

Nilo-Saharan – Models and Descriptions

NILO – SAHARAN
Linguistic Analyses and Documentation

Edited by
*Norbert Cyffer, Osamu Hieda, Angelika Mietzner,
Doris Payne and Anne Storch*

ISSN 0932 - 1993

Volume

28

RÜDIGER KÖPPE VERLAG · KÖLN

Nilo-Saharan – Models and Descriptions

Edited by

Angelika Mietzner & Anne Storch



RÜDIGER KÖPPE VERLAG · KÖLN

Bibliographic information published by the Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

ISBN 978-3-89645-665-6

© 2015 The Editors

RÜDIGER KÖPPE VERLAG

P.O. Box 45 06 43

50881 Cologne

Germany

www.koeppe.de

All rights reserved.

Production: Klever GmbH, Bergisch Gladbach / Germany

♾ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

CONTENTS

Preface	7
Number and case inflection in Tennessean noun and adjective <i>Adelino Amargira</i>	9
Case alignment(s) in Sinyar <i>Pascal Boyeldieu</i>	21
Case marking in Tugen <i>Prisca Jerono</i>	37
Case as meta-categorical heuristic in Ik grammaticography <i>Terrill Schrock</i>	49
Classifier, reflexive, and beyond: a synchronic and diachronic exploration of the morpheme /tʃa/ ‘body’ in Gumuz <i>Colleen Ahland</i>	61
Locational nouns in Datooga <i>Stefan Bruckhaus</i>	73
The use of body part lexemes in Berta <i>Susanne Neudorf</i>	93
The Ama dual suffix: an internal reconstruction <i>Russell Norton</i>	113
Dual number in Ethiopian Komo <i>Manuel A. Otero</i>	123
The pronominal system of Baka <i>Christine Waag & Martin Phodunze</i>	135
Stamp morphs in Central Sudanic languages <i>Gregory D.S. Anderson</i>	151
Syntactic sensitivity and preferred clause structure in Majang <i>Andreas Joswig</i>	169

Complex verbal predicates in Uduk <i>Don Killian</i>	177
The philosophy of walking: motion and the verbs of walking in Cherang'any (Kalenjin) <i>Angelika Mietzner</i>	199
Perspectives on Nilotic verb composition: why can't we agree on the Maa verb? <i>Doris L. Payne</i>	211
Mixed pivot constraints in Toposa clause chaining <i>Helga Schröder</i>	231
Tone and tongue root as ventive morphemes in Endo-Marakwet <i>Chelimo Andrew Kiprop</i>	245
The relationship between syllable weight and stress in Dholuo <i>Jane Akinyi Ngala Oduor</i>	255
Eastern Nilotic vowel harmony and optimality theory: what can the optimal system look like? <i>Diane Lesley-Neuman</i>	273
Berta and the East Jebel subfamily: reinvestigating a Nilo-Saharan isolate <i>Nate D. Bremer</i>	311
The diachronic developments of KI Constructions in the Luo and Koman families <i>Richard Griscom</i>	357
Linguistic etiquette in a frontier situation: a case study from Chopi <i>Anne Storch</i>	373
Is Aroid Nilo-Saharan or Afro-Asiatic? Some evidences from phonological, lexical and morphological reconstructions <i>Moges Yigezu</i>	381

PREFACE

Nilo-Saharan studies are one of the most vibrant and innovative fields in African linguistics. Besides representing a large number of language groupings, individual linguistic codes, and ways of speaking, Nilo-Saharan as a phylum extends over a huge geographical area, roughly from Nigeria in the West to Ethiopia, Kenya, and Tanzania in the East. It is the close cooperation between linguists of the various universities in these countries, the speakers of those languages, members of NGOs, and scholars from universities outside Africa that make Nilo-Saharan studies an exciting, diversified and highly productive field.

The present volume aims at illustrating the current approaches to modelling language change and the description of phonological, morphological and syntactic structures, as well as sociolinguistic and pragmatic features of these languages. We are pleased that the contributions to these manifold topics come from both established and aspiring young scholars, pointing at the highly innovative potential of the field and its continuing development.

The present volume is the outcome of the 11th Nilo-Saharan Linguistics Colloquium which took place in May 22nd-24th, 2013 at the University of Cologne. The event was generously sponsored by the German Research Foundation (DFG), to whom we remain deeply grateful. We are grateful for the fantastic cooperation and assistance rendered by the University of Cologne, and in particular all our colleagues of the Institute of African Studies and Egyptology.

The entire conference was intellectually stimulating and exciting, full of wonderful discussions, and a firework of ideas. We are grateful to all the participants who took part in the colloquium, and to all our colleagues who provided feedback, input and support.

This volume would never have been published within a reasonable time-frame without the expertise of Marvin Kumetat in copy-editing and formatting. We also remain grateful to Monika Feinen for providing maps and graphics for the volume.

We hope that the materials and analyses presented in this volume will ignite further interest in this fascinating field of linguistics.

Cologne, January 2015
Angelika Mietzner & Anne Storch

NUMBER AND CASE INFLECTION IN TENNET NOUN AND ADJECTIVE, A SURMIC LANGUAGE OF SOUTH SUDAN

Adelino Amargira

1 Introduction

Tennet is one of the Southwest Surmic languages that are found in South Sudan, belonging to the Eastern Sudanic branch of the Nilo-Saharan language family. Other members of the Southwest Surmic family inside South Sudan are Murle, Didinga, Naarim and Baale (Kachepeo spoken in Boma plateau both in South Sudan and Ethiopia). Tennet, Didinga and Laarim are found in Eastern Equatoria State of South Sudan, but Murle is found in Jonglei State. Numerically, Tennet people are the smallest member of the Southwest Surmic family. Its home area consists of six villages on the northern tip of the Lopit Mountains, 65 kilometers northeast of the town of Torit in South Sudan (see map in Appendix A).

The Tennet have adopted some aspects of the neighboring Lopit culture and have intermarried with the Lopit, but Tennet is still their primary language. An estimate of today's population is 10,000, which includes the Tennet living in other parts of South Sudan. Dimmendaal (1983) was the first to publish linguistic data specifically on Tennet. More language data is available in Randal (1995), and Tennet Grammar Book, by Waag (2012). Tennet is a VSO language.

Inflectional morphology is that which is determined by the rules of syntax, or which interacts crucially with syntax in detailed manner. According to Payne (1997: 27), "Inflectional operations are those which are required by the syntactic environment in which a root appears. Inflectional operations do not normally alter the basic meaning of the concept expressed. Rather they ground the concept expressed by a root according to place, time, participant reference, etc."

In this paper, we discuss the inflectional suffixes that attach to nouns and adjectives to indicate number or case. Tennet inflection marks number on nouns and adjectives, and nominative, ablative, and genitive case on nouns. The paper begins with a brief description of phonemes, tone and orthography in sections 2–3. Number marking on nouns is introduced in section 4, followed by case marking on nouns in section 5. Each case is then discussed in turn in sections 6 to 9. Number marking on adjectives is discussed in 10. Verb morphology will be dealt with separately.

2 Tennet vowels

Tennet has two sets of vowels, which are orthographically distinguished. One set of five vowels is produced with the tongue root pushed forward, having Advanced Tongue Root [+ATR]. The other set of five vowels is produced with the tongue root further back, [-ATR]. The [-ATR] set is written without any mark and the [+ATR] set is written with an underline mark. Here is the list of the two sets of the vowels:

Table 1: Vowel chart (adapted from Amargira et al 1999)

- ATR		+ ATR	
Tennet Orthography	Phonetic	Tennet Orthography	Phonetic
i	ɪ	<u>i</u>	ɪ̄
e	ɛ	<u>e</u>	ɛ̄
o	ɔ	<u>o</u>	ɔ̄
u	ʊ	<u>u</u>	ū
a	a	<u>a</u>	ā

Vowel harmony is a type of assimilation which takes place when vowels come to share certain features with contrastive vowels elsewhere in a word or phrase (Crystal 1991: 168). This feature of

vowel harmony is found in Tennet. [+ATR] and [-ATR] vowels are generally not mixed in the same words. [+ATR] quality spreads from roots to suffixes or from suffixes to roots. In the examples below, the [+ATR] quality of the suffixes *-êc*, *-tîi*, and *-it* spreads to the roots.

Vowel harmony

Noun	singular	plural	
	dól- êc	dóólí	‘child’
	unge-c	unge- tîi	‘nose’
Verb	3 rd pl. progressive	3 rd pl. completive	
	écéb	écéb- it	‘tie’
	ádák	ada- it	‘eat’

Tone carries an important function in the language, both lexically and grammatically. There are two level tones. In this paper, high tone is marked by acute accent and low tone unmarked but considered. The falling tone is marked by circumflex on the vowel. Rising tone is indicated by using two vowels with a low high pattern.

