responses of confirmation with *hèè* or disconfirmation with *òghò* (18b). Locative query constructions attract confirming LOC response constructions that show *hèè* and predicate affirmation (18c) or a disconfirming response construction that incorporates *òghò* and predicate negation (18d).

- (18) a. *àlèkè rî vî ékóà?* LOC
Aleke BE LCT room
'Is Aleke in the room? / Aleke is in the room, isn't she?'
- b. *hèè / òghò*
yes no
'Yes.' 'No.'
- c. *hèè, ò rî vî ékóà* LOC
yes 3S BE LCT room
'Yes, she is in the room.'
- d. *òghò, ó ì è vî ékóà* LOC
no 3S NEG BE LCT room
'No, she is not in the room.'

PRP *u* and LOC *ri* each manifests a corresponding polar query with predicate negation, the latter again requiring the suppletive form *e*. As responses framed with PRP and LOC forms respectively, they admit disconfirming responses with *òghò* but not confirming responses with *hèè*, as indicated in (19) and (20).

A PRP query evincing predicate negation (19a) accepts a PRP response. PRP admits a disconfirming lexical response consisting of *òghò* but not a confirming lexical response composed of *hèè* (19b). Construction responses align with these lexical responses. A predicate negation PRP query allows disconfirming PRP response constructions manifesting *òghò* and predicate negation (19c), not a confirming PRP construction response incorporating *hèè* and predicate affirmation (19d).

- (19) a. *ólì ògèdè í ì ù kísín?* PRP
the plantain SC NEG BE tiny
'Isn't the plantain tiny? / The plantain isn't tiny, is it?'
- b. *òghò / !hèè*
no yes
'No.' 'Yes.'
- c. *òghò, ó ì ù kísín* PRP
no 3S NEG BE tiny
'No, it is not tiny.'
- d. *!hèè, ó ú kísín* PRP
YES 3S BE tiny
'Yes, it is tiny.'

A comparable overall pattern exists for LOC queries with predicate negation, since they receive LOC responses. A LOC query with predicate negation (20a) allows lexical responses with disconfirming *òghò* but not confirming *hèè* (20b). As construction responses, predicate negation LOC queries accept disconfirming LOC response units characterized by *òghò* and predicate negation (20c) but not confirming LOC response units comprised of *hèè* and predicate affirmation (20d).

- (20) a. *àlèkè í ì è vbí ékóà?* LOC
 Aleke SC NEG BE LCT room
 'Isn't Aleke in the room? / Aleke isn't in the room, is she?'
 b. *òghò / !hèè*
 no yes
 'No.' 'Yes.'
 c. *òghò, ó ì è vbí ékóà* LOC
 no 3S NEG BE LCT room
 'No, she is not in the room.'
 d. *!hèè, ò rî vbí ékóà* LOC
 yes 3S BE LCT room
 'Yes, she is in the room.'

Query-response frames associated with polar interrogation of the predication forms CMB, PRP, and LOC exhibit limited take over behavior. In fact, polar interrogative frames reveal no instances of partial takeover within the prediction dimension. It is only in the relationship between predication and identification, witnessed in the preceding section, that take over by a predication form occurs. Recall that the CMB form was more closely linked to identification forms than any of the other prediction forms through its role as response in polar interrogation frames with SPE queries about a human referent. Indeed, CMB is more restricted in its distribution than LOC and PRP forms. It serves as query only in negative predication constructions, never in affirmative constructions. PRP and LOC, on the other hand, appear without restraint across the dimensions affirmative and negative in response and query units of polar interrogation frames. Each appears as response to its own query, although LOC requires a suppletive form in the case of predicate negation. Overall, polar interrogation frames reveal little difference in the distributional character of PRP and LOC. Finally, since none of the predication forms admitted an identification form as a response unit in polar interrogation frames, the dimensional character of our postulated BE hierarchy distinguishing predication from identification is again reinforced.

LOC < PRP < CMB < SPE < EID

4. Information interrogatives

Additional relationships among BE forms are suggested by morphosyntactic shapes in information questions. Identification forms never serve as query for predication responses but predication queries accept at least one identification form as response. For one identification function and one predication function, the respective information question queries require a form on the BE hierarchy to the left of the one otherwise expected. There is thus a tendency for leftward shift in formation of query information questions. Essentially, this constitutes a partial takeover by one BE form of a function

associated with another BE form. SPE and PRP are both affected by a takeover process. Each shifts to the left to secure a lexical BE resource for an information-question query, although not for the response. As a response to information question queries, SPE exhibits wide distribution, not only for identification forms, as one might expect, but also for all predication forms except the locative argument of LOC.

4.1 Information Interrogatives for Identification

Consider, first, information questions for BE forms on the identification dimension. EID queries admit EID and SPE responses. For EID information questions (21a), responses with EID (21b) or SPE (21c) are acceptable.

- (21) a. óé' *í* *khì* *óì*? EID
 who SC BE thief
 'Who is a thief?'
 b. *àlèkè* *lí* *í* *khì* *óì* EID
 Aleke PF SC BE thief
 'It is Aleke who is a thief.'
 c. *àlèkè* óò SPE
 Aleke BE
 'It is Aleke.'

For SPE information questions, SPE responses are acceptable, although the query unit requires the PRP shape. Where the SPE form might be expected in its corresponding information question query (22a), the PRP shape *u* is obligatory (22b). As response, SPE is acceptable but PRP is not (22c).

- (22) a. *óé' ó óò? SPE
 who 3s BE
 'Who is it?'
 b. óé' ó *ú*? PRP
 who 3s BE
 'Who is it?'
 c. *àlèkè* óò SPE
 Aleke BE
 'It's Aleke.'

These query-response frames for identification forms reveal partial takeover affecting both EID and SPE forms. However, EID again appears more stable or restricted in its distribution, while SPE ranges over a wider range of frames. For EID, takeover is evident only in response units of interrogative information frames and then only optionally. EID queries allow either EID responses or SPE responses. Never, though, does EID occur in the response unit for an SPE query. For SPE, takeover is evident in its query unit. What obligatorily takes over in the SPE query is the PRP form. Thus, SPE exhibits a leftward shift in the search for a BE form in formation of its query unit for

information questions, utilizing the PRP form. SPE's lack of distributional restriction is also evident in its response role for SPE and EID queries.

4.2 Information interrogatives for predication

The behavior of predication forms in information interrogatives adds confirming detail to the postulated BE hierarchy. CMB and PRP information questions require additional morphemes in their query but LOC does not. CMB queries admit CMB and SPE responses. CMB information questions about the subject argument require a locative *vbi* phrase with partitive significance (*vbi úsèé vbá*) in their query form (23a). Acceptable responses include constructions with emphatic *òkpá* 'alone' in positive focus position with CMB (23b) or SPE (23c).

- (23) a. *óé' ó ì vbi òí vbí úsèé vbá?* CMB
 who 3s NEG BE thief LCT midst your
 'Who is not a thief among you?'
 b. *àlèkè òkpá lí ó ì vbi òí* CMB
 Aleke alone PF 3s NEG BE thief
 'It is Aleke alone who is not a thief.'
 c. *àlèkè òkpá óò* SPE
 Aleke alone BE
 'It is Aleke alone.'

No information question query corresponds to a non-partitive CMB *vbi* construction. This is true regardless of predicate negation (24a) or predicate affirmation (24b). Similarly, no information question that queries the post-*vbi* argument in CMB constructions is acceptable (24c).

- (24) a. **óé' ó ì vbi òí?* CMB
 who 3s NEG BE thief
 'Who is not a thief?'
 b. **óé' ó vbi òí?* CMB
 who 3s BE thief
 'Who is a thief?'
 c. **óé' áléké í ì vbi?* CMB
 who Aleke SC NEG BE
 'Who is Aleke not?'

Turning now to the PRP form, we find that PRP information question queries receive PRP and SPE responses. PRP information questions that query the subject argument require deictic *àin* 'that' with sortal (*kind of*) significance (25a). Failure to include *àin* in the query frame leads to ungrammaticality (25b). As responses to PRP-*àin* queries, positive focus constructions with PRP forms (25c) or SPE forms (25d) are acceptable.

- (25) a. *émé' ó ú nwènénwèné àin?* PRP
 what 3s BE spotted that
 'What kind of thing is spotted?'
 b. **émé' ó ú nwènénwèné?* PRP
 what 3s BE spotted
 'What is spotted?'
 c. *èkpèn lí ó ú nwènénwèné* PRP
 leopard PF 3s BE spotted
 'It is a leopard that is spotted.'
 d. *èkpèn óò* SPE
 leopard BE
 'It's a leopard.'

Our ranking of PRP relative to LOC is affected by the next set of information questions. As shown in (26a), the PRP form never occurs in an information query focused on its property value. To frame an acceptable information interrogation for property, the LOC form *ri* must occur in the query (26b) and the PRP form *u* in the response (26c). No other combination is grammatically acceptable.

- (26) a. **ébé' ólí úbélé í ù?* PRP
 how the gourd MAN BE
 'How is the gourd?'
 b. *ébé' ólí úbélé í rî?* LOC
 how the gourd MAN BE
 'In what state is the gourd? / How is the gourd?'
 c. *ó ú' kîsîn* PRP
 3s BE tiny
 'It is tiny.'

Unlike PRP, Emai's LOC form is not restricted across information interrogative frames. LOC queries about the locative argument admit LOC responses. Information questions focused on the locative argument in LOC constructions do not exhibit the distributional asymmetry witnessed with queries about the PRP predicate. The query unit of an information interrogative for LOC *ri* (27a) requires a LOC response (27b), never an SPE response (27c). And in contrast to PRP frames for information interrogation, a LOC response such as (27b) is never linked to a query formed with PRP *u* (27d).

- (27) a. *ébé' ólí ógéde rî?* LOC
 where the plantain BE
 'Where is the plantain?'
 b. *ìtébù lí ó rî* LOC
 table PF 3s BE
 'It's on the table.'
 c. *!ìtébù óò* SPE
 table BE
 'It's a table.'

- d. **ébé'* *ó**lí* *ó**gédé* *í*? PRP
 where the plantain BE
 'Where is the plantain?'

Lastly, we note that the subject argument of LOC constructions accepts as response either LOC or SPE constructions (28b–c).

- (28) a. *ó**é'* *ó* *rî* *v**bí* *ìwè*? LOC
 who 3s BE LCT house
 'Who is in the house?'
 b. *àlèkè* *lí* *ó* *rî* *v**bí* *ìwè* LOC
 Aleke PF 3s BE LCT house
 'It is Aleke who is in the house.'
 c. *àlèkè* *ó**ò* SPE
 Aleke BE
 'It is Aleke.'

As the preceding query-response frames for information questions indicate, takeover within the predication dimension is limited but revealing. CMB appears more restricted in its distribution than either remaining predication form. In fact, neither CMB nor PRP stand grammatically unsupported in their respective query units for information questions, although CMB appears to require a bit more grammatical scaffolding. In addition to obligatory negation, CMB requires a partitive *vbi* phrase. PRP, on the other hand, demands only sortal *àin*. However in contrast to the preceding two predication forms, only LOC does not require any grammatical scaffolding. It is also the only predication form that demonstrates partial takeover within the predication dimension. LOC takes over for PRP in the latter's query unit. Perhaps this is related to the takeover behavior by PRP of the query unit for the identification form SPE and by CMB of the response unit for SPE. Nonetheless, neither PRP nor CMB exhibit takeover behavior within the predication dimension. Finally, we note that SPE optionally takes over in the response unit for all predication forms, although not for the locative argument of LOC forms and the property of PRP. In sum, CMB appears more restricted distributionally than either PRP or LOC, PRP is more restricted than LOC, and LOC is the least restricted in its occurrence across interrogative frames. This overall pattern of distribution supports the relative ranking among BE forms expressing predication that is represented in our hierarchy.