3 Tennet consonants

The following chart shows the consonants in the Tennet orthography. Wherever a symbol differs from the IPA, the IPA symbol of the consonant is included as it would sound intervocalically. The intervocalic form is chosen because several of the consonants are subject to phonological changes that apply in other environments (word-initial, word-final, and preceding another consonant).

Table 2: Consonant chart (adapted from Randal, 1995)

	Labial		Dental		Alveolar		Postalveolar/ Retroflex		Velar	
	Lenis	Fortis	Lenis	Fortis	Lenis	Fortis	Lenis	Fortis	Lenis	Fortis
-vc stop	p	pp [p:]			t	tt [t:]	t	[t]	k	kk [k:]
+vc stop	b [b]	bb [b:]	dh [d̪]		d	dd [d:]			g	gg [g:]
fricative	v	vv [v:]	z	zz [ð:]					h	hh [χ]
-vc affricate							c	cc [tʃ]		
+vc affricate							j	jj [dʒ]		
nasal	m	mm [m:]			n	nn [n:]	ny [ɲ]		ng	nng [ŋ:]
flap/trill					r	rr [r:]				
approx	w	ww [w:]			l	ll [l:]	y	yy [j:]		

4 Number marking in Tennet

Number in Tennet is marked by inflectional suffixes attached to the noun stem as in other Surmic languages. Dimmendaal explains the system for the Surmic language Baale, found in Ethiopia:

“Number marking in Baale is expressed through a variety of suffixes attached to an inherently singular noun stem (plural marking), to an inherently plural noun stem (singular marking), or to be the singular and the plural in replacement pattern.” (2000: 220)

In Tennet, as in Baale, singular suffixes attach to the stem in some nouns, plural suffixes attach to the stem in other nouns, and singular and plural suffixes both attach to the stem in still other nouns. See the following examples:

Number marking in Tennet

	Singular noun	Plural noun	Gloss
Singulative marking	lábí- tot	lábí	‘grain’
	arte- nit	arte	‘grass’
Plurative marking	íllá	illa- nya	‘spear’
	tátûg	tatug- et	‘door’
Replacement marking	tuúli- c	tuuli- nya	‘chicken’
	goolé- c	goolé- n	‘calf’

5 Case marking in Tennet

Case marking is an important feature of Tennet grammar. Although the accusative is unmarked, suffixes distinguish between subject (nominative), object (accusative), locative/instrumental (ablative) and possessive (genitive) case. Case suffixes agree in number with the noun to which they attach. In the examples below, singular case suffixes attach to the singular noun *cerem* and plural suffixes to the plural noun *m̄ining*. The plural form *cerem-wa* and singular form *m̄ining-ít* are given in parentheses for comparison.

Case suffixes on singular and plural nouns

	Accusative (ACC)	Nominative (NOM)	Ablative (ABL)	
SG	cerem	cerem- i	cerem- a	‘clothing item’
	(cerem-wa)	(cerem-wa- na)	(cerem-wa- ne)	
PL	m̄ining- ít	m̄ining- a	m̄ining- e	‘evil spirit’
	(m̄ining-ít)	minig-a	minig-ewane	
	m̄ining	(m̄ining-ít- i)	(m̄ining-t- a)	
	m̄ining	m̄ining- a	m̄ining- e	
	Genitive SG (GEN)	Genitive PL (GEN)		
SG	cerem- o	cerem- ak		
	(cerem-wa- nu)	(cerem-wa- nu)		
PL	m̄ining- u	m̄ining- u		
	minig-wanu	minig-ak		
	(m̄ining-t- o)	(m̄ining-t- ak)		
	m̄ining- u	m̄ining- u		

6 Absolutive case

Absolutive case in Tennet is the unmarked grammatical case of a core argument of a verb as the case in nominative, which is used as the citation form of a noun. However, absolutive (ABS) nouns are only marked for number.

- (1) a. Ányáha gaala-nit-i túré
brought soldier-SG-NOM gun.ABS
‘The soldier brought a gun.’
- b. Ányáha gaala-nit-**í** túré-t
brought soldier-SG-NOM gun-PL.ABS
‘The soldier brought guns.’

- (2) a. Idímán ngáá dhongku
made lady basket.ABS
'The lady made a basket.'
- b. Idímán ngáá dhongkw-**et**
made lady basket-PL.ABS
'The lady made baskets.'
- (3) a. Okkóyé logoz-it-i illá Láárim-a
taking person-SG-NOM spear.ABS Laarim-ABL
'The person is taking a spear to a Laarim person.'
- b. Okkóyé logoz-it-i illa-**nyá** Láárim-a
taking person-SG-NOM spear-PL.ABS Laarim-ABL
'The person is taking spears to a Laarim person.'

As previously mentioned, singular and plural formation occurs by adding suffixes. There are several singular suffixes and several plural suffixes. Some nouns attach singular suffixes, some nouns attach plural suffixes, and some nouns attach both singular and plural suffixes. Suffixes attaching to nouns are generally not predictable by either phonology or semantics.

Absolutive nouns with plural suffixes

<i>Suffix</i>	<i>Singular Noun</i>	<i>Plural Noun</i>	<i>Gloss</i>
-wa	l̥h̥oḥ	l̥h̥oḥ-d- wâ	'cow bell'
	meger	meger- wá	'gazelle'
	keelang	keelang- wa	'leopard'
	cerem	cerem- wa	'clothing item'
-nya	illá	illá- nyá	'spear'
	waraga	waraga- nya	'book'
	úwât	uwa- nya	'snake'
-et	tátûg	tatug- et	'door'
	deere	deer- et	'gourd'
	joore	joor- et	'mountain'
-e	udúc	uduc- e	'calabash'
	loloc	loloc- e	'spoon'
	guzul	guzul- e	'hyena'
-i	cééz	ceez- i	'house'
	buwec	buwec- i	'lady'
	aḥat	aḥat- i	'tongue'

Absolutive nouns with singular suffixes

<i>Suffix</i>	<i>Singular Noun</i>	<i>Plural Noun</i>	<i>Gloss</i>
-c	orobíya- c	orobíya	'money'
	angaya- c	angaya	'suger cane'
	boroboro- c	boroboro	'kind of tree'
-coc	buru- coc	buru	'egg'
	imma- coc	imma	'hair'
	ínyâ- coc	ínyá	'louse'
-tot	labi- tot	labi	'grain'
	ommogi- tot	ommog	'kind of plant'
	juwa- tot	juwa	'faeces'

-it	mīning- it tēnnetēn- it laarimēn- it	mīning tennet laarim	‘evil spirit’ ‘Tennet people’ ‘Laarim people’
-nit	arte- nit uluta- nit bolbole- nit	arte ulut bolbolen	‘grass’ ‘worms’ ‘kind of tree’

Absolutive nouns with singular and plural suffixes

Suffix	Singular Noun	Plural Noun	Gloss
-t/-n	āzī- t eezi- t áhá- t	azzi- n eezi- n aháni- n	‘hand’ ‘breast’ ‘food’
-c/-én	goolé- c lobore- c ūrē- c	gool- én lobor- en ūr- en	‘calf’ ‘lizard’ ‘kind of gazelle’
-c/-nya	tuúli- c oróbīya- c manga- c	tuuli- nya oróbīya- nya manga- nya	‘chicken’ ‘money’ ‘mango’

In addition, there are some nouns with irregular plural formation, as shown below. This irregularity may be similar to the irregular inflectional morphology in Dinka discussed by Ladd, Remijsen, and Manyang (2009).

Absolutive nouns with irregular inflection

Singular Noun	Plural Noun	Gloss
tang	riḡong	‘cow’
ziirac	temedek	‘antelope’
eēt	ol	‘person’

7 Nominative case

In Tennet, nominative (NOM) case marks the subject, the doer of the action. Singular and plural nominative nouns are inflected. The most common nominative singular suffix is *-i*, and it is often attached to subjects that end with a consonant as in (4). A tone change commonly marks singular subjects ending in a vowel as in (5).

- (4) a. Ákátí éét-**i** móólet
 spear person-NOM calf.ACC
 ‘A person is spearing a calf.’
- b. Ákátí móólet-**i** éét
 spear calf-NOM person.ACC
 ‘A calf is butting a person.’
- (5) a. Apán dóléc ngaa
 slap child.NOM woman.ACC
 ‘A child slaps a woman.’
- b. Apán ngáá dóléc
 slap woman.NOM child.ACC
 ‘A woman slaps a child.’

In (4), the tone of *éét* ‘person’ is high, both when used as a subject in (a), and when used as an object (b). However in (5), the tone of *dóléc* ‘child’ is LH when used as a subject (a) but LHL when used as an object (b). Also in (5), the tone of *ngáá* ‘woman’ is H when used as a subject (b), but L when used as an object (a).

The suffix *-i* attaches to nominative singular nouns and the suffixes *-a*, *-na* attach to nominative plural nouns. A tone change from the tone of the accusative noun can mark both nominative singular and plural nouns.

Nominative singular noun suffixes

<i>Suffix</i>	<i>Accusative Noun</i>	<i>Nominative Noun</i>	
-i	meger	meger- i	‘gazelle’
	tuúlic	tuulic- i	‘chicken’
	keelang	keelang- i	‘leopard’

Nominative plural noun suffixes

	<i>Accusative Noun</i>	<i>Nominative Noun</i>	
-a	ture-t	ture-t- a	‘gun/s’
	tulele-t	tulele-t- a	‘flutes’
	deere-t	deere-t- a	‘gourds’
-na	máála	maalá- nâ	‘calves’
	tiin	tii- na	‘cows’
	ongólwâ	ongólwâ- na	‘elephants’

Nominative singular noun tone changes

	<i>Accusative</i>	<i>Nominative</i>	
L/H	ngaa	ngáá	‘woman’
L/H	tang	táng	‘cow’
H/L	máa	maa	‘lion’
LHL/LH	dóléc	dóléc	‘child’

Nominative plural noun tone changes

	<i>Accusative</i>	<i>Nominative</i>	
H/L	ól	ol	‘humans’
H/L	dóolí	dooli	‘children’
HL/LH	lógôz	logóz	‘young men’

8 Oblique case

Oblique (OBL) case primarily indicates meaning having to do with locations or instruments. For locations, the ablative suffix indicates the specific location of an object, the direction an agent is heading towards, or from where a state or action existed or took place. The ablative suffix can also indicate the instrument of the action. In the ablative examples of (6–8), accusative examples are given in (6a, 7a) for comparison.

- (6) a. *Ávīl* ngáá *cééz*
 sweep lady.NOM house.ACC
 ‘The lady is sweeping the house.’
- b. *Óbóda* ngáá *cééz-a*
 return lady.NOM house-ABL
 ‘The lady returns from the house.’

- c. *Ááve udúc-i cééz-a*
 stay calabash-NOM house-ABL
 ‘The calabash is in the house.’
- (7) a. *Anyúk gaala gool cí Juba*
 closes government.NOM road.ACC of Juba.POSS
 ‘The government closed the Juba road.’
- b. *Óvvó ol Juba-kté*
 going people.NOM Juba-ABL
 ‘People are going to Juba.’
- (8) a. *úruúk éét-i ngaa keet-a*
 beat person-NOM woman.ACC stick-ABL
 ‘The man beat the woman with a stick.’
- b. *Ókkó éét-i Toric-a zoo-nê*
 going person-NOM Torit-ABL foot-ABL.PL
 ‘The person is going to Torit on foot.’

The suffix *-a* attaches to ablative singular nouns and the suffixes *-e*, *-ne* attach to ablative plural nouns. There are also a few less common ablative suffixes such as the singular suffixes *-kté*, *-toto* and the plural suffix *-i*.

Oblique singular noun suffixes

Suffix	Accusative Noun	Ablative Noun	Gloss
-a	<i>cééz</i> keet	<i>cééz-a</i> keet-a	‘house’ ‘tree’
-kté	<i>Toric</i> Juba Lovi Apoeta	<i>Toric-a</i> Juba- kté Lovi- kte Apoeta- kte	‘Torit (town name)’ ‘Juba (town name)’ ‘Lovi (town name)’ ‘Kapoeta’ (town name)’
-to	<i>mana</i> orog <i>ceez</i>	<i>mana-tó</i> orga- tó caa- tó	‘field’ ‘village’ ‘house’

Oblique plural noun suffixes

Suffix	Accusative Noun	Ablative Noun	Gloss
-e	<i>mining</i> keen	<i>mining-e</i> keen-e	‘evil spirits’ ‘sticks’
-ne	<i>allen</i> <i>ceez-i</i> baaru	<i>allen-e</i> <i>ceez-i-né</i> baaru- né	‘whips’ ‘houses’ ‘robs’
-i	<i>ira</i> caava tuulinya tuwenya	<i>ira-nê</i> caava- i tuulinya- i tuwenya- i	‘milk’ ‘shoes’ ‘chickens’ ‘ravens’

Some place names such as *Juba* (7b) and *Lovi* (9b) attach the ablative suffix *-kte* and some place names such as *Leteji* (10b) and *Toric* (11b) attach the ablative suffix *-a*.

- (9) a. *Ákkú* *ngáá* *Lovi* 'The woman is coming to Lovi.'
 b. *Óbódá* *ngáá* *Lovi-kte* 'The woman returned from Lovi.'
- (10) a. *Úzúcá* *dóléc* *Leteji* 'The child reached Leteji (village).'
Údúng *dóléc* *Letéj-a* 'The child slept in Leteji.'
- (11) a. *Ávìl* *gaal* *Torìc* 'The government is cleaning Torit.'
 b. *Údúng* *ééti* *Torìc-a* 'The person slept in Torit.'

9 Genitive case

The genitive (GEN) case indicates possession. The genitive particle *ci/cik* introduces the possessor noun and agrees in number with the possessed noun it follows. Genitive suffixes attach to possessor nouns. As in other cases, suffixes depend on whether the noun is singular or plural. In genitive case, if the possessor noun is singular, the suffixes also differ according to whether the possessed noun is singular or plural. In (12a), the genitive suffix *-o* marks a singular noun possessing a singular noun ('lady's dog'). In (b), the genitive suffix *-ak* marks a singular noun possessing a plural noun ('lady's dogs'). In (c-d), the genitive suffix *-nu* marks a plural noun possessing a singular or plural noun ('ladies' dog' or 'ladies' dogs').

- (12) a. *Ícín* *elegit-i* *uruz* *cí* *buuwec-o*
 saw animal-NOM dog.ACC of lady-GEN
 'The animal saw the dog of the lady (lady's dog).'
- b. *Ícín* *elegit-i* *uruz-i* *cik* *buuwec-ak*
 saw animal-NOM dog-PL.ACC of lady-GEN
 'The animal saw the dogs of the lady (lady's dogs).'
- c. *Ícín* *elegit-i* *uruz* *cí* *buuwec-i-nu*
 saw animal-NOM dog.ACC of lady-PL-GEN
 'The animal saw the dog of the ladies (ladies' dog).'
- d. *Ícín* *elegit-i* *uruz-i* *cik* *buuwec-i-nu*
 saw animal-NOM dog-PL.ACC of lady-PL-GEN
 'The animal saw the dogs of the ladies (ladies' dogs).'

The suffixes *-o*, *-wo* attach to singular nouns possessing singular nouns, the suffixes *-ak*, *-wak* attach to singular nouns possessing plural nouns, and the suffixes *-u*, *-nu* attach to plural nouns possessing singular or plural nouns.

Genitive suffixes on singular nouns possessing singular nouns

Suffix	Accusative Noun	Genitive Noun	Gloss
-o	<i>Tamutáddên</i> <i>dóléc</i> <i>ééz</i>	<i>Tamutáddên-o</i> <i>dóléc-ô</i> <i>ééz-o</i>	'God' 'child' 'goat'
-wo	<i>ngáá</i> <i>ínyá</i> <i>ibâ</i>	<i>ngáá-wo</i> <i>ínyâ-wo</i> <i>ibâ-wo</i>	'wife' 'neck' 'arm'

Genitive suffixes on singular nouns possessing plural nouns

Suffix	Accusative Noun	Genitive Noun	Gloss
-ak	buuwec	buuwec- ak	‘lady’
	tátûg	tátûg- ak	‘door’
	lopilding	lopilding- ak	‘window’
-wak	ngáá	ngáá- wak	‘wife’
	tamu	tamu- wak	‘rain’
	bbambbu	bbambbu- wak	‘tobacco’

Genitive suffixes on plural nouns possessing singular or plural nouns

Suffix	Accusative Noun	Genitive Noun	Gloss
-u	ééza	eeza- u	‘goats’
	ázza	ázza- u	‘sheep’
	tiin	tiin- u	‘cows’
-nu	ngáá	ngaa- nu	‘wives’
	buuwec	buuwec-i- nu	‘ladies’
	meger	meger-wa- nu	‘gazelle’

10 Singular and plural adjectives

As in other Surmic languages, Tennet attaches suffixes to adjectives to express plural number agreement with the nouns they describe. The genitive particle *ci/cik*, which introduces the adjectives, also agrees in number with the nouns described.

- (13) a. Ányáhá eet-i tang ci olle
 brought person-NOM cow.ACC of black
 ‘The man brought a black cow.’
- b. Ányáhá eet-i tiin cik olle-k
 brought person-NOM cows.ACC of black-PL
 ‘The man brought black cows.’

Plural adjective suffixes are *-k*, *-ttik*, or *-wa*. As with noun suffixes, adjective suffixes are not predictable by phonology or semantics. Unlike nouns, only plural adjectives have suffixes.