LOC < PRP < CMB < SPE < EID

5. Conclusion

Preceding sections have demonstrated that across basic construction types Emai's BE forms exhibit varying degrees of distributional stability. This asymmetry of distri-

bution affects predication and identification forms, although not all are affected to the same degree. Some forms are more restricted than others across query-response frames. The resulting conditions of distributional stability appear proportionally related to the temporal stability of BE form types ranked within and across predication and identification. The left-to-right arrangement of forms stands in a proportional relation to the temporal stability of these same forms. Indeed, it seems that a temporal stability scale respecting the dimensions predication and identification may underlie the behavioral stability of BE forms in Emai. If so, the characterization of language forms and constructions in terms of a temporal underpinning may play a significant role alongside spatial underpinning (Anderson 1971; Lyons 1979). For our part, however, we hope that the tentative hierarchy proposed here might prove useful in the typological characterization of BE constructions in other languages, especially those in Africa and the Edoid family.

Notes

1. Data incorporated in this paper were collected as part of research support to the first author from the National Science Foundation, BNS #9011338 and SBR #9409552, and Southern Illinois University Edwardsville as well as assistance to the second author from an Alexander von Humboldt Fellowship. We thank these institutions for their generous support, while not extending to them any responsibility for our interpretation of the data. An earlier version of this paper was presented at the International Symposium: Typology of African Languages sponsored by the Institut für Afrikanistik, Universität zu Köln, from May 21–24, 2001. We would like to thank symposium participants for their comments on that earlier version.
2. Orthographic conventions for Emai are consistent with those in Schaefer (1987), where *ɔ* represents a lax mid back vowel, *ɛ* a lax mid front vowel, and *vb* a voiced bilabial approximant. High tone is marked by an acute accent, low tone by a grave accent, and high downstep by an acute accent followed by an apostrophe.
3. Abbreviations used throughout this study are the following: EMP=Emphatic, LCT=Locative, MAN=Manner, NEG=Negative, NF=Negative Focus, PF=Positive Focus, s=singular, sc=Subject Category.

References

- Anderson, J. (1971). *The Grammar of Case: Towards a Localistic Theory*. Cambridge: Cambridge University Press.
- Bendor-Samuel, J. (Ed.), (1989). *The Niger-Congo Languages*. New York, NY: University Press of America.
- DeBose, C., & Faraclas, N. (1993). An Africanist approach to the linguistic study of black English: Getting to the roots of the tense-aspect-modality and copula systems in Afro-American. In S. Mufwene (Ed.), *Africanisms in Afro-American Language Varieties* (pp. 364–387). Athens, GA: The University of Georgia Press.

- Declerck, R. (1988). *Studies in Copular Sentences, Clefts and Pseudo-Clefts*. Leuven: Leuven University Press.
- Declerck, R. (1991). A taxonomy of copular sentences: A reply to Keizer (1990). *Linguistics*, 29, 521–536.
- Ellis, J., & L. Boadi (1969). ‘TO BE’ in Twi. In J. N. M. Verhaar (Ed.), *The Verb Be and its Synonyms*, Part 4 (pp. 1–71). Dordrecht: Reidel.
- Frawley, W. (1992). *Linguistic Semantics*. Hillsdale, NJ: Lawrence Erlbaum.
- Givón, T. (1984). *Syntax: A Functional-typological Introduction*. Amsterdam: John Benjamins.
- Keizer, M. E. (1990). A typology of copular sentences. *Linguistics*, 28, 1047–1060.
- Keizer, M. E. (1991). A typology of copular sentences: Response to Declerck. *Linguistics*, 29, 537–542.
- Lyons, J. (1979). Knowledge and truth: A localistic approach. In D. Allerton et al. (Eds.), *Function and Context in Linguistic Analysis* (pp. 111–141). Cambridge: Cambridge University Press.
- Schaefer, R. P. (1987). *An Initial Orthography and Lexicon for Emai: An Edoid Language of Nigeria*. Bloomington, IN: Indiana University Linguistics Club.
- Schaefer, R. P., & F. O. Egbokhare (1999). *Oral Tradition Narratives of the Emai People*, Part I and II. Hamburg: LIT Press.
- Schaefer, R. P., & F. O. Egbokhare. In preparation. *An Emai-English Dictionary*. Berlin: Mouton de Gruyter.
- Stassen, L. (1997). *Intransitive Predication*. Oxford: Clarendon Press.
- Verhaar, J. W. M. (Ed.). (1967–1973). *The Verb “BE” and its Synonyms*. Parts 1–5. Dordrecht: Reidel.
- Welmers, W. (1973). *African Language Structures*. Los Angeles, CA: University of California Press.

Intrinsic focus and focus control in two varieties of Hausa

H. Ekkehard Wolff

University of Leipzig

This paper takes another look at both *intrinsic focus* properties of certain conjugational paradigms of verbs and at language-specific manifestations of *pragmatic control* and *grammatical control* in the domain of *assertive predication focus*. Or, in terms which are more familiar to specialists on African languages, the paper deals with dichotomized sub-systems which are made up of so-called “general” vs. “relative” aspects/tenses. Such sub-systems have been reported to exist in several African languages across the continent irrespective of their genealogical affiliation. They, therefore, provide an interesting field for cross-linguistic typological comparison.

Introduction

While the present scope of our ongoing research on the typology of predication focus in African languages in general and in Chadic languages in particular is much wider,¹ the present paper is solely concerned with one particular Chadic language: Hausa. However, this language will be represented by two of its several varieties, i.e. Standard Hausa (SH) and Damagaram Hausa (DH) which is a non-Standard variety as partially described in Attouman (2000). Generally speaking, there is little in-depth linguistic research on language-internal dialectal variation published for even the most widely spoken African languages including Hausa. Further, when working with so-called Standard varieties of languages as represented in their major reference works, one tends to overlook the vast array of possible forms which non-Standard varieties display and which the major reference works do not always account for. “Standard” varieties are by definition and usage more normative than natural, often incomplete and somewhat artificial, as opposed to natural non-Standard varieties. The two Hausa varieties dealt with in this paper, for instance, display a quite surprising degree of variation in terms of focus control in the domain of assertive predication focus.

The paper will refrain from a detailed theoretical discussion of the in some quarters still controversial notion and scope of “focus” and restricts itself, within its theoretical frame, to the presentation of results and illustrative examples which are particu-

larly interesting from a typological point of view. For the paper's underlying theoretical framework, the reader is referred to Dik (1989) and Hyman & Watters (1984).

1. The pragmatics of focus

In a classic paper on the topic, Hyman & Watters (1984) have sketched out the existence and distribution in Africa (as far as their language sample allowed) of a category which they refer to as "auxiliary focus", which intimately interacts with the inflectional verbal morphology of a given language. I will use matching definitions by Hyman & Watters (1984) and Dik (1989) as starting point for our discussion.

The focal information in a linguistic expression is that information which is relatively the most important or salient in the given communicative setting, and considered by S[peaker] to be the most essential for A[ddressee] to integrate into his pragmatic information. (Dik 1989:277ff.) Focus relates to that information in an utterance which the speaker believes, assumes, or knows that the hearer does not share with him/her.

... the focused part of an utterance ... is said to be *asserted*, or is the *assertion*, while the out of focus part, i.e. which the speaker "believes, assumes, or knows the hearer shares with him/her", is said to be *presupposed*, or simply, the *presupposition* of that utterance. (Hyman & Watters 1984:237)

Parameters of relevance within the underlying theory of focus for this study are the following distinctions based on Hyman & Watters (1984: 239):

... we shall distinguish between "assertive focus" vs. "contrastive focus" as broad categories ... *Assertive focus* can be defined as asserted information projected against neutral background. By using the term "neutral background", we mean that the "slot" occupied by the focused element(s) is judged by the speaker not to have been assigned any *conflicting* value by the listener. By "value" is meant any referent, verb action or state, truth value etc.

Hence the following graphic representation:

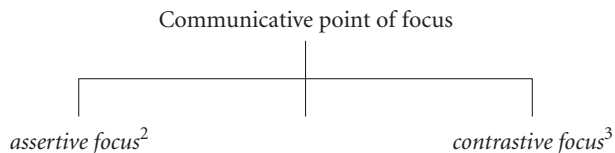


Figure 1.

In this paper I will show that and how the distinction between these two different kinds of focus, hitherto completely overlooked in traditional Hausa studies, operate in the grammar of at least some varieties of Hausa.

Another important parameter is the *scope* of focus. In Figure 2 I use labels taken from Güldemann (1996: 159):

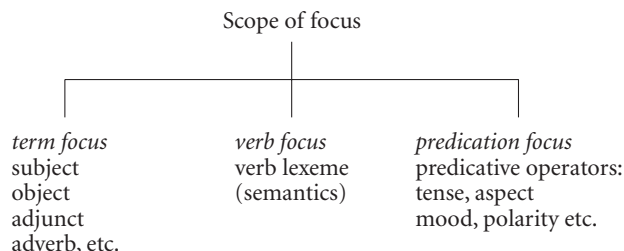


Figure 2.

1.1 Predication focus

Predication focus has become a focal point of interest in African linguistics since Hyman & Watters' seminal paper of 1984 who refer to it as "auxiliary focus" (a label which I consider slightly misleading). However, the category as such had been addressed prior to that, implicitly or even explicitly, in some grammatical descriptions of African languages albeit under quite different labels and hardly in a theoretically satisfying way.

1.1.1 *Focus control: Pragmatic vs. grammatical*

Hyman & Watters (1984:242ff.) introduce yet another important parameter into the typology of focus constructions, i.e. "control of focus". This control has two manifestations (emphasis mine):

pragmatic control of focus: the *speaker* determines the element(s) on which the grammar will express focus.

grammatical control of focus: the *grammar* determines how the speaker will express focus.

This distinction highlights the observation that, in some languages and for at least some constructions, the speaker is free to choose between constructions or verb forms marked for [+Focus] and [-Focus]. In other languages or other constructions within the same language, the speaker has no choice, the grammar itself enforces the choice between [+Focus] or [-Focus] marked constructions or verb forms.

Both instances of control operate in Hausa, but differently in the two varieties under description.

1.2 Intrinsic focus

Observations as illustrated in (1) and (2) below for the Hausa PERFECT and PROGRESSIVE can be referred to as “intrinsic focus” (Hyman & Watters 1984:259ff.), meaning that certain inflectional categories of verbs (predicative operators like aspects, tenses, moods etc.) show an affinity to predication focus. This affinity may show up in the following ways:

Like in Hausa, an inflectional category may display a particular *in-focus* vs. *out-of-focus* morphology; cf. Hausa PERFECT *sun/sukà*, PROGRESSIVE *sunàa/sukèè* under (1) and (2) below.

Like in Gwari (cf. Hyman & Watters 1984 – again it is the category of PERFECT), an obligatory maker must be used which, in Non-Perfect constructions, is characteristic for *in-focus* morphology. This is taken to be sign of intrinsic [+Focus] qualities of the inflectional category in question.

Or, the occurrence of some inflectional categories may be restricted to exactly those pragmatically or grammatically defined environments, in which characteristically *out-of-focus* morphology is used (e.g. in relative clauses, or because other constituents, like question words etc. are focused, i.e. where we have instances of *term focus* of some sort). This is taken to be a sign of intrinsic [–Focus] qualities of the inflectional category in question.

This paper looks at the relationship between focus control and intrinsic focus in two varieties of Hausa.