Singular and plural adjectives inflection suffix

Singular Adjective	Plural Adjective	Gloss
culai	culai- k	‘green’
olle	olle- k	‘black’
voor	voori- k	‘white’
wuun̩i	wuun- ttik	‘long’
appwa	appin- ttik	‘huge’
bunybunya	bunybuny- ttik	‘giant’
guur	guur- wá	‘lazy’
ariz	ariz- wá	‘huge bull’
ilel	ilel- wá	‘barren’

11 Conclusion

Inflection plays an important role in Tennes nouns and adjectives, marking number and case. Singular and plural suffixes attach to nouns, but only plural suffixes attach to adjectives. Some nouns attach singular suffixes (*lábi-tot / lábí* ‘grain’), some nouns attach plural suffixes (*illá / illa-nya* ‘spear’), and some nouns attach both singular and plural suffixes (*tuúli-c / tuuli-nya* ‘chicken’).

Absolutive case is unmarked by any suffix, but other cases are marked by suffixes agreeing in number with the noun. The nominative singular suffix is *-i* (*meger-i* ‘gazelle’) and common nominative plural suffixes are *-a*, *-na* (*ture-t-a* ‘cows’), although a tone change from the accusative form can also indicate nominative singular or plural case (*ngaa / ngáá* ‘wife’, *lógóz / logóz* ‘young men’). The oblique singular suffix is *-a* (*ceez-a* ‘in house’) and oblique plural suffix is *-e* (*mining-e* ‘evil spirits’), although for place names the ablative suffix can be *-kte* (*Lovi-kte* ‘from Lovi’). The genitive suffix on singular nouns possessing a singular noun is *-o*, *-wo* (*ngáá-wo* ‘wife’), on singular nouns possessing plural nouns *-ak*, *-wak* (*ngáá-wak* ‘wife’), and on plural nouns possessing either singular or plural nouns *-u*, *-nu* (*ngaa-i-nu* ‘wives’).

Finally, plural adjectives can be marked with the suffixes *-k*, *-wa*, and *-ttik* (*culai-k* ‘green’, *wuuni-ttik* ‘long’, *guur-wá* ‘lazy’).

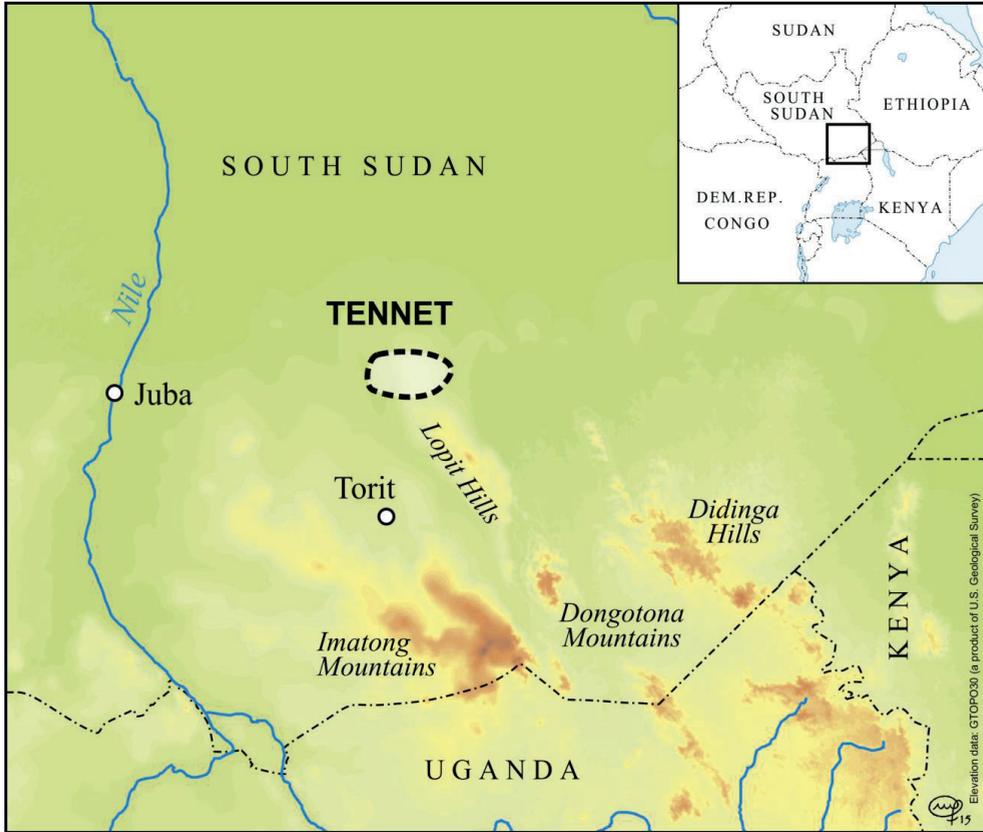
Abbreviations

ABL	ablative
ACC	accusative
GEN	genitive
NOM	nominative
pl., PL	plural

References

- Amargira, Adelino P., Rev. Severino Maira, Rev. Karlo Kolong, Luka Aribok & Richard P. Margetts. 1999. *Writing Tennes*. Kitale, Kenya.
- Crystal, David. 1991. *A Dictionary of Linguistics and Phonetics*. Cambridge, MA: Basil Blackwell.
- Dimmendaal, Gerrit J. 1983. Tenet (Surma): Ein Fall partiellen Sprachwechsels. In Rainer Voßen & Ulrike Claudi (eds.), *Sprache, Geschichte und Kultur in Afrika: Vorträge, gehalten auf dem III. Afrikanistentag*, Köln (Lectures held at the third African Exhibit, Cologne, October 1982), pp. 331—344. Hamburg: Helmut Buske.
- Dimmendaal, Gerrit J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42.2: 214—261.
- Ladd, D. Robert, Caguor Adong Manyang & Bert Remijnsen. 2009. Short report on the distinction regular and irregular inflectional morphology evidence from Dinka. *Language* 85.3: 659—670.
- Payne, Thomas. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press.
- Randal, Scott. 1995. *Nominal Morphology in Tennes*. MA thesis, Dallas: University of Texas at Arlington.
- Waag, Christine, Amargira Adelino, Lomojo Cirino Osfaldo & Loromo Peter Bartholomeo. 2012. *Tennes Grammar Book*. Juba: SIL-South Sudan.

Map 1: Location of the Tennet area¹



¹ This map is adapted from Randal (1995).

CASE ALIGNMENT(S) IN SINYAR

Pascal Boyeldieu

1 Introduction

The Sinyar language (*tàar fàmjà*) is spoken in an area that is shared between Chad (main settlement: Mongororo, Waddai province) and Sudan (main settlement: Foro Boranga, South to Darfur). For Chad only the number of speakers was estimated 12,300 in the year 2000 (Lewis, Simons & Fennig 2013). A roughly similar amount may be suggested for Sudan. Their closest neighbours are Daju of Sila (Chad) and Fur (Sudan).

Up to now Sinyar has been classified within the (Sara-)Bongo-Bagirmi (henceforth SBB) sub-branch of Central Sudanic (Tucker & Bryan 1956: 14; Haaland 1978; Doornbos undated; Doornbos & Bender 1983, Bender 1992). However this position relied on very limited data and recent fieldwork of my own revealed that things are not so straightforward. Sinyar does indeed share many lexical and morphosyntactic features with SBB languages but, as shown in table 1 below, it also differs from them in other features of the same domains, the origin or relationship of which remains undetermined, so that the social and historical genesis of Sinyar still represents an open question in the present state of research.

Table 1: Sinyar features converging with/diverging from SBB features

SBB-like	SBB-unlike
<ul style="list-style-type: none">• part of the lexicon (including ‘basic’)• part of the personal pronouns• personal indices 2 & 3 distinguished for number by plural suffixes on the verb• infinitive in <i>t-</i> (vowel-initial verbs)• relative verb forms in <i>k(V)-</i>• adjectives in <i>k-</i> (vowel-initial verbs)	<ul style="list-style-type: none">• part of the lexicon (including ‘basic’)• no tonal alternation on verbs• some double (related) radicals for verbs• intensive verbs in <i>-r-</i>• factitive verbs suffixed with <i>-oo/-uu</i>• noun plural in (<i>-jà</i>)• some noun plurals in <i>-àar</i> (animates)• <i>case marker system</i>

The aim of this paper is to give an account of Sinyar’s (double) case marker system, probably the most original and striking feature in comparison with the SBB language group. After displaying the markers and their complementarity (section 3), and commenting on morphosyntactic principles (section 4), I will review the way cases are involved in the expression of the different functions¹ that may be assumed by noun phrases (section 5), summarize the relations between function and case (section 6), and finally give concluding remarks (section 7).

2 Preliminary remarks

Sinyar contrasts two level tones (Low/High) and has frequent downstepped Lows (the tonal transcriptions should not be considered as definitive).