2. Predication focus in Hausa

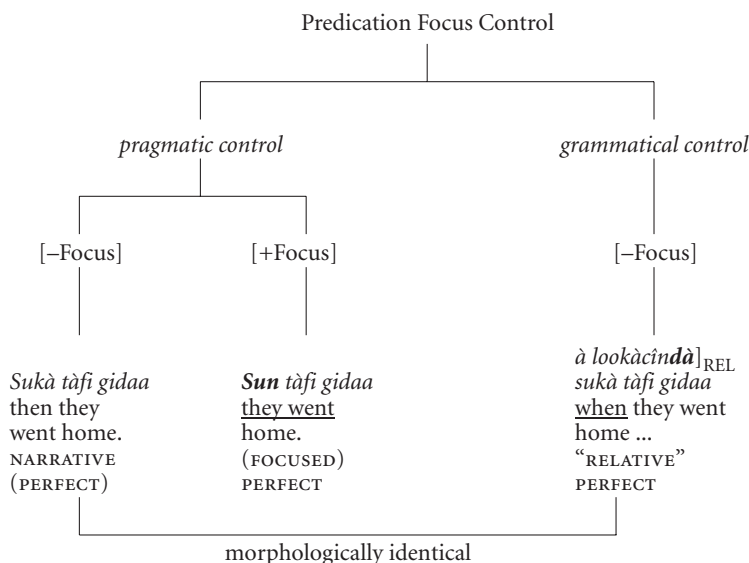
To the best of my knowledge and apart from the somewhat sketchy treatment in Hyman & Watters (1984), predicative focus has not yet been made the subject of in-depth linguistic study of Hausa, neither Standard Hausa, nor in a diachronic or pan-dialectal perspective, nor have the distinctions between assertive and contrastive focus nor between pragmatic and grammatical control.⁴ The present paper attempts to remedy the situation.

2.1 Standard Hausa (Nigeria)

The situation in Standard Hausa as reflected in the most recent reference grammars of the language (Wolff 1993; Newman 2000) can be illustrated as follows. Aspect and tense are encoded in mostly bi-morphemic units with morphophonologically fused subject-pronoun + aspect/tense markers which are pre-posed to the verb. Note that the verb, but only in the case of the PROGRESSIVE, must take a particular shape referred to as verbal noun.

2.1.1.1 *The Standard Hausa PERFECT*

(1) PERFECT



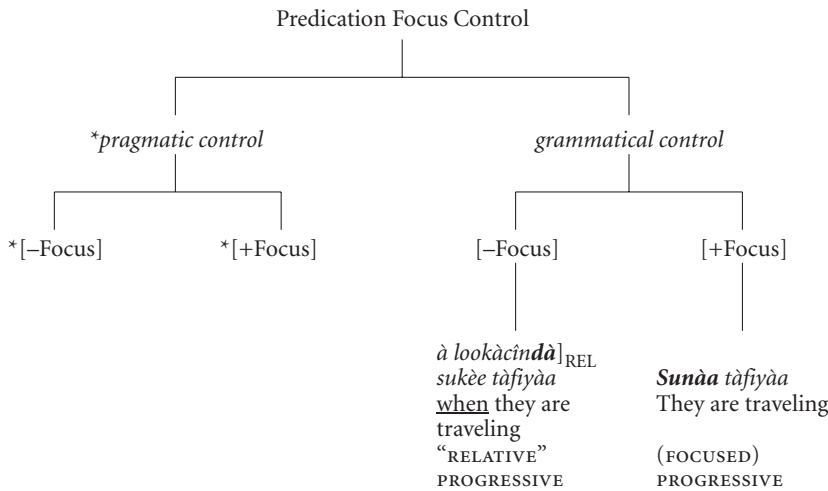
Summary of descriptive facts – PERFECT (Standard Hausa):

1. For SH, we introduce the threefold distinction between NARRATIVE PERFECT,⁵ RELATIVE PERFECT,⁶ and FOCUSED PERFECT.⁷
2. Synchronically and generally, the out-of-focus pre-verb marker *sukà* (for 3rd pl), for instance, may be used by the speaker to mark NARRATIVE PERFECT, when he is also pragmatically free to use the in-focus pre-verb marker *sun* (for 3rd pl) for FOCUSED PERFECT.
3. However, when “preceded by the relativizer *dà* ‘that’, a question word or whomever-type expression (which inherently carry focus), or any other focused element” (Newman 2000:567), the out-of-focus pre-verb marker *sukà* (for 3rd pl) must be chosen to mark the inflectional category which is then referred to as RELATIVE PERFECT.
4. The RELATIVE PERFECT must be chosen under the same conditions as the RELATIVE PROGRESSIVE (cf. below) is used.

2.1.1.2 *The Standard Hausa PROGRESSIVE*

With regard to the inflectional category of PROGRESSIVE, it appears that the similar dichotomy between “general” and “relative” (in traditional Hausaist terms) is exclusively grammatically controlled, i.e. the speaker has no choice in terms of pragmatic control:

(2) PROGRESSIVE



Summary of descriptive facts – PROGRESSIVE (Standard Hausa):

1. For SH, we introduce a distinction between RELATIVE PROGRESSIVE,⁸ and FOCUSED PROGRESSIVE.⁹
2. Synchronically and generally in sentences with verbal predicates, when the speaker chooses the inflectional category of PROGRESSIVE, he must use the in-focus pre-verb marker *sunàa* (for 3rd pl; FOCUSED PROGRESSIVE), unless any of the grammatical conditions which require the RELATIVE PROGRESSIVE are met.
3. When “preceded by the relativizer *dà* ‘that’, a question word or whoever-type expression (which inherently carry focus), or any other focused element” (Newman 2000: 567), the out-of-focus pre-verb marker *sukèè* (for 3rd pl) must be chosen to mark the inflectional category which is then referred to as RELATIVE PROGRESSIVE.
4. The RELATIVE PROGRESSIVE must be chosen under the same conditions as the RELATIVE PERFECT is used.

2.2 Predication focus control (SH)

We summarize the situation in SH with reference to our notion of focus control with regard to predication focus:

1. With PERFECT, pragmatic control of predication focus is allowed in certain constructions (i.e. main clause choice between NARRATIVE PERFECT and FOCUSED PERFECT), whereas grammatical control enforces the out-of-focus forms (RELATIVE PERFECT) in other constructions;
2. with PROGRESSIVE, pragmatic control is disallowed, it is exclusively grammatical control which governs the distribution of the in-focus- and out-of-focus forms exactly along the same lines as with grammatically controlled (RELATIVE) PERFECT.

3. In other words and possibly interesting from a universal typological point of view, FOCUSED PROGRESSIVE is the default category when the conditions for the RELATIVE PROGRESSIVE are not met, while with PERFECT there is no default category, i.e. pragmatic control is unavoidable.

Note that this analysis as, for instance, represented in Wolff (1993) differs from traditional descriptions and teaching materials in which the pragmatic control of the perfect/narrative distinction tends to be overlooked and both PERFECT and PROGRESSIVE, in a grossly oversimplifying manner, are insinuated to represent parallel subsystems with complete complementary distribution of “general” and “relative” forms.

2.2.1 *Intrinsic predication focus (SH)*

In terms of intrinsic focus, we can now generalize for Standard Hausa and provide a historical explanation for the non-symmetry between PERFECT and PROGRESSIVE.

1. We observe, first of all, the absence of symmetry in the two categories of inflectional verb morphology which are subject to intrinsic focus, i.e. PERFECT and PROGRESSIVE:¹⁰

(3) Intrinsic predicative focus (SH)

[+Focus] = in-focus form: “General” Perfect (I), “General” Progressive (I)

[−Focus] = out-of-focus form: Narrative (Perfect), “Relative” Perfect (II), “Relative” Progressive (II)

PERFECT	<i>pragmatic control</i>
[+ assertive PREDICATION focus]	
main clause	
“General” Perfect	<i>Sun tàfi gidaa.</i> They went home.
[− assertive PREDICATION focus]	
main clause	
Narrative (Perfect)	<i>Sukà tàfi gidaa.</i> Then they went home.
<i>grammatical control</i>	
[− assertive PREDICATION focus]	
relative clause	
“Relative” Perfect	<i>(à lookàcìndà) sukà tàfi gidàa</i> (when) they went home

PROGRESSIVE	<i>grammatical control</i>
<hr/>	
[+ assertive PREDICATION focus]	
main clause	
<i>Progressive I</i>	<i>sunàa tàfiyàa</i> they are travelling
<hr/>	
[– assertive PREDICATION focus]	
relative clause	
<i>Progressive II</i>	<i>(à lookàcìndà) sukèe tàfiyàa</i> (while) they were traveling
<hr/>	

2. In Standard Hausa, both PERFECT and PROGRESSIVE have intrinsic focus; further, both categories show complementary in-focus- and out-of-focus forms.
- (4) In-focus- and out-of-focus forms (SH)

<hr/>		
PERFECT		
	in-focus	out-of-focus
<hr/>		
1.sg.c.g.	<i>naa</i>	<i>na</i>
2.sg.m.	<i>kaa</i>	<i>ka</i>
f.	<i>kin</i>	<i>ki-kà</i>
3.sg.m.	<i>yaa</i>	<i>ya</i>
f.	<i>taa</i>	<i>ta</i>
1.pl.c.g.	<i>mun</i>	<i>mu-kà</i>
2.pl.c.g.	<i>kun</i>	<i>ku-kà</i>
3.pl.c.g.	<i>sun</i>	<i>su-kà</i>
impers.	<i>'an</i>	<i>'a-kà</i>
<hr/>		
PROGRESSIVE		
	in-focus	out-of-focus
<hr/>		
1.sg.c.g.	<i>i-nàa</i>	<i>na-kèe</i>
2.sg.m.	<i>ka-nàa</i>	<i>ka-kèe</i>
f.	<i>ki-nàa</i>	<i>ki-kèe</i>
3.sg.m.	<i>ya-nàa</i>	<i>ya-kèe</i>
f.	<i>ta-nàa</i>	<i>ta-kèe</i>
1.pl.c.g.	<i>mu-nàa</i>	<i>mu-kèe</i>
2.pl.c.g.	<i>ku-nàa</i>	<i>ku-kèe</i>
3.pl.c.g.	<i>su-nàa</i>	<i>su-kèe</i>
impers.	<i>'a-nàa</i>	<i>'a-kèe</i>
<hr/>		

3. The Standard Hausa PROGRESSIVE is inherently marked for [α Predication Focus], its in-focus- and out-of-focus forms are in full complementary distribution.

4. The Standard Hausa PERFECT, however, appears at least historically, not to be inherently marked for [α Predication Focus] and therefore did not, at least in diachronic perspective, originally distinguish in-focus- and out-of-focus forms. I.e., the reconstructable system from which Standard Hausa arose can be assumed to have been insensitive to any kind of control, be it grammatical or pragmatic in nature. The only existing form would have been the source of the modern out-of-focus form (*su-kà tàfi* in the examples above).¹¹
5. However, as was already argued for convincingly in Newman & Schuh (1974), a “new” in-focus PERFECT form was innovated at one point in time in the linguistic history of pre-Standard Hausa dialects, not surprisingly making use of the then “independent” set of personal pronouns, represented by *sun* (for 3rd pl) in the in-focus examples above.¹²

3. Damagaram Hausa (Niger)

Recently, Attouman (2000) provided a detailed account of his own native Damagaram Hausa (*Damagaranci*) which allows us to take the analysis of intrinsic focus and both grammatical and pragmatic control in Hausa even further. It appears from Attouman’s description that at least some non-Standard Hausa varieties have streamlined and generalized their aspectual subsystems in three ways which are quite interesting in terms of universal typology and the dynamics of change:¹³

1. the system is extended beyond PERFECT (*accompli*) and PROGRESSIVE (*inaccompli*) to also include “FUTURE” (*visée*);
2. all three categories now allow pragmatic control, in addition to grammatical control;
3. meaning that predicative focus is allowed to co-occur with other types of focus (i.e. term focus indicated by fronting the term-in-focus to clause-initial position, and interrogatives).