Kind of ‘floating’ tones – symbolised by (L) / (H) put after the term – condition the tonal realisation of subsequent toneless suffixes (or clitics), a property which is, among others, the case of certain case markers. Contrast for instance the behaviour of the Nominative Singular *-Ni* after *bís* (L) ‘cat’ (1) and *lél* (H) ‘donkey’ (2):

¹ In doing so I was appreciably inspired by an article of Kießling (2007) on the marked nominative in Datooga.

- (1) *bís-ni*
cat-NOM:SG
'cat' (Nominative)
- (2) *lél-li*
donkey-NOM:SG
'donkey' (Nominative)

The nominal Plural *-ɲà* however is transparent in the sense that the 'floating' tone 'jumps' it and shifts its effect onto the following unit (*-ɲà* may be dropped, especially when it is followed by a case marker whose plural number is explicit). Compare the behaviour of the Nominative Plural *-si* after the two nouns *ingà* (L) 'mountain' (3) and *màasi* (H) 'spear' (4).

- (3) *ingà(-ɲà)-si*
mountain(-PL)-NOM:PL
'mountains' (Nominative)
- (4) *màasi(-ɲà)-sí*
spear(-PL)-NOM:PL
'spears' (Nominative)

Basic syntactic order is SVO with great freedom as to the main constituents (pragmatic effects are likely but not well known).

Numerous examples have been elicited.²

3 Case markers

Sinyar is characterized by two sets of case markers according to the nature of the head of the noun phrase.

i.) The first set concerns phrases whose head is a common noun, except if modified by a deictic (see below). This set, that includes the dummy noun *ɲàkwà* (H) 'thingy, thingummy, what-do-you-call-it' and the partial interrogative *ingá* (H) 'what?', contrasts an unmarked form ('Absolute'), a Nominative, and an Adverbial (a label that covers all values of 'locative', 'instrumental', and 'comitative').

The specific markers of set i. are illustrated in table 2 below:

Table 2: Case markers for phrases whose head is a common noun

	Absolute form	Nominative	Adverbial
Sg.	∅	-n / -Ni ³	-ti
Pl.		-si	

ii.) The second set concerns phrases whose head is a proper name, a deictic or a personal pronoun, as well as cases where a common noun is modified by a deictic modifier. 'Proper names' include personal names – and names of personified animals –, 'God', 'chief', 'grandmother', names of places, rivers, mountains, and countries, the dummy noun *ɲàkwà* 'what's-his/her-name, So-and-So' and the partial interrogative *dè¹e-* (L) 'who?', some animal names: '(home) cat', '(home) duck', 'fox', 'owl

² Data were collected on the occasion of two recent fieldworks in Khartoum (2010, 2011) with the kind assistance of MM. Bakheet Abdulkarim and Zakaria Abdulkarim. The research was supported by the Llacan (UMR 8135), a joint research unit of the CNRS and Inalco (France). I am especially grateful to the Department of Linguistics of the University of Khartoum for its warm welcome and help.

³ These two complementary variants are respectively suffixed to nouns ending in a short vowel vs. nouns ending in a long vowel or a consonant.

sp.', 'bird sp.', 'stork', etc. as well as the points of the compass borrowed from Arabic, e.g. *Rîi* 'North'.

This second set contrasts an unmarked, 'Absolute' form (restricted to proper names *stricto sensu*), a Nominative, a Genitive, an Accusative, and an Adverbial.

The specific markers and specific forms of set ii. are illustrated in tables 3. to 5. below:

Table 3: Case markers for phrases whose head is a proper name (and assimilated)

	Absolute form	Nominative	Genitive	Accusative	Adverbial
Sg.	∅	-n / -lè ⁴	-na (L)	-(y)àa (H)	-lèe (H)
Pl.		-ngè (H) ⁵		-ngàá (H)	-ngèer (H)

Table 4: Deictic modifiers (and pronouns)⁶

		Nominative	Genitive	Accusative	Adverbial
'near'	Sg.	-àní	kin.nàà (L)	kin	(?)íngèr
	Pl.	-àgí	kígí.ningè (H)	kígí	kígí.tí
'far'	Sg.	-àná	wín.nàà (L)	wín	wísèr
	Pl.	-àgá	wísí.ningè (H)	wísí	wísí.tí

Table 5: Personal pronouns

	Nominative	Genitive (Associative)	Accusative		Adverbial
			Independent	Indices	
1S	màalé, màalè, (màá)	-ándà (L)	màá	-má/-mà	màárèe
2S	illé, illè	-nún (L)	làá	-lá/-là	làárèe
3S	nàalé, (nàá)	-náà (L)	niyáá	-nà	niyáárèe
1Pd	cíngé	-náakú (H)	cíngáa	-singà/-singá	cíngáangèer
1P[?]	ciyé		ciyáa		ciyáangèer
1Pp	cèesá	-náasà (H)	cèsáa	-sàasà/-sáasà	cèesáangèer
2P	síngé	-núkú (H)	síngáa	-singá/-singá	síngáangèer
2P[?]	sèsésá		sèsésáa		sèsésáangèer
3P	ningé	-níngè (H)	ningáa	-ningá	ningáangèer

As can be seen from tables 1 and 2, case markers are usually but not always distinct for number. Deictics (table 4) and personal pronouns (table 5) have their own, usually irreducible forms. Note however the genitive forms of the Deictics that combine the accusative forms plus the genitive of the Personal Pronouns 3S and 3P. Several other formal links may be observed between the three tables 3 to 5.

4 Morphosyntactic principles

As a general rule case markers appear at the end of the phrase and may be considered as postpositions. Thus the Nominative is suffixed to the sole head noun representing the subject in (5) but to the last modifier (adjective) following the head noun in (6) (as a convention, italics will highlight illustrated *head* and *case marker*, if any):

⁴ These two complementary variants are respectively suffixed to nouns ending in a short vowel vs. nouns ending in a long vowel or a consonant.

⁵ Note that the plural markers of proper names do not refer to several individuals bearing the same name (see comment on (29) below) but to the group of people associated to a single individual ('associative plural'), e.g. *Gindà-ngè* 'Ginaa (man's name) and his mates/his friends/his family/his clan/etc.' (Nominative and Genitive) or, in Arabic, *nas Ginaa*.

⁶ Deictic pronouns, the Genitive forms of which have not been systematically registered, are very similar to the deictic modifiers.

- (5) *kàjjà-n àmbíy¹à*
 sun-NOM:SG 3:go:PST2
 ‘The time is over’ (lit. the sun has gone)
- (6) *kàndìnè-ṅà-náà kiccí-sì àm¹àn irri*
 cloth-PL-3S:GEN old-NOM:PL 3:have:PRS:PL dirt
 ‘His old clothes are dirty.’

Since common nouns do not know any accusative markers, objects represented by common nouns appear in the unmarked (absolute) form (7a) unless they are modified by a deictic that must appear in the accusative (7b) in the same way as the deictic pronoun (7c):

- (7) a. *ngùbbò-n tèll¹è kwàrjà*
 old_woman- NOM:SG 3:twist:PST1 polenta
 ‘The old woman prepared polenta.’
- b. *ngùbbò-n tèll¹è kwàrjà kìn*
 old_woman- NOM:SG 3:twist:PST1 polenta this:ACC:SG
 ‘The old woman prepared this polenta.’
- c. *ngùbbò-n tèll¹è èkìn*
 old_woman- NOM:SG 3:twist:PST1 this_one:ACC:SG
 ‘The old woman prepared this one.’

Finally genitive case markers (8) or specific forms (9-10) of set ii. may appear within a noun phrase as modifiers of a common noun and be followed by the last case marker governed by the head:

- (8) *ḃáà-ṅà Kùnyáà-nà-sì wèṅpinn¹àa firàngà*
 father-PL Kunyaa-GEN:SG-NOM:PL 3:come:PST2:PL yesterday
 ‘Kunyaa’s fathers (i.e. father and/or uncles) have come yesterday.’
- (9) *kàndìnè-ṅà-náà kiccí-sì àm¹àn irri (= (6) above)*
 cloth-PL-3S:GEN old-NOM:PL 3:have:PRS:PL dirt
 ‘His old clothes are dirty.’
- (10) *ḃáà céè wín.ná-n úwè*
 father person that:GEN:SG-NOM:SG 3:FUT:come
 ‘The father of that person will come.’

Such are the main principles along which case markers function. Further situations and contexts will be illustrated throughout section 5.

5 Function and case

This section will review and comment the use of case along the main functions noun phrases may assume in the Sinyar utterance.

a. Subject of any predicate and focussed subject of verb predicate

The subject is normally marked as nominative. Examples (11-14) illustrate subjects of a verbal predicate (with focussed subject in (12)). (15) is an adverbial utterance with an adverbial noun phrase as predicate while (16) is an equative utterance resulting from the mere juxtaposition of both subject and predicate represented by noun phrases in the nominative (see b. below).

- (11) *ngàar-ri wéppí*
stew-NOM:SG 3:be_nice:PRS
'The stew is nice.'
- (12) *ingá-n k-ùj-1^à*
what?:NOM:SG FOC-hurt:PRS-2S:ACC
'What hurts you?' (lit. what is it that hurts you?)
- (13) *Sùdàan-nè ùnjàbí Íngiltàràa*
Sudan-NOM:SG 3:be_hot:pass:PRS England:ACC:SG
'Sudan is hotter than England.'
- (14) *gèe-ní ùmbò-tá tádá*
child-this:NOM:SG 3:hear:PRS-NEG speech
'This child is disobedient.' (lit. doesn't hear speech)
- (15) *nqùbbòl-lí ùmmù kàanà-tí*
snake-NOM:SG inside grass-ADV
'There is a snake in the grass'
- (16) *Bàkùit-lè cé-k-íjjá¹-n*
Bakheet-NOM:SG person-REL-3:count-NOM:SG
'Bakheet is the accountant.'

In rare cases subjects have been noted as unmarked, i.e. without any nominative. Although not well understood this behaviour could be restricted to nouns referring to inanimate, maybe non-countable notions whose unmarked form would then express indefiniteness. Contrast (17) and (18):

- (17) *kùcùl úngw^{1à}*
money 3:not_be:PRS
'There is no money (money doesn't exist, e.g. 'I/we have no money.').'
- (18) *kùcùl-lí úngw^{1à}*
money-NOM:SG 3:not_be:PRS
'The money is not here (the money is absent, e.g. 'The money is somewhere else.').'

b. Nominal predicate

As already mentioned in connection with example (16) above, the nominal predicate of a simply juxtaposed equative utterance is itself marked as nominative. Examples (19) and (20) illustrate phrases whose head is a common noun or a proper name respectively. In (21) the nominal predicate is modified by a relative clause.

- (19) *Bàkiit-lè cé-k-íjjá¹-n* (= (16) above)
Bakheet:NOM:SG person-REL-3:count-NOM:SG
'Bakheet is the accountant.'
- (20) *èerà-ándà-n Bàkiit-lè*
name-1S:GEN-NOM:SG Bakheet-NOM:SG
'My name is Bakheet.'
- (21) *ángá-sí kàndinè-ṅà k-ingillé-sí*
which_one?:NOM:PL cloth-PL REL-3:get_lost:PST1-NOM:PL
'Which clothes were lost?' (lit. which ones (are) the clothes that were lost?)

- (28) *nàalé áamá¹à nù dē¹e-nà*
 3S:NOM 3:marry:PST2 daughter who?-GEN:SG
 ‘Whose daughter has he married (he has married the daughter of whom)?’
- (29) *Bàkùt mùtjà c-àmà béerè*
 Bakheet three 1P-have:PRS at_home
 ‘We have three Bakheet at home.’
- (30) *Nàmùrò-n ámbóngá¹à Mindéelè-ngàá*
 Spider-NOM:SG 3:visit:PST2 Rat-ACC:PL
 ‘Spider has paid a visit to Rat’s family.’
- (31) *màalé m-igàakà cèè-kín*
 1S:NOM 1S-ignore:PRS person-this:ACC:SG
 ‘I don’t know this person.’

f. Object of ditransitive or applicative verbs

Objects of ditransitive verbs both behave in the same way like patients: common nouns are unmarked (32) while proper names and assimilated are marked with accusative (33). Note that the order of objects is free.

- (32) a. *nàalé ndóolú¹ù gí tàar jàmjà*
 3S:NOM 3:teach:PRS child language Sinyar
 ‘S/he teaches the child the Sinyar language.’
- b. *nàalé ndóolú¹ù tàar jàmjà gí*
 3S:NOM 3:teach:PRS language Sinyar child
 ‘S/he teaches the child the Sinyar language.’
- (33) a. *nàalé injóoy¹à kùcùl Bàkùt-yàa*
 3S:NOM 3:let_count:PST2 money Bakheet-ACC:SG
 ‘S/he let Bakheet count the money.’
- b. *nàalé injóoy¹à Bàkùt-yàa kùcùl*
 3S:NOM 3:let_count:PST2 Bakheet-ACC:SG money
 ‘S/he let Bakheet count the money.’

Several verbs may receive an applicative extension that increases the number of their arguments. In such cases the ‘applicative’ object may refer to different values (namely benefactive, locative, or instrumental) and is again unmarked in the case of common nouns (34) or marked in the accusative in the case of proper names (35-36).

- (34) *ingá áarinnáa wèrfi dúkùm*
 what? 3:FUTdo:APPL:PL knife blunt
 ‘What will they do with the blunt knife?’
- (35) *màalé m-èbì kâyyáa Fòrbárángàá*
 1S:NOM 1S-PRGR be_near:PRS:APPL F_B:ACC:SG
 ‘I am living close to Foro Boranga.’

- (36) k-ikàlmbiyáa-tá *Zàkàriyà-ngè* ngàar
 2:PRHB-cook:APPL-NEG Zakariya-ACC:PL polenta
 ‘Don’t cook polenta for Zakariya’s people!’

g. Adverbial (locative, instrumental, and comitative)

Adverbial adjuncts are usually marked as such, the same markers expressing all the values of locative, instrumental and comitative. Common nouns are illustrated in (37-42) and proper names or assimilated in (43-45). (40-42) exemplify cases of complex adpositions resulting from modified nouns (*kùbàl*...ADV ‘near, close to’, *tà*... ADV ‘because of, according to’).

- (37) wèɛp¹à *kàndinè-ɲà* kwii-tí
 3S:come:PST2 cloth-PL other-ADV
 ‘S/he has come with other clothes.’
- (38) nàalé *sùdò-tí* górnú¹
 3S:NOM sleep-ADV 3:snore:PRS
 ‘S/he snores while sleeping.’
- (39) *mùttà-níngè-tí* wèɛɲinnà¹a
 three-3P:GEN:ADV 3:come:PST2:PL
 ‘They have come in a group of three.’ (lit. with the three of them)
- (40) cijji *kùbàl-ándá-tí*
 2:put:IMP nearness-1S:GEN-ADV
 ‘Put (it) beside me!’ (lit. in my nearness)
- (41) *tà* *céè-sì* *t-àd¹àn-tí* ...
 matter person:PL-NOM:PL DEP-3:say:PRS:PL-ADV
 ‘according to what people say...’
- (42) màalé m-èbì *kùbàl* Fòrbáráɲá-ná-tí
 1S:NOM 1S-live:PRS nearness F_B-GEN:SG-ADV
 ‘I live near Foro Boranga.’ (lit. in the nearness of F.B.)
- (43) m-ábi *Éli-lèe* sùk-tí
 1S-FUT:go Ali-ADV:SG market[Ar.]-ADV
 ‘I’ll go to the market with Ali.’
- (44) illé *tàb¹à-là* *ciyáangèer*
 2S-NOM 2:leave:IMP-2S:NOM 1P[?]:ADV
 ‘Stay with us!’ (lit. you (sg.) leave yourself with us!)
- (45) wèɛɲngònà¹a *dè¹e-lèe* kwii-tí
 3:come:again:PST2 who?-ADV:SG other-ADV
 ‘With whom else has he come?’

However toponyms remain unmarked when used as absolute locatives (46).

- (46) nàalé windéɲ *Sàalé*
 3S:NOM 3:bathe:PST1 Salih[river]
 ‘S/he bathed in Salih (river).’

h. Complement of preposition kándáaká ‘like’

Noun phrases governed by the preposition *kándáaká* ‘like, as’ behave like objects, i.e. they appear complementarily as unmarked (47) or as accusative (48). This fact might suggest a verbal origin for the preposition.

(47) nàalé èrfi kándáaká *jèk*
 3S:NOM 3:speak:PRS like chief [Ar.]
 ‘He speaks like a/the chief.’

(48) nàalé èrfi kándáaká *ningáa*
 3S:NOM 3:speak:PRS like 3P:ACC
 ‘He speaks like them.’

i. Associative=Genitive (modifying a noun, including complex adpositions)

Modifying nouns, i.e. nouns functioning as the associative (genitive) expansion of a head remain unmarked if they are common nouns (49-51). They are marked as genitive in the case of proper names (52-53) and assimilated, e.g. partial interrogative ‘who?’ (54) or deictics (55-56).

(49) ningé álann¹áa kándinè *nàfá*
 3P:NOM 3:put:PST2:PL cloth each_other
 ‘They have put on each other’s cloth.’

(50) ningé èði àd¹àn tádá *báa*-ndà
 3P:NOM 3:PRGR 3:say:PRS:PL matter father-1S:GEN
 ‘They are discussing the matter of my father.’