Note that, contrary to Attouman’s data and analysis and with regard only to Standard Hausa, in the traditional reference works the two “futures” have not been systematically treated as pairing up with PERFECT and PROGRESSIVE in terms of a “general” vs. “relative” dichotomy!¹⁴ As a matter of fact, their semantics and distributional overlap remain much of a mystery until today. This is also reflected in the rich choice of descriptive labels given to them by various authors,¹⁵ and the largely open question whether one or both represent “tenses” or “aspects”, or even “moods”.¹⁶ For easy reference and quite *contre coeur*, however, I shall henceforth refer to the two conjugational paradigms in question as “FUTURE”.

3.1 Triple in-focus vs. out-of-focus system in Damagaram Hausa

The Hausa variety described by Attouman shows a neat triple structure with regard to in-focus and out-of-focus forms:

- (5) Triple structure with regard to in-focus and out-of-focus forms (DH)

PERFECT (<i>accompli</i>)		
	in-focus	out-of-focus
1.sg.c.g.	<i>naa</i>	<i>na</i>
2.sg.m.	<i>kaa</i>	<i>ka</i>
f.	<i>kin</i>	<i>ki-kà</i>
3.sg.m.	<i>yaa</i>	<i>ya</i>
f.	<i>taa</i>	<i>ta</i>
1.pl.c.g.	<i>mun</i>	<i>mu-kà</i>
2.pl.c.g.	<i>kun</i>	<i>ku-kà</i>
3.pl.c.g.	<i>sun</i>	<i>su-kà</i>
impers.	<i>'an</i>	<i>'a-kà</i>
PROGRESSIVE (INACCOMPLI)		
	in-focus	out-of-focus
1.sg.c.g.	<i>i-nàa</i>	<i>na-kèe</i>
2.sg.m.	<i>ka-nàa</i>	<i>ka-kèe</i>
f.	<i>ki-nàa</i>	<i>ki-kèe</i>
3.sg.m.	<i>ya-nàa</i>	<i>ya-kèe</i>
f.	<i>ta-nàa</i>	<i>ta-kèe</i>
1.pl.c.g.	<i>mu-nàa</i>	<i>mu-kèe</i>
2.pl.c.g.	<i>ku-nàa</i>	<i>ku-kèe</i>
3.pl.c.g.	<i>su-nàa</i>	<i>su-kèe</i>
impers.	<i>'a-nàa</i>	<i>'a-kèe</i>
"FUTURE" (VISÉE)		
	in-focus	out-of-focus
1.sg.c.g.	<i>n-àà</i>	<i>zâ-n</i>
2.sg.m.	<i>k-àà</i>	<i>zaa-kà</i>
f.	<i>k-yàà</i>	<i>zaa-kì</i>
3.sg.m.	<i>y-àà</i>	<i>zâ-y</i>
f.	<i>t-àà</i>	<i>zaa-tà</i>
1.pl.c.g.	<i>m-àà</i>	<i>zaa-mù</i>
2.pl.c.g.	<i>kw-àà</i>	<i>zaa-kù</i>
3.pl.c.g.	<i>s-àà</i>	<i>zaa-sù</i>
impers.	<i>'àà</i>	<i>zaa-'à</i>

The following examples are provided by Attouman but have been re-arranged to match our proposed systematic for Standard Hausa (cf. above). We will attempt some generalizations concerning co-occurrence restrictions with various focus types (assertive and contrastive), and also with regard to topicalization which, although semantically quite different from focus, interacts in a systematic way with predication focus.

(6) Systematic of focus control in DH

[+Focus] = in-focus form: *accompli I*, *inaccompli I*, *visée I*

[−Focus] = out-of-focus form: *accompli II*, *inaccompli II*, *visée II*

PERFECT	<i>pragmatic control</i>
[α assertive PREDICATION focus] main clause:	
<i>accompli I</i>	Àlì yáa zóo Ali came/has come
<i>accompli II</i>	Àlì yá zóo Ali came
& in combination with assertive TERM focus:	
<i>accompli I</i>	á káhwà yáa jée áykìi (it's) by foot he went to work
<i>accompli II</i>	á káhwà yá jée áykìi (it's) by foot he went to work
PERFECT	<i>grammatical control</i>
[− assertive PREDICATION focus] relative clause:	
<i>accompli II</i>	góonâl dà Léekò yá nóomàa the farm which Leko cultivated
interrogative:	
<i>accompli II</i>	wàa yá jée gárkáa? who has gone to the garden?
contrastive TERM focus (with/out topicalization):	
<i>accompli II</i>	Àlì, shíi yá zóo (as for) Ali, he (is the one who) came
[+ assertive PREDICATION focus] topicalization:	
<i>accompli I</i>	Àlì, yáa zóo (as for) Ali, he came Àlì, shíi, yáa zóo (as for) Ali, (as for) him, he came/has come

PROGRESSIVE ¹⁷	<i>pragmatic control</i>
<hr/>	
[α assertive PREDICATION focus] main clause	
in combination with assertive TERM focus:	
<i>inaccompli I</i>	<u>á káhwà</u> yá nàa zúwàa áykii (it's) by foot he's going to work
<i>inaccompli II</i>	<u>á káhwà</u> yá kèe zúwàa áykii (it's) by foot he's going to work
<hr/>	
PROGRESSIVE	<i>grammatical control</i>
<hr/>	
[+ assertive Predication focus] main clause:	
<i>inaccompli I</i>	Àlì yá nàa zúwàa Ali is coming
<hr/>	
[– assertive PREDICATION focus] relative clause:	
<i>inaccompli II</i>	<u>góonàl dà</u> Léekò yá kèe nóomàa the farm which Leko cultivates
interrogative:	
<i>inaccompli II</i>	<u>wàa</u> yá kèe zúwàa gárkáa? who is going to the garden?
contrastive TERM focus (with/out topicalization):	
<i>inaccompli II</i>	Àlì, <u>shii</u> yá kèe zúwàa (as for) Ali, he (is the one who) is coming
<hr/>	
“FUTURE”	<i>pragmatic control</i>
<hr/>	
[α assertive PREDICATION focus] main clause:	
<i>visée I</i>	Léekò yàà jée gárkáa Leko will go to the garden
<i>visée II</i>	Léekò zây jée gárkáa? Will Leko go to the garden?
& in combination with assertive TERM focus:	
<i>visée I</i>	<u>á káhwà</u> yàà jée áykii (it's) by foot he'll go to work
<i>visée II</i>	<u>á káhwà</u> zây jée áykii (it's) by foot he'll go to work
& in combination with topicalization:	
<i>visée I</i>	<u>Àlì</u> , yàà tàhí gárkèe (as for) Ali, he will go to the sheep-farm
<i>visée II</i>	<u>Àlì</u> , zây tàhí gárkèe (as for) Ali, he will go to the sheep-farm
<hr/>	

FUTURE	<i>grammatical control</i>
<hr/>	
[– assertive PREDICATION focus] relative clause:	
<i>visée II</i>	<u>góonàl dà Léekò</u> <i>zây nóomàa</i> the farm which Leko will cultivate
interrogative:	
<i>visée II</i>	<u>wàa</u> <i>zây jée gárkáa?</i> who will go to the garden?
contrastive TERM focus (with/out topicalization):	
<i>visée II</i>	<i>Àlì, shíi</i> <i>zây tàhí gárkèe</i> (as for) Ali, <u>he</u> (is the one who) will go to the sheep-farm <i>Àlì nèe</i> <i>zây tàhí gárkèe</i> (it's) Ali (who) will go to the sheep-farm
<hr/>	

3.2 Re-gaining pragmatic control in Damagaram Hausa

Given a particular type of construction in Hausa, grammatical control is automatic, yet there are strategies to avoid the rigidity of grammatical control and re-gain pragmatic control. Attouman (2000) lists examples which show how this can be achieved in DH. There are at least four strategies available, i.e.

1. introducing *assertive term focus*;
2. interrupting the linear construction with the *insertion* of certain syntactic material, often involving front-shifting;
3. *overtly filling the systematic slots* which are structurally located immediately to the right of the verb (i.e. direct object with transitive verbs, and adverbials with intransitive verbs);
4. changing *scope of interrogative* (generalization yet unclear).

Without claiming completeness of description, Attouman (2000) provides excellent and highly interesting examples along these lines which will be presented here under our entirely different approach in terms of predication focus and focus control. The spread and productivity of these (and possibly other strategies) remain for future research to be ascertained for Standard Hausa and further regional varieties.

Re-gaining pragmatic control by using assertive term focus. As illustrated in (6) above, pragmatic control of predication focus can be re-gained in the main clause by introducing assertive term focus:

1. In the PERFECT, the main clause allows assertive predication focus – both alone and in combination with another assertive term focus.
2. In the PROGRESSIVE, pragmatic control is also allowed in the main clause, as it seems, but only if assertive predication focus combines with another assertive term focus.

3. In the “FUTURE”, pragmatic control is allowed in the main clause for assertive predication focus, which is also allowed to combine with another assertive term focus.

As soon as contrastive focus is involved, on the other hand, grammatical control takes over in DH as well.

Re-gaining pragmatic control by interruption of linear ordering. A subsystem worth investigation also in Standard Hausa is that of pragmatic variability with questions beginning with *dón mée* “why” (cf. Attouman 2000, Section 4) which appear to allow pragmatic control – an observation which, at first glance, would appear to be incompatible with traditional wisdom. Traditional wisdom would state that in questions with interrogatives (like SH *mee* / DH *mù* “what”, as contained in SH *dón mée* / DH *dán mù* “why”) only out-of-focus forms can be used. But, as Attouman shows, certain syntactic devices can be used to re-establish the possibility of pragmatic control which would “normally” be excluded by the interrogative.

Among the syntactic devices to re-gain pragmatic control with *dán miù* questions the interruption of the linear sequence by inserted or front-shifted constituents appears to play a prominent role (examples provided by Attouman for PERFECT only).

- (7) Interrupting the linear construction with an *apposition*:

dán mǐ Dílǎa shíi dáyá yá sàamí sháanúu góomà kúma ...

why (is it) Dila, he alone, (who) got ten cows ...

dán mèi Dílāa shíi dáyá yáa sàamí sháanúu góomà kúma ...

why (did) only Dila, he alone, get ten cows ...

- (8) Interrupting the linear construction with a *vocative/address*:

dán m̀ì, Hásànáá, ákà sáyóo m̀atà díyál wàasáa?

why, Hasana, did one buy her a doll?

dán mì, Hásànáá, án sáyóo mátà díyál wàasáa

why, Hasana, did one buy her a doll?

- (9) Interrupting the linear construction with a fronted *term-focused constituent* (e.g. the indirect object *Hasana* in the following examples)

dán m̀ì Hásànáá ákà sáyóo m̀atà díyál wàasáa...?

why for Hasana (did) one buy a doll ...?

dán m̀ì Hásànáa áń sáyóo m̀atà díyál wàasáa...?

why Hasana for whom one bought a doll ...?

- (10) Interrupting the linear construction with a fronted *topicalized constituent* (e.g. an indirect object)

dán mèi Hásànáá, itá, ákà sáyóo wàà díyál wàasáa?

why (for) Hasana, (for) her, (did) one buy a doll for ?

dán mèi Hásànáá, itá, án sáyóo mātà díyál wàasáa?

why for Hasana, for her, one bought her a doll?