(51) káalitin tà *ùmmà*-tì
 3:quarrel:PRS:PL matter woman-ADV
 ‘They quarrel concerning (with the matter of) a/the woman.’

(52) érbí térbí *Bórnò*-nà
 3:run:PRS running fox-GEN:SG
 ‘He runs like a fox.’ (lit. he runs a running of fox, i.e. very fast)

(53) ningé èðin illè *Sábáa*-nà-tì
 3P:NOM 3:live:PRS:PL village[Ar.] East-GEN:SG-ADV
 ‘They live in a village of the East.’

(54) ègí ingè(-nà) *dè¹e-ngè*-sí
 this:NOM:PL house(-PL) who?-GEN:PL-NOM:PL
 ‘Whose (pl.) houses are these?’ (lit. these (are) the houses of whom (pl.)?)

(55) báa cèè *kín.ná*-n úwè (= (9) above)
 father person this:GEN:SG-NOM:SG 3:FUT:come
 ‘The father of this person will come.’

(56) nàalé úwèn báa cèè *wísí.ningè*-tí
 3S:NOM 3:FUT:come:PL father person those:GEN:PL-ADV
 ‘He will come (lit. he, they will come) with the father of those people.’

j. Topicalized subject

Sinyar has the ability to strengthen the topicalization of the subject – or making it a frame – by duplicating it with a subject pronoun that appears after a light pause. The extraposed subject itself appears in the nominative. Contrast (57a) and (58a) (subject) with (57b) and (58b) (strengthened topicalization):

- (57) a. *wèrfi kící-n jùwè-t^là jàmmá^ll*
 knife old-NOM:SG 3:cut:PRS-NEG 3:be_nice
 ‘The old knife does not cut well.’
- b. *wèrfi kící-n nàalé jùwè-t^là jàmmá^ll*
 knife old-NOM:SG 3S:NOM 3:cut:PRS-NEG 3:be_nice
 ‘The old knife, it does not cut well.’
- (58) a. *Kùmùlà-ngè ámbínn^làa nùnnùù*
 Kumula:NOM:PL 3:go:PST2:PL far_away
 ‘Kumula’s people have gone away.’
- b. *Kùmùlà-ngè níngé ámbínn^làa nùnnùù*
 Kumula:NOM:PL 3P:NOM 3:go:PST2:PL far_away
 ‘Kumula’s people, they have gone away.’

Another process – that might be characterized as ‘contrastive topicalization’ – consists of embedding the extraposed noun phrase in the frozen formula *tùbà... cè(n)* ‘as for... (also)’. The extraposed subject is again marked as a nominative (59).

- (59) *tùbà Éli-n cè(n) ámbíy^là firàngà*
 as_for Ali:NOM:SG also 3:go:PST2 yesterday
 ‘As for Ali, he has gone yesterday.’

The term *tùbà* obviously derives from the verb *ùbà* ‘to become’ (*t-ùbà* = dependent form?); *cè* ‘also’ is attested elsewhere but it may appear here with an additional *n* according to conditions that are not understood. All instances of this specific construction (see also k. and m. below) concern human proper names.

These cases should not be assimilated to the behaviour of *ùbà* as a plain verb (see section d. above) since the topicalized noun phrase may be marked at least as a nominative, an accusative or an adverbial according to the function copied by the topic (see sections k. and m. below).

k. Topicalized object (patient)

The object may be topicalized through extraposition to the beginning of the clause. It is unmarked in the case of common nouns (60–61) and marked as accusative in the case of proper names (62) or personal pronouns (63). Note that the inanimate object is not necessarily represented by a pronoun *in situ*: compare (60) (no pronoun) and (63) (object pronoun present).

- (60) *wèrfi kící báandâ-n indéŋŋ^là fâfèré*
 knife old father:1S:GEN-NOM:SG 3:take:PST2 this_morning
 ‘The old knife, my father has taken (it) this morning.’
- (61) *márfâ-náà m-àndŋŋ-nà súk-ti*
 husband-3S:GEN 1S-see:PST1-3S:ACC yesterday
 ‘Her husband, I saw him on the market.’

- (62) *Kùmùlàa* m-àndjɲ-nà súk-ti
 Kumula:ACC:SG 1S-see:PST1-3S:ACC market[Ar.]-ADV
 ‘Kumula, I saw her on the market.’
- (63) *kàggà-kín* náa kùun-nà
 tree-this:ACC:SG how? 3:call:PRS:PL-3S:ACC
 ‘This tree, how is it called? (lit. how do they call it?)’

The topicalized noun phrase (here a personal pronoun) may sometimes be marked either as accusative or as nominative: compare (64a) and (65a) with (64b) and (65b) respectively. In these cases the *animate* feature of the object (patient or experiencer) seems to explain the attraction of a more ‘active’ or ‘controlling’ case – the nominative – although the control is non-existent in both examples (‘be born’, ‘be hungry’).⁷

- (64) a. *niyáà* wjɲèn-n¹áa Fòrbáráŋá
 3S:ACC 3:beget-3S:ACC:PST2 F_B
 ‘He, he was born (lit. they have begotten him) in Foro Boranga.’
- b. *nàalé* wjɲèn-n¹áa Fòrbáráŋá
 3S:NOM 3:beget-3S:ACC:PST2 F_B
 ‘He, he was born (lit. they have begotten him) in Foro Boranga.’
- (65) a. *niyáà* mùsù-n k-ùɲ-n¹à
 3S:ACC hunger-NOM FOC-3:pain:PRS-3S:ACC
 ‘He, he is hungry (lit. hunger pains him).’
- b. *nàalé* mùsù-n k-ùɲ-n¹à
 3S:NOM hunger-NOM FOC-3:pain:PRS-3S:ACC
 ‘He, he is hungry (lit. hunger pains him).’

Finally a contrastive topicalization may appear in a similar way like with the subject (see j. above). The topicalized noun phrase (here a proper name) is marked in the accusative (66).

- (66) *tùbà* *Éli-yàa* cè m-àndjɲ-nà firàangà
 as_for Ali:NOM:SG also 1S-see:PST1-3S:ACC yesterday
 ‘As for Ali, I saw him yesterday.’

1. Topicalized object of ditransitive or applicative verbs?

Sinyar’s ability of topicalizing the supplementary object of a ditransitive or applicative verb is not clear in the present state of the data. Most probably there is a possibility of moving some arguments with pragmatic effects. Contrast for instance (67), where the second object (instrumental) of an applicative verb is put in front of the sentence, with (68), where a similar object appears after the verb predicate (as expected the common noun object is unmarked in either case).

- (67) *wèrɲi* kící ingá n-áaryáa
 knife old what? 2-do:PST1:APPL
 ‘The old knife, what did you do with (it)?’
- (68) ingá áarinnáa wèrɲi dúkùm (= (34) above)
 what? 3:FUTDO:APPL:PL knife blunt
 ‘What will they do with the blunt knife?’

⁷ In English and French however the experiencer is treated as a subject in these precise cases.

But other examples, involving both ditransitive and applicative verbs, are contradictory as to the actual possibility to duplicate the presumed topicalized noun phrase – even referring to an animate – with a personal pronoun *in situ*. On the other hand we lack examples of embedded contrastive topicalization (as in j. and k. above) for this type of argument.

m. Topicalized adverbial

Here again Sinyar may move some adjuncts with a likely pragmatic effect, although this does not appear clearly in the translation. Compare (69a) and (70a) with (69b) and (70b) respectively where adverbial phrases are marked as such in either case. Pronominal duplication seems to be excluded in such cases, even with animate nouns (proper names).

- (69) a. *ingá n-áary¹à wèrfi kiccí-tí*
 what? 2-do:PST2 knife old-ADV
 ‘What have you done with the old knife?’
- b. *wèrfi kiccí-tí ingá n-áary¹à*
 knife old-ADV what? 2-do:PST2
 ‘What have you done with the old knife?’
 (or better ‘The old knife, what have you done with (it)’?)
- (70) a. *c-ábi Kùmùlà-lèe súk-ti*
 1P-FUT:go Kumula-ADV:SG market[AR.]-ADV
 ‘I (lit. we) will go to the market with Kumula.’
- b. *Kùmùlà-lèe c-ábi súk-ti*
 Kumula-ADV:SG 1P-FUT:go market[AR.]-ADV
 ‘I (lit. we) will go to the market with Kumula.’
 (or better Kumula, I will go to the market with (her)’?)

Here however the embedding construction of the contrastive topicalization may apply and, as shown in example (71), the extraposed proper name may appear in the nominative, in the accusative, or in the adverbial case. Despite some slight variations in these three sentences,⁸ the three choices of case marking are clearly possible here.