Re-gaining pragmatic control by filling right-of-verb slots. With relative clauses, i.e. the other classical domain of exclusive usage of out-of-focus forms in the traditional wisdom of Hausaists, there are further syntactic devices to regain pragmatic control by the speaker. The general idea is to overtly fill the systematic slots which are structurally located immediately to the right of the verb, i.e. a *direct object* with transitive verbs, and *adverbials* with intransitive verbs.

- (11) For PERFECT, Attouman shows that a resumptive pronoun strategy allows to use predicative focus even within relative clauses (at least with transitive verbs and extracted direct object):

góonâl dà ákà nóomèè ø zàa kà cèè bàà mà̀y itá!
 the farm which one has cultivated (about which) you would say (that it has) no owner!
góonâl dà án nóomèè tá zàa kà cèè bàà mà̀y itá!
 the farm which one has cultivated (and about which) you would say (that it has) no owner!

- (12) Still for PERFECT, an appropriate adverbial group will do the same job with intransitive verbs:

kíihîn dà yá kwán ø ká kèè cîi, bà̀y yí wáarîi bá?
 the fish which was kept over night (and which) you ate, didn't it stink?
kíihîn dà yáa kwán à pìrigóo, mîi gàréè shì?
 the fish which was kept over night in the frig, what's (wrong) with it?

- (13) For PROGRESSIVE, a modal discourse particle like *máa* will restore the possibility of pragmatic control within relative clauses:

lókàcîn dà ká nàa bárcîi máa Léekò yáa hùtá
 also while you were sleeping Leko left
lókàcîn dà ká kèè bárcîi máa Léekò yáa hùtá
 also while you were sleeping Leko left

- (14) For "FUTURE", the same situation as in the PERFECT appears to hold, i.e. a resumptive pronoun strategy restores pragmatic control with transitive verbs:

móotàl dà náa gyáaràa tá, zàa kà kí áràa míni!
 the car which I will repair, you will refuse to let me borrow it!
móotàl dà zàn gyáaràa tá,
 (it's) the car which I will repair (and which)
zàa kà kí áràa míni!
 you will refuse to let me borrow!

Re-gaining pragmatic control by scope of interrogative (?). The following set of examples shows neither syntactic interruption of the linear construction nor the filling of any right-of-verb slots. Rather, two different readings could be linked, as it will

appear from the just one example available in Attouman (2000), to the scope of the interrogative itself:

- (15) *dán mì ákà sáyóo wàà Hàsànáa díyál wàasáa bà à sáyóo wàà Húsàynáa bá?*
 why not did one buy a doll for Hassana (and) buy (one) for Husayna?

dán mì án sáyóo wàà Hàsànáa díyál wàasáa bà à sáyóo wàà Húsàynáa bá?
 why did one buy a doll for Hassana (and) did not buy (one) for Husayna?

4. Conclusion

This paper aimed to show the empirical relevance of some theoretical notions about the nature of focus in Hausa and, by implication and to be elaborated upon at another occasion (cf. Wolff 2003), in Chadic and African languages in general. This was done in view of the potential of studying and exploiting, variants of focus constructions for cross-linguistic typological comparisons.

It has been shown that, beyond the first results of the valid pioneer analysis provided by Hyman & Watters (1984) for Standard Hausa,

1. both *assertive* and *contrastive focus* operate as distinct semantico-syntactic categories in Hausa;
2. both PERFECT and PROGRESSIVE have *intrinsic focus* properties in terms of *assertive predication focus*;
3. *assertive predication focus* is a valid and dynamic semantico-syntactic operation with some history and variance in modern Hausa varieties including Standard Hausa;
4. there are two different control mechanisms at work in the language, i.e. speaker's choice (i.e. *pragmatic control*) and system-internal *grammatical control*;
5. Hausa varieties tend to differ in terms of presence or absence of automatic *default* choices in terms of predication focus in main clauses; default choice appears to be only active for the PROGRESSIVE;
6. in the case of the PERFECT, assertive predication focus appears to be diachronically more recent in the language, at its origin lies the grammaticalization of a "new" FOCUSED PERFECT – as opposed to the older NARRATIVE PERFECT – which made use of the Old Hausa independent Personal Pronoun set (i.e. *mun* for 3rd pl); this diachronic development explains the lack of symmetry in Standard Hausa between the usages of the in-focus- and out-of-focus forms for the PERFECT and the PROGRESSIVE;
7. innovative Non-Standard Hausa varieties may show a tendency to analogically restructure their grammatical systems with regard to encoding information structure by increasing the respective pragmatic domains, e.g. allowing pragmatic control not only in the PERFECT, but also in the PROGRESSIVE and "FUTURE";

8. it is possible for Hausa speakers to circumvent automatic grammatical control and re-gain pragmatic control by using various semanto-syntactic strategies;
9. the significant distinction between pragmatic and grammatical control in Hausa, as in many other Chadic languages, is not restricted to issues of information structure, i.e. various dimensions of focus, but also operates in the domain of verbal plurality which, in Chadic, tends to cross the borderline between inflectional and derivational morphology in both directions – but that is yet another story.

Notes

1. The paper takes up an issue previously touched upon in an unpublished paper put together by Tom Güldemann, René Kriegler, H. Ekkehard Wolff under the title *Prädikationsfokus als Flexionskategorie in afrikanischen Sprachen* which was presented as a work-in-progress report to the 2nd Colloquium of the *DFG-Schwerpunktprogramm 'Sprachtypologie'* (Günzburg, October 3–5, 1997). I gladly acknowledge the financial support of the project *Strukturell-typologische Parameter der verbalen Flexion in afrikanischen Sprachen* by the DFG. I would also like to acknowledge the particularly valuable input provided by René Kriegler (as part of the research team) and Tom Güldemann (as external consultant) in the course of that project.

2. I prefer Hyman & Watters' label "assertive" focus over Dik's term "completive" focus albeit both mean basically the same:

This type of Focus requests or presents new information pertaining to an information gap on the part of S[peaker]; there is no contrast involved with any kind of similar information. (Dik 1989:282)

3. "Contrastive" focus has a variety of subtypes as outlined, for instance, by Dik (1989). Our use of this label is probably uncontroversial in all quarters and could be generalized in Dik's words:

All other Focus types . . . involve some contrast between the Focus constituent and alternative pieces of information which may be explicitly presented or presupposed. (1989:282)

4. Even the most recent and authoritative reference grammar (Newman 2000:187–195) restricts the discussion of focus in Hausa to *contrastive* focus, thereby missing to address the topic of *assertive* focus under which predicative focus would have to be treated.

5. This conjugational paradigm has been treated in the literature under various labels (cf. Wolff 1993:423): *Historicus*, *Preterite*, *Past*, *Aorist*, *Aorist II*, *Relative Perfect*, *Relative Completive*, *Relative Past*, *Perfect II*, *Accompli II*.

6. Treated in the literature as being "the same" as the Narrative perfect (cf. Note 5).

7. This conjugational paradigm has been treated in the literature under various labels (cf. Wolff 1993:422): *General Past*, *Aorist I*, *Perfect (I)*, *Completive*, *Accompli I*.

8. This conjugational paradigm has been treated in the literature under various labels (cf. Wolff 1993:425) such as *Relative Continuous*, *Continuous II*, *Inaccompli II*.

9. This conjugational paradigm has been treated in the literature under various labels (cf. Wolff 1993:425) such as *General Continuous*, *Continuous I*, *Inaccompli I*.

10. Note again that the occurrence of the *out-of-focus* forms in general is by no means restricted to relative or subordinate clauses, as a cursory glance into the literature and the selective examples below might – quite wrongly – suggest.

11. It is widely considered received wisdom among scholars of Hausa that the **-kà* marker of the out-of-focus “relative” perfect (II) forms reflects an old and reconstructable Chadic “perfect” marker (cf. Newman & Schuh 1974 who were probably the first to suggest this).

12. This analysis is supported by the observation that the pronominal paradigm which is used in the “general perfect (I)” does not lend itself to a reanalysis of being constructed as a combination of {person+aspect marker} as compared to the more transparent “relative perfect (II)”, the more so if we compare some north-western non-Standard Hausa forms for the latter in which

- I. the marker **-kà* is more generally used (and if only as reflected in a consonant “C” which automatically assimilates to the following initial consonant of the verb stem), and which
- II. uses the reconstructable independent pronouns, at least in some forms with regional distribution, to attach the marker **-ka* (Wolff 1993:98):

in-focus form PERFECT		out-of-focus form PERFECT	
Standard Hausa		Standard Hausa	Northwestern Hausa dialects
1.sg.c.g. <i>naa</i>	1.pl.c.g. <i>mun</i>	<i>na</i> <i>mu-kà</i>	<i>na-C</i> <i>mun-kà</i>
2.sg.m. <i>kaa</i>	2.pl.c.g. <i>kun</i>	<i>ka</i> <i>ku-kà</i>	<i>ka-C</i> <i>kun-kà</i>
f. <i>kin</i>		<i>ki-kà</i>	<i>ki-C ~ kin-kà</i>
3.sg.m. <i>yaa</i>	3.pl.c.g. <i>sun</i>	<i>ya</i> <i>su-kà</i>	<i>ya-C ~ shin-kà sun-kà</i>
f. <i>taa</i>		<i>ta</i>	<i>ta-C</i>
impers. <i>'an</i>		<i>'a-kà</i>	<i>'an-kà</i>

13. When referring to Attouman’s data and description, for reasons of convenience and distinction of the Hausa varieties, I shall maintain his slightly idiosyncratic labels for the conjugational paradigms involved.

14. Nevertheless, restrictions of distribution have occasionally been noted, yet without placing these observation within a more general framework together with PERFECT and PROGRESSIVE; cf. for instance Newman (2000: 588) who notes that “the potential does not occur in Rel environments; instead, one has to use the normal future (or some other alternative)”. Hyman & Watters (1984: 249 fn.) had already quoted information by Russell Schuh to the same effect, adding, however, that “the [–focus] future construction differs from its two counterparts in its ability to be used in an independent clause” (here they have overlooked the NARRATIVE PERFECT!), and that “another divergence from the pattern is that the two futures have distinct negative forms. In the completive and progressive negatives, the focus distinction is neutralized.” Note also that the limited information which we have on Hausa dialectal variation suggests that variability in the functional range of the “modal future” appears to be fairly common: cf. the “Eventual (I/II)” in Tibiri Hausa and a second future paradigm in Tahoua Hausa (*Aderanci*); cf. Wolff (1993: 104f.).

15. Wolff (1993: 430f.) lists the following functional labels which have been used for SH in the literature:

- I. *Future I, Ingressive, Intentional*;
- II. *Future II, Indefinite Future, Prédicatif, Certative, Potential*.

16. Wolff (1993) takes a clear position on the issue: Future I (formed with the help of a grammaticalized motion verb “go” **zaa* plus suffixed subject-pronouns) is taken as a true tense in the sense of “immediate future”, whereas the sometimes so-called Future II is referred to as

“Potentialis” in the sense of an aspectual category (with, however, strong modal connotations). Newman (2000:584, 586) follows basically the same line of analysis by referring to the conjugational paradigms in question as “future” and “potential” (short for “potential-future”, Newman mentions semantically better motivated alternative labels such as “conditional future” or “modal future”).

17. Note that Attouman (2000) does not provide any example for the use of topicalization and Progressive. I cannot say at this point in time whether this an accidental gap, although I would assume it to be. (In SH, for instance, there would be no restriction against using the Progressive with topicalized constituents.)