- (71) a. *tùbà Éli-n cè m-ámbí niyáàrèe súk-ti*
 as_for Ali:NOM:SG also 1S-go:PST1 3S:ADV market[AR.]-ADV
 ‘As for Ali, I went with him to the market.’
- b. *tùbà Éli-yàa cè c-ámbí niyáàrèe súk-ti*
 as_for Ali:ACC:SG also 1P-go:PST1 3S:ADV market[AR.]-ADV
 ‘As for Ali, I (lit. we) went with him to the market.’
- c. *tùbà Éli-lèe cè(n) c-ámbí niyáàrèe súk-ti*
 as_for Ali:ADV:SG also 1P-go:PST1 3S:ADV market[AR.]-ADV
 ‘As for Ali, I (lit. we) went with him to the market.’

n. Topicalized associative

The associative (genitive) modifier of a head noun may be extraposed and duplicated by a personal pronoun *in situ* with an effect of topicalization. As illustrated below a common strategy is to mark the

⁸ The plural *c-ámbí* ‘we went’ (71b-c) is preferred to the singular *m-ámbí* ‘I went’ as in (71a). Shouldn’t it be the case also in (71a)? As mentioned before (see note 12 above) the possible variants of *cè(n)* ‘also’ are not well understood.

topicalized modifier as a nominative in both cases of a common noun (72-73) or a proper name (74-75).

- (72) *wèrfi kící-n táar-nâ-n áffá-tá*
 knife old-NOM:SG edge-3S:GEN-NOM:SG 3:be_sharp:PRS-NEG
 ‘The old knife, its edge is not sharp.’
- (73) *ùmmà-n ningé káalítin tòor-náà-ti*
 woman:NOM:SG 3P:NOM 3:quarrel:PRS:PL top-3S:GEN-ADV
 ‘The woman, they quarrel concerning (lit. at the top of) her.’
- (74) *Kùmùlà-n báa-nâ-n ámbiy¹à nùnnúù*
 Kumula-NOM:SG father-3S:GEN-NOM:SG 3:go:PST2 far_away
 ‘Kumula, her father has gone away.’
- (75) *Máamàt-lè màalé c-ámbí gí-náà-tí màarà-tí*
 Maamat-NOM:SG 1S:NOM 1P-go:PST1 child-3S:GEN-ADV hunting-ADV
 ‘Maamat, I went hunting with his son.’

But accusative (compare (76) with (73) above) or genitive (77) may also be chosen in this context, at least for noun phrases (proper name, interrogative ‘who?’)⁹ that make it possible.

- (76) *Bàkíit-yàa ningé káalítin tòor-náà-ti*
 Bakheet-ACC:SG 3P:NOM 3:quarrel:PRS:PL top-3S:GEN-ADV
 ‘Bakheet, they quarrel concerning (lit. at the top of) him.’
- (77) *dè¹e-nà èðì àd¹àn tádá-náà*
 who?-GEN:SG 3:PRGR 3:say:PRS:PL matter-3S:GEN
 ‘Who are they discussing the matter of?’ (lit. of whom, they are discussing his matter?)

A last example (78) illustrates again the free variation of a nominative with an accusative in marking a proper name.

- (78) a. *Bàkíit-ngè màalé m-èrf¹i báa-ningè-tí firàangà*
 Bakheet-NOM:PL 1S:NOM 1S-speak:PST1 father-3P:GEN-ADV yesterday
 ‘Bakheet’s people, I spoke with their father yesterday.’
- b. *Bàkíit-ngàá màalé m-èrf¹i báa-ningè-tí firàangà*
 Bakheet-ACC:PL 1S:NOM 1S-speak:PST1 father-3P:GEN-ADV yesterday
 ‘Bakheet’s people, I spoke with their father yesterday.’

o. Citation

Considering the lexical lists that have been previously published as well as my own data, the behaviour of nouns in isolation is not really conclusive: according to their nature (common noun or proper name) they may appear as unmarked, accusative or nominative, without obvious link with a feature of animacy. Citation form should therefore not be taken as decisive evidence in Sinyar.

6 Summary of function and case

The different relations between function and case that have been inspected in the preceding section are summarized in Table 6.:

⁹ If the analysis is right, the topicalization of a partial interrogative is rather unexpected.

Table 6: Summary of function and case

	head = common nouns	head = proper names & al.
a. Subject of any predicate and focussed subject of verb predicate	NOM (/ <i>unmarked</i>)	NOM
b. Nominal predicate	NOM	
c. Complement of <i>èbì</i> ‘be, exist, live’ and <i>úngw¹à</i> ‘not be, be absent, not exist’	<i>unmarked</i> / NOM (≈ ‘circumstantial’ / ‘essential’?)	
d. Complement of <i>ùbà</i> ‘become’	<i>unmarked</i>	ACC ?
e. Object (patient)	<i>unmarked</i>	ACC
f. Object of ditransitive or applicative verbs	<i>unmarked</i>	ACC
g. Adverbial (locative, instrumental, and comitative)	ADV	ADV / <i>unmarked</i> (toponyms as absolute locatives)
h. Complement of preposition <i>kándáaká</i> ‘like, as’	<i>unmarked</i>	ACC
i. Associative (modifying a noun, including complex adpositions)	<i>unmarked</i>	GEN
j. Topicalized subject	NOM	
k. Topicalized object	<i>unmarked</i>	ACC (/ NOM)
l. Topicalized object of ditransitive or applicative verbs?	?	
	ADV	ADV
m. Topicalized adverbial	(contrastive topicalization:)	
	?	NOM / ACC / ADV
n. Topicalized associative	NOM	NOM / ACC / GEN
o. Citation	not conclusive	

7 Concluding remarks

In order to conclude this presentation of Sinyar’s case marking system(s), several remarks can be made.

1. From a typological point of view König (2008) makes the following observation concerning African languages:

“[...] two types of marked-nominative languages are to be distinguished with regard to the morphological markedness of nominative and accusative: Type 1 (the most common one), in which the [nominative] is the morphologically marked form and the [accusative] the morphologically unmarked form, and Type 2, in which both case forms, nominative and accusative, are morphologically marked.” (König 2008:254).

If so, Sinyar results from an interesting combination of the two types: the first one is valid for common nouns and the second one is valid for proper names and assimilated (deictics and personal pronouns). They both contribute to establish the specificity of the language under this respect.

2. According to the same author, the distribution of marked-nominative languages is as follows: Berber, Cushitic, Omotic, Nilotic, and Surmic (König 2008: 271). It must then be underlined that Sinyar is spoken in an area that is rather characteristic of accusative languages: Maba, Merarit, Masalit, Tama, and Fur (König 2008:266). This unexpected situation raises again the question of the origin and historical constitution of Sinyar that looks, at first sight at least, as resulting from the contact between different, only partially identified sources (see Introduction).

3. Coming back to the case system, it has been shown that the two sets constituted by common nouns on one hand and proper names, deictics and pronouns on the other hand do not function in a strictly parallel way. This situation does not result so much from the imbalance between the number of cases available for either group as from the features of animacy and specificity that are strongly dominant in the second set.

4. Finally the choices offered – or imposed – by the system for marking the different functions assumed by noun phrases are not equally straightforward. The most uncertain strategies appear in the case of topicalized arguments or adjuncts: variation may appear in the case of topicalized objects and, even more, with topicalized adverbials and associatives. This uncertainty seems to result from the fact that the topicalized element is functionally disconnected from the sentence and unloaded from a syntactic role that is usually taken over by its pronominal duplicate.

Abbreviations

1P	1 st person pl.	FUT	Future
1PD	1 st person pl. (dual)	GEN	Genitive
1PP	1 st person pl.	H	High tone
1S	1 st person sing.	IMP	Imperative
2	2 ^d person	L	Low tone
2P	2 ^d person pl.	NEG	Negative, Negation
2S	2 ^d person sing.	NOM	Nominative
3	3 ^d person	PL	Plural
3P	3 ^d person pl.	PRGR	Progressive
3S	3 ^d person sing.	PRHB	Prohibitive
ACC	Accusative	PRS	Present
ADV	Adverbial	PST1	Past 1
APPL	Applicative	PST2	Past 2
AR.	Arabic	REL	Relative
DEP	Dependent (verb form)	SBB	(Sara-)Bongo-Bagirmi languages
FOC	Subject focalizing (verb form)	SG	Singular

N.B. The characterization of TAMs as ‘Present’, ‘Future’, ‘Past 1’, and ‘Past 2’ is tentative and provisional.

References

- Bender, M. Lionel. 1992. Central Sudanic segmental and lexical reconstructions. *Afrikanistische Arbeitspapiere* 29: 5-61.
- Doornbos, Paul [undated]. *The Language of the Senyar*.
- Doornbos, Paul & M. Lionel Bender. 1983. Languages of Wadai-Darfur. In M. Lionel Bender (ed.), *Nilo-Saharan Language Studies*, pp. 42-79. East Lansing: African Studies Center, Michigan State University.
- Haaland, Gunnar. 1978. Ethnic groups and language use in Darfur. In Robin Thelwall (ed.), *Aspects of Language in the Sudan* (Occasional papers