References

- Attouman, M. B. (2000). L'opposition occurrence ouverte – occurrence fermée à travers le système aspectuel du hawsa. In H. E. Wolff & O. D. Gensler (Eds.), *Proceedings of the 2nd World Congress of African Linguistics, Leipzig 1997* (pp. 427–457). Köln: Rüdiger Köppe.
- Dik, S. C. (1989). *The Theory of Functional Grammar*. Part I: *The structure of the clause*. Dordrecht: Foris [Functional Grammar Series 9].
- Güldemann, T. (1996). *Verbalmorphologie und Nebenprädikation im Bantu*. Bochum: Universitätsverlag Dr. N. Brockmeyer [Bochum-Essener Beiträge zur Sprachwandelforschung 27].
- Güldemann, T., Kriegler, R., & Wolff, H. E. (1997). Prädikationsfokus als Flexionskategorie in afrikanischen Sprachen. Work-in-progress report presented to the 2nd Colloquium of the DFG-Schwerpunktprogramm 'Sprachtypologie' (Günzburg, October 3–5, 1997). Ms.
- Hyman, L., & Watters, J. R. (1984). Auxiliary focus. *Studies in African Linguistics*, 15, 233–273.
- Newman, P. (2000). *The Hausa Language. An encyclopedic reference grammar*. New Haven, CT: Yale University Press.
- Newman, P., & Schuh, R. G. (1974). The Hausa aspect system. *Afroasiatic Linguistics*, 1, 1–39.
- Wolff, H. E. (1993). *Referenzgrammatik des Hausa*. Münster: Lit (Hamburger Beiträge zur Afrikanistik 2).
- Wolff, H. E. (2003). Predication focus in Chadic languages. In H. Ekkehard Wolf (Ed.), *Papers in Chadic Linguistics*. From the 1st Biennial International Colloquium of the Chadic Languages Family, Leipzig 2001 (pp. 137–159). (Chadic Languages – Linguistique tchadique – Tschadistik Vol. 1.) Cologne: Ruediger Koppe.

Language index

A

Afar 132–134, 145, 304–305
 Agaw 304–305
 Aiki 139–140, 142–143, 145–146
 Alagwa 304–306, 309–310, 312, 313–315, 317–318, 320–322
 Amele 212–213
 Amharic 80, 227–228
 Anyi 46, 52–53
 Anywa 56, 74, 75, 77, 195, 227–228
 Arabic 210, 227–228
 Arbore 30, 34, 160–161, 304–305, 309–310, 312, 313–317–318, 320–321
 Ateso 227–228

B

Baale 77, 84
 Babungo 211–212, 219–220, 227–228
 Bagirmi 227–228
 Baiso 304–305
 Baka 227–228, 362ff.
 Bambara 213–215, 227–228
 Bantu x–xi, 1, 93ff., 215
 Bari 75, 81, 222–223, 227–228
 Basque 59
 Bedauey 138, 142–145
 Beja 304–305
 Ben (Moba) 231
 Birom 141
 Bole 227–228
 Boni 304–305, 307, 309–311, 312, 315, 317–318, 320–321
 Bulgarian 211
 Burji 304–305
 Burunge 34, 304–306, 309–310, 312, 313–315, 317–318, 320–321
 Byali 337ff.

C

Chaha 80
 Chai 74, 87
 Chikuhane 9
 Cibak 217, 227–228
 Cushitic xii, 303ff.

D

Dagbani 227–228
 Dahalo 34, 304–305, 308–310, 312, 317–318, 320–321
 Dan 55
 Datooga 14, 29, 75, 83
 Dawuro 221, 227–228
 Dhaasanac 304–305, 309–310, 312, 313–314, 317–321
 Dholuo 227–228
 Didinga 84
 Dinka 75
 D'irayta 161–162
 Ditammari 337ff.
 Dongola 139, 142, 144–145
 Dullay 304–305, 307, 309–310, 312, 313–314, 317–318, 320–321
 Dzamba 227–228

E

Eastern Sudanic 73
 Egyptian 135
 Elmolo 304–306, 309–310, 312, 313–317–318, 320–321
 Emai xiii, 377ff.
 English 13, 27, 36, 115–117, 362ff.
 Ewe 52, 55
 Ewe (Anglo) 220–221, 227–228

F

French 27, 47, 59, 363, 368
 Fula 46
 Fur 73, 79, 141–143, 227–228
 Fyem 227–228

G

Gamo 155–156, 158, 227–228
 Ganda xi, 101, 107, 171ff., 227–228
 Gciriku 255
 German 15, 29, 31–32, 34, 37, 362ff.
 Gidar 227–228
 Gidole (Dirayta) 304–305
 Gitonga (Inhambane) 3, 5
 Giziga 221–222, 227–228
 Gorwa 304–305
 Greek 59
 Gujarati 27

H

Haddiyya 304–305
 Hadza 74, 87
 Harar Oromo 160, 227–228, 307
 Haro 227–228
 Hausa xiii, 30–32, 68–69, 195, 227–228, 397ff.
 Hausa (Ader) 263, 266, 273
 Haya 101, 107
 Herero 9
 Hmong 38

I

Igbo 54
 Ik xi, 135, 197ff., 227–228
 Ila 99
 Iraqw 73, 304–305, 308–310, 312, 313–315, 317–318, 320–321
 Irob Saho 158, 162
 Italian 27, 29, 50, 363, 368

J

Jacalteco 30, 32
 Japanese 27, 29, 49, 164–165

K

Kaansa 337ff.
Kabba xii, 277ff.
Kabiye xi, 229ff.
Kabyile xi, 263ff.
Kadu 74
Kalenjin 75, 83
Kambaata 304–305
Kana 54
Kanuri 61–62, 140, 142, 145–146, 205
Karachay-Balkar 218
Kimanyanga 327ff.
Kimbundu 95
Kinyarwanda 227–228
Kiowa 28
Kipsikiis 72, 78
Kituba 327–328
Kiyansi xii, 327ff.
Konso 304–305, 307, 309–311, 312, 313–314, 317–318, 320–321
Koozime 227–228
Koranko 227–228
Korean 31–32
Koromfe 227–228
Koyaga 213–214
Koyra 220, 227–228
Koyraboro-Senni 227–228
Kposo 65
Krongo 135–136, 142, 144–146, 227–228
Kuliak 80
Kumari 28–29
Kunama 73, 157, 227–228
Kwangali 251ff.
Kwanyama 251ff.
Kxoe 221, 227–228, 362ff.

L

Lamang 135
Lango 227–228
Latin 54
Lele 227–228
Lezgian 135
Lingala 227–228
Logooli 2, 9
Loma 195–196
Lotuxo 75
Lotuxo-Maa 81
Luba 99, 102, 108
Luo 75

M

Maale xi, 53, 72, 222, 227–228
Maasai 78, 82, 227–228
Maban 73
Macedonian 59
Majang 77, 84, 87
Malayalam 28
Mamvu 227–228
Mandara 195
Mandarin 128–129
Mankono 213
Manyo 255
Marathi 27
Maricopa 211
Mbukushu 255
Mbum 222
Mende 47, 53, 55, 59–60
Mina x, 121ff.
Misantla Totonac 37
Mohave 33–34
Mparntwe Arrernte 28
Murle 77, 84
Mwera 9

N

Nahuatl 47
Nandi 78, 227–228
Nara 156–157, 159, 163, 166
Nawdem 229ff.
Ndonga 104, 108, 251ff.
Nera 137, 142, 144–146
Ngbaka 222
Ngemba xi, 223, 227–228
Ngiti 54
Nguni 3, 135
Nilo-Saharan 73
Nilotic 9, 71ff.
Northern Sotho 4, 5, 8, 100, 227–228
Nubian 73–74, 79
Nuer 77
Nyimang 73–74

O

Ongamo-Maa 75
Oromo 54, 157, 304–305, 309–310, 312, 317–318, 320–321

P

Paez 54
Päri 74, 75, 77, 195–196
Pero 14, 29, 227–228

Persian 213

Piedmontese 50
Pirahã 211
Polish 118–121
Proto-Bantu 5–6

Q

Qafar 156, 158, 162, 166

R

Rendille 304–307, 309–310, 312, 315, 317–318, 320–321
Ronga 103, 109
Ronga (Tsonga) 7

S

Saharan 73, 79
San 263, 266, 273
Sebirwa 5
Sekgalagadi 5
Seme 33
Serbo-Croatian 54
Shambyu 255
Shilluk 74, 75, 195
Shimaore 61
Shona 97–98, 102, 134–135
Sidamo 304–305
Sinhalese 37
Somali x, 13ff., 157, 160, 304–307, 310–311, 312, 313–315, 317–318, 320–321
Songhay (Gao) 54
Sonrai 227–228
Sotho ix, 3
Southern Min 287
Southern Sotho 4–6, 8, 10, 63–64
Spanish 13, 27, 29, 50, 363, 368
Supyire 227–228
Surmic 74, 77, 84
Swahili 2, 9, 48–49, 60–61, 216, 227–228, 264

T

Tahitian 27
Tama 73
Tangale 195
Tashelhiyt 263, 266, 272
Tem xi, 229ff.
Tennet 77, 79, 84–86
Teso-Turkana 81
Thai 28, 38

Tigre 80, 137–138, 142–143, 145

Tirma 87

Tokelauan 28

Tonga 98

Tongan 31, 32

Tsonga 3, 102–103

Tswana 3, 5, 7, 47, 48, 56, 62–64,
109

Tuareg (Ahaggar) 263, 266, 272

Turkana 76, 78, 82, 85, 89

Turkish 212

Tzotzil 30, 32

U

Umbundu 227–228

Uzbek 29

V

Venda 3, 4, 7, 99–100, 227–228

W

Walpiri 212

Wambon 135

Wolaitta 227–228

Wolof 14, 29, 34, 50–51, 57–58,
62–63

X

Xhosa 6, 100

Y

Yaaku 304–305

Yoruba 66, 218–219, 227–228,
263, 266, 273

Z

Zayse 136, 142–143, 145, 155, 158,
165–166

Zulu ix, 4, 6, 8, 100–101,
216–217, 227–228

Name index

A

Abdoulaye, M. L. 31
 Aikhenvald, A. Y. 171, 174, 287, 299
 Alpher, B. 145
 Amador, E. M. 29
 Amborn, H. 307
 Ameka, F. 55
 Amha, A. 53, 72, 222
 Andersen, T. 74, 75, 195
 Anderson, J. 395
 Andrzejewski, B. W. 157
 Appleyard, D. L. 317, 319
 Armbruster, C. H. 131, 139
 Asher, R. E. 28
 Ashton, E. O. 101, 176, 179
 Attouman, M. B. 397, 405, 407, 409–410
 Awobuluyi, O. 218

B

Baker, M. C. 361, 371
 Banti, G. 157
 Barre, M. Cisse 26
 Batibo, H. M. ix, 3
 Beaton, A. C. 141
 Biber, D. 317, 319
 Bisang, W. 38
 Bitima, T. 54
 Black, P. D. 148
 Blake, B. J. 72, 196
 Blanchon, J.-A. 195
 Bloomfield, L. 172
 Boadi, L. 379
 Boeschoten, H. 29
 Böhm, G. 195
 Bosch, S. E. 100–101
 Bouquiaux, L. 141
 Bourdin, P. x, 24, 28, 35
 Broadwell, A. 113
 Burssens, A. 99, 102, 108
 Byarushengo, E. R. 101
 Bybee, J. L. ix, 2, 6, 10

C

Carlson, R. 338, 351, 357
 Caron, B. 263, 266, 273
 Chao, Y. 128
 Chappell, H. 287, 299
 Chumbow, S. B. 222–223
 Claudi, U. 278
 Cohen, D. 131–133, 135
 Cole, D. T. 179
 Collins, B. 98
 Comrie, B. xi, 210–213, 218–219, 284, 287
 Contini-Morava, E. 264
 Craig, C. G. 30, 279, 299
 Crass, J. 131
 Crazzolara, J. P. 77
 Creider, C. 78
 Creider, J. T. 78
 Creissels, D. 213–215
 Culioli, A. 35, 267, 269
 Cutler, A. 143
 Cyffer, N. 62

D

Dahl, O. 2
 Dammann, E. 257
 Danon-Boileau, L. 269
 de Bose, C. 379
 Deutsche Forschungs-gemeinschaft xiii
 Diesel, H. 231, 240–242
 Dik, S. 398
 DiMeola, C. 16
 Dimmendaal, G. J. x, 74, 76, 82, 85–87, 89, 196, 222–223
 Dixon, R. M. W. 171, 174, 195–196, 287
 Doke, C. M. 2–3, 7
 dos Santos, E. 96
 Dubois, J. 1

E

Edgar, J. 139
 Egbokhare, F. O. xiii
 Ehret, C. 74, 304
 Eichinger, L. M. 362
 Eklo, A. A. 65
 Ellis, J. 379
 Emanation, M. 5, 9
 Everett, D. 211
 Eze, E. 54

F

Fagan, B. M. 3
 Faraclos, N. 379
 Fehderau, H. 327
 Ferguson, C. A. 131, 146
 Fillmore, C. J. 16
 Finer, D. 113
 Finlayson, R. 3
 Fivaz, D. 108
 Fleisch, A. x
 Fortune, G. 98
 Frajzyngier, Z. x, 14, 113–114, 118, 195, 217, 222
 Frawley, W. 377

G

Galboran, L. 306–307
 Gemetchu Megerssa 157
 Giacalone, R. A. 280
 Gilley, L. G. 74, 75, 195
 Gilligan, G. 143
 Givón, T. 93, 95, 265, 280, 351, 355, 377
 Gordon, L. 211
 Greenberg, J. H. 74, 82, 89
 Griefenow-Mewis, C. 54
 Groussier, M.-L. 35

- Güldemann, T. 54, 131–132, 135, 143, 399
 Guthrie, M. 3, 5, 251, 258, 260, 333
- H**
 Hagège, C. 222
 Haiman, J. 113–114
 Halliday, M. A. K. 340–341
 Haspelmath, M. 95, 135, 338, 352
 Haviland, J. B. 30
 Hawkins, J. A. 143
 Hayward, R. J. 30, 72, 132–133, 136, 155–156, 160–162, 165, 220, 306, 308, 316–317
 Heath, J. 54, 145
 Hegel, G. W. F. 266
 Heine, B. xiii–ix, 6, 9, 73, 81–83, 149, 160, 205, 215–216, 261, 278, 287, 299, 306–307, 316–317, 322, 334
 Heny, F. 177, 182
 Hetzron, R. 148, 304
 Hill, D. 14, 27, 29
 Himmelmann, N. P. 144
 Hooper, R. 28
 Hopper, P. J. 278, 280
 Horn, L. 266–267
 Howard, K. 28
 Hulstaert, G. 327
 Hünemeyer, F. 278
 Hutchison, J. 140
 Hyman, L. M. xi, 2, 157, 174, 179, 187, 189, 254, 398–400
- I**
 Ikoro, S. M. 54
 Innes, G. 53, 55, 59–60
- J**
 Janson, T. 3
 Jo, In-hee 31
 Junod, H. A. 103
- K**
 Katamba, F. X. xi, 174, 187, 189, 254
 Kawasha, B. 93
 Keenan, E. 94, 219
 Kemmer, S. 109
 Kendall, M. B. 34
 Kibrik, A. 209–10
- Kiessling, R. 304, 306
 Kilian-Hatz, C. xii, 221
 Kimenyi, A. 93
 Klavens, J. L. 178
 Klimov, G. A. 196
 Koch, H. 13
 König, C. 80, 199, 206
 Kooij, J. 306
 Kurcz, I. 118
 Kuryłowicz, J. 35
 Kuteva, T. xi, 215
 Kutsch-Lojenga, C. 54
- L**
 Last, M. 74
 Launey, M. 47
 Lawrenz, B. 362
 Leben, W. 156
 Lebigkaza, K. K. xi, 231
 Leguil, A. 272
 Lehmann, Ch. 35, 210
 Leslau, W. 73, 137, 148
 Levin, B. 339, 352
 Levinson, S. C. 231
 Lewis, M. 220–221
 Li, C. N. 129
 Lichtenberk, Fr. 25, 28
 Lloyd, P. M. 362
 Longacre, R. 132–133
 Louw, F. 3
 Louwrens, L. J. 100
 Lucassen, D. 74
 Lyons, J. 242, 95
- M**
 MacKay, C. J. 37
 Maho, J. 253–254
 Mallinson, G. 196
 Martin, S. E. 27
 Masica, C. P. 72
 Mathews, P. 113
 McCawley, J. 164
 McGregor, W. 145
 McLaren, J. 100
 Meillet, A. 278
 Mettouchi, A. xi, 264, 268–269
 Miller, C. L. 74, 75, 195
 Mithun, M. 35, 113, 362, 370–371
 Moges, Y. 87
 Moreno, M. M. 157
 Moser, R. xii
 Mous, M. xii, 34, 304, 306–309, 315, 322
- Mpaayei, J. Tompo Ole 78
 Mufwene, S. S. xii, 327–328, 332
 Munro, P. 33, 113–114
- N**
 Nespor, M. 173
 Newman, J. 282
 Newman, P. 69, 400, 405
 Nichols, J. 71, 79, 149
 Nicole, J. 357
 Nougayrol, P. 139–140
 Nurse, D. 2
- O**
 O'Neil, J. 97
 Ouwehand, M. 102–103
 Owens, J. 157, 160, 307
- P**
 Palmer, F. R. 131
 Parker, E. M. 162, 165
 Pillinger, S. 306–307
 Platiel, S. 263, 266, 273
 Poser, W. J. 172
 Poulos, G. 99–100, 216–217
 Prasse, K.-G. 269
 Prost, A. 33
- Q**
 Quaireau, A. 46, 53
- R**
 Raible, W. 372
 Rainer, F. 370
 Randal, S. 79, 84–86
 Rappaport Hovac, M. 352
 Raz, S. 137–138
 Reh, M. 56, 74, 75, 195, 278
 Reinisch, L. 136–138, 144
 Reynolds, C. H. B. 37
 Ribeiro, A. 103
 Ricca, D. 14
 Richardson, I. 260
 Riehl, C. xii
 Roberts, J. 212
 Rojas Curieux, T. 54
 Roper, E. M. 138
 Ross, M. 87
 Rottland, F. 14, 75, 83
 Rude, N. 195–196

S

Sachnine, M. 263, 266, 273
 Saeed, J. I. 160–161, 306
 Samb, A. 14
 Sandilands, A. 109
 Sapir, E. 71
 Sasse, H.-J. 148, 311
 Schadeberg, T. C. 52, 150
 Schaefer, R. XIII
 Schaub, W. 211–212, 219–220
 Schuh, R. G. 405
 Schultze-Bernd, E. 145
 Selkirk, E. O. 173
 Smith, E. W. 99
 Spencer, A. 361
 Stassen, L. 379–380
 Stauder, J. 87
 Stevick, E. W. 177, 179
 Stirling, L. 113–114
 Sundström, R. 138

Svolacchia, M. 307, 312
 Sweetser, E. 5

T

Taljaard, P. C. 100–101
 Talmy, L. 17, 287
 Tao, L. 129
 Tchekhoff, C. 31
 Thomas, J. 222
 Thompson, E. D. 136–137
 Thompson, S. A. 129
 Tosco, M. 73, 80, 304, 306, 308,
 310, 314, 319
 Towet, T. 72
 Traugott, E. C. 278
 Tucker, A. N. 78, 177, 179,
 206–207
 Tunbridge, D. 13

V

Voeltz, F. K. E. 217, 338
 Vogel, I. 173

Voßen, R. 75, 81–83
 Vries, L. J. de 135
 Vries-Wiersma, R. de 135
 Vuillaume, M. 16, 31

W

Waley, A. 131
 Walusimbi, L. 174, 187
 Watkins, L. J. 28, 87
 Watters, J. 2, 216, 398–400
 Welmers, W. E. 379
 Westphal, E. O. J. 257
 Whaley, L. J. 361
 Wilkins, D. P. 14, 27, 29
 Williams, W. 219
 Woedemann, H. 221
 Wolff, H. E. XIII, 135, 400,
 412

Z

Zwicky, A. M. 174

Subject index

A

ablative 197, 320
absolutive 74–78, 81
accessibility hierarchy of
 relativization 219
accusative 197, 199
agent x, 96ff., 365, 367
agreement 46ff., 238–239
Aktionsart 15, 16
allative 320
anti-passive 355
aorist 268
apposition 410
aspect 3
ATR harmony 52
auxiliary 4

B

Bauplan 89–90
benefactive 279, 284, 368

C

calendar adverbs 234–235
case x1, 195ff., 319–320
case marking 81
case on conjunctions 197–202
case on verbs 203
categorical misbehavior 196–197,
 205–206
causative 279, 287–291, 294–295,
 298
clitic doubling 48
comitative 97, 320
complex predicate 131ff.
compound auxiliary 8
compounds x11, 361ff.
conditional 7
connectives 297
copula x111, 100–101, 197, 377ff.
copula, class membership 378,
 380ff.

copula, equation 378, 380ff.
copula, location 378, 380ff.
copula, property 378, 380ff.
copula, specification 378, 380ff.
co-reference x

D

dative 197, 203–204, 284–286,
 320
dative clitic 59
deictic center 28
deixis x–x1, 13ff., 16, 229ff., 321
deixis & social functions
 245–247
deixis, directional 27, 29, 34
deixis, discourse 242–243
demonstrative 231, 236–237, 248
dependent-marking 71ff.
diathesis 337ff.
different-subject-constraint 23,
 29, 32–34
different-subject-logic 26
discourse marker 295–297
distal 230–231
ditransitive x11, 58ff.

E

endophoric level 30
ergative 74–77, 81, 195
ethnic names 366
exophoric level 30

F

focus x111, 51, 244–245, 308, 312,
 397ff.
focus, control 399, 411
focus, intrinsic 400, 403–404
focus, predication 399, 402–403
free pronoun 50
future 1–3, 5–6, 9, 280–281, 406,
 409, 411

G

gender 81, 83
generic auxiliary 131ff.
genitive 197
Goddard's law 86
grammaticalization x, 200–202,
 204–205
grammaticalization chain 278ff.

H

head-marking ix, x11, 71ff.

I

imperfective 268, 280–281
inchoativity x1
indexation 62–63, 66
indirect object 281–282
instrumental 82, 84, 99–100,
 320, 365, 369
interrogatives 382ff.
interrogatives, information
 390ff.
interrogatives, polar 387ff.
intransitive passive 341

L

locative 78, 98–99, 320, 368

M

manner 369
modality 2

N

negation x1, 7, 263ff.
nominative 74–78, 81, 197
noun classes 251, 253
noun modifier 256–260

O

object incorporation 355–356
object marking 43ff., 313
object pronouns 318, 329
obligatory contour principle 156
oblique 197
one-high-only-constraint
(OHO) 155ff.
optative 204

P

passive 93ff., 348–351
patient 368
pejoratives 366–367
perfect 1, 268, 280–281, 332, 401,
404, 406–407, 411, 414
postposition 79–80
potential 7
pragmatics 20
preposition 80, 86, 286
preprefix 251ff.
progressive 331, 343, 356,
401–402, 404, 406, 408, 411
pronoun 117ff.
pronoun marking 50
prosodic domain theory 173

proximal 230–231
purposive 291–293, 298

Q

quotative verb 134, 136

R

recipient 279, 285
reference 239–240
relativization XI, 209ff., 237–238
resultative 279, 291–293, 298

S

same-subject-constraint 18–19,
33–34
selectors XII, 303ff.
serialization 65ff.
space deixis 18, 19
space-antideictic-constraint
21–26, 32, 33
space-deictic-constraint 21–26,
32
strict layer hypothesis 173
subject 282
subject markers 43ff., 313
subject pronouns 314–315

subject-verb agreement 328
subjunctive 204
subordination 236
switch-reference x, 35, 113ff.

T

TAM IX, 1ff., 51, 55, 135, 264,
272–273, 275, 319, 321–322
topic 410–411
transfer verbs 59, 60, 277ff.
transitive active 347–348

V

valence marker 85
valence-changing strategy 79,
338–340, 351ff.
ventive 29, 31
verb conjugation 327–328
verb extensions 353ff.
verb-initial language 73–74
vocative 410

W

weak pronoun 54
word order 160, 265
wordhood 171ff.

In the series *Typological Studies in Language* the following titles have been published thus far or are scheduled for publication:

- 66 HICKMANN, Maya and Stéphane ROBERT (eds.): *Space in Languages. Linguistic Systems and Cognitive Categories*. xii, 317 pp. + index. *Expected September 2006*
- 65 TSUNODA, Tasaku and Taro KAGEYAMA (eds.): *Voice and Grammatical Relations. In Honor of Masayoshi Shibatani*. 2006. xviii, 342 pp.
- 64 VOELTZ, F. K. Erhard (ed.): *Studies in African Linguistic Typology*. 2006. xiii, 426 pp.
- 63 FILIMONOVA, Elena (ed.): *Clusivity. Typology and case studies of the inclusive–exclusive distinction*. 2005. xii, 436 pp.
- 62 COUPER-KUHLEN, Elizabeth and Cecilia E. FORD (eds.): *Sound Patterns in Interaction. Cross-linguistic studies from conversation*. 2004. viii, 406 pp.
- 61 BHASKARARAO, Peri and Karumuri Venkata SUBBARAO (eds.): *Non-nominative Subjects. Volume 2*. 2004. xii, 319 pp.
- 60 BHASKARARAO, Peri and Karumuri Venkata SUBBARAO (eds.): *Non-nominative Subjects. Volume 1*. 2004. xii, 325 pp.
- 59 FISCHER, Olga, Muriel NORDE and Harry PERRIDON (eds.): *Up and down the Cline – The Nature of Grammaticalization*. 2004. viii, 406 pp.
- 58 HASPELMATH, Martin (ed.): *Coordinating Constructions*. 2004. xcv, 578 pp.
- 57 MATTISSEN, Johanna: *Dependent-Head Synthesis in Nivkh. A contribution to a typology of polysynthesis*. 2003. x, 350 pp.
- 56 SHAY, Erin and Uwe SEIBERT (eds.): *Motion, Direction and Location in Languages. In honor of Zygmunt Frajzyngier*. 2003. xvi, 305 pp.
- 55 FRAJZYNGIER, Zygmunt and Erin SHAY: *Explaining Language Structure through Systems Interaction*. 2003. xviii, 309 pp.
- 54 AIKHENVALD, Alexandra Y. and R.M.W. DIXON (eds.): *Studies in Evidentiality*. 2003. xiv, 349 pp.
- 53 GIVÓN, T. and Bertram F. MALLE (eds.): *The Evolution of Language out of Pre-language*. 2002. x, 394 pp.
- 52 GÜLDEMANN, Tom and Manfred von RONCADOR (eds.): *Reported Discourse. A meeting ground for different linguistic domains*. 2002. xii, 425 pp.
- 51 NEWMAN, John (ed.): *The Linguistics of Sitting, Standing and Lying*. 2002. xii, 409 pp.
- 50 FEIGENBAUM, Susanne and Dennis KURZON (eds.): *Prepositions in their Syntactic, Semantic and Pragmatic Context*. 2002. vi, 304 pp.
- 49 WISCHER, Ilse and Gabriele DIEWALD (eds.): *New Reflections on Grammaticalization*. 2002. xiv, 437 pp.
- 48 SHIBATANI, Masayoshi (ed.): *The Grammar of Causation and Interpersonal Manipulation*. 2002. xviii, 551 pp.
- 47 BARON, Irène, Michael HERSLUND and Finn SØRENSEN (eds.): *Dimensions of Possession*. 2001. vi, 337 pp.
- 46 AIKHENVALD, Alexandra Y., R.M.W. DIXON and Masayuki ONISHI (eds.): *Non-canonical Marking of Subjects and Objects*. 2001. xii, 364 pp.
- 45 BYBEE, Joan and Paul J. HOPPER (eds.): *Frequency and the Emergence of Linguistic Structure*. 2001. vii, 492 pp.
- 44 VOELTZ, F. K. Erhard and Christa KILIAN-HATZ (eds.): *Ideophones*. 2001. x, 436 pp.
- 43 GILDEA, Spike (ed.): *Reconstructing Grammar. Comparative Linguistics and Grammaticalization*. 2000. xiv, 269 pp.
- 42 DIESSEL, Holger: *Demonstratives. Form, function and grammaticalization*. 1999. xii, 205 pp.
- 41 FRAJZYNGIER, Zygmunt and Traci S. CURL (eds.): *Reciprocals. Forms and functions. Volume 2*. 2000. xii, 201 pp.
- 40 FRAJZYNGIER, Zygmunt and Traci S. CURL (eds.): *Reflexives. Forms and functions. Volume 1*. 2000. xiv, 286 pp.
- 39 PAYNE, Doris L. and Immanuel BARSHI (eds.): *External Possession*. 1999. ix, 573 pp.
- 38 SIEWIERSKA, Anna and Jae Jung SONG (eds.): *Case, Typology and Grammar. In honor of Barry J. Blake*. 1998. 395 pp.
- 37 GIACALONE-RAMAT, Anna and Paul J. HOPPER (eds.): *The Limits of Grammaticalization*. 1998. vi, 307 pp.
- 36 NEWMAN, John (ed.): *The Linguistics of Giving*. 1998. xv, 373 pp.
- 35 GIVÓN, T. (ed.): *Grammatical Relations. A functionalist perspective*. 1997. viii, 350 pp.
- 34 GIVÓN, T. (ed.): *Conversation. Cognitive, communicative and social perspectives*. 1997. viii, 302 pp.
- 33 FOX, Barbara A. (ed.): *Studies in Anaphora*. 1996. xii, 518 pp.
- 32 BYBEE, Joan and Suzanne FLEISCHMAN (eds.): *Modality in Grammar and Discourse*. 1995. viii, 575 pp.

- 31 GERNSBACHER, Morton Ann and T. GIVÓN (eds.): Coherence in Spontaneous Text. 1995. x, 267 pp.
- 30 DOWNING, Pamela A. and Michael NOONAN (eds.): Word Order in Discourse. 1995. x, 595 pp.
- 29 KAHREL, Peter and René van den BERG (eds.): Typological Studies in Negation. 1994. x, 385 pp.
- 28 GIVÓN, T. (ed.): Voice and Inversion. 1994. viii, 402 pp.
- 27 FOX, Barbara A. and Paul J. HOPPER (eds.): Voice: Form and Function. 1994. xiii, 377 pp.
- 26 LORD, Carol: Historical Change in Serial Verb Constructions. 1993. x, 273 pp.
- 25 SVOROU, Soteria: The Grammar of Space. 1994. xiv, 290 pp.
- 24 PERKINS, Revere D.: Deixis, Grammar, and Culture. 1992. x, 245 pp.
- 23 KEMMER, Suzanne: The Middle Voice. 1993. xii, 300 pp.
- 22 PAYNE, Doris L. (ed.): Pragmatics of Word Order Flexibility. 1992. viii, 320 pp.
- 21 DOWNING, Pamela A., Susan D. LIMA and Michael NOONAN (eds.): The Linguistics of Literacy. 1992. xx, 334 pp.
- 20 CROFT, William, Suzanne KEMMER and Keith DENNING (eds.): Studies in Typology and Diachrony. Papers presented to Joseph H. Greenberg on his 75th birthday. 1990. xxiv, 243 pp.
- 19:2 TRAUGOTT, Elizabeth Closs and Bernd HEINE (eds.): Approaches to Grammaticalization. Volume 2: Types of grammatical markers. 1991. xii, 558 pp.
- 19:1 TRAUGOTT, Elizabeth Closs and Bernd HEINE (eds.): Approaches to Grammaticalization. Volume 1: Theoretical and methodological issues. 1991. xii, 360 pp.
- 18 HAIMAN, John and Sandra A. THOMPSON (eds.): Clause Combining in Grammar and Discourse. 1988. xiii, 428 pp.
- 17 HAMMOND, Michael, Edith A. MORAVCSIK and Jessica WIRTH (eds.): Studies in Syntactic Typology. 1988. xiv, 380 pp.
- 16 SHIBATANI, Masayoshi (ed.): Passive and Voice. 1988. x, 706 pp.
- 15 AUSTIN, Peter (ed.): Complex Sentence Constructions in Australian Languages. 1988. vii, 289 pp.
- 14 HINDS, John, Shoichi IWASAKI and Senko K. MAYNARD (eds.): Perspectives on Topicalization. The case of Japanese WA. 1987. xi, 307 pp.
- 13 Never published.
- 12 NEDJALKOV, Vladimir P. (ed.): Typology of Resultative Constructions. Translated from the original Russian edition (1983). Translation edited by Bernard Comrie. 1988. xx, 573 pp.
- 11 TOMLIN, Russell S.: Coherence and Grounding in Discourse. Outcome of a Symposium, Eugene, Oregon, June 1984. 1987. viii, 512 pp.
- 10 RANSOM, Evelyn N.: Complementation: its Meaning and Forms. 1986. xii, 226 pp.
- 9 BYBEE, Joan: Morphology. A Study of the Relation between Meaning and Form. 1985. xii, 235 pp.
- 8 SLOBIN, Dan I. and Karl ZIMMER (eds.): Studies in Turkish Linguistics. 1986. vi, 294 pp.
- 7 CRAIG, Colette G. (ed.): Noun Classes and Categorization. Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983. 1986. vii, 481 pp.
- 6 HAIMAN, John (ed.): Iconicity in Syntax. Proceedings of a symposium on iconicity in syntax, Stanford, June 24-26, 1983. 1985. vi, 402 pp.
- 5 RUTHERFORD, William E. (ed.): Language Universals and Second Language Acquisition. 1984. ix, 264 pp.
- 4 CHISHOLM, William, Louis T. MILIC and John A.C. GREPPIN (eds.): Interrogativity. A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages, Cleveland, Ohio, October 5th 1981-May 3rd 1982. 1984. v, 302 pp.
- 3 GIVÓN, T.: Topic Continuity in Discourse. A quantitative cross-language study. 1983. vi, 492 pp.
- 2 HAIMAN, John and Pamela MUNRO (eds.): Switch Reference and Universal Grammar. Proceedings of a symposium on switch reference and universal grammar, Winnipeg, May 1981. 1983. xv, 337 pp.
- 1 HOPPER, Paul J. (ed.): Tense-Aspect. Between semantics & pragmatics. 1982. x, 350 pp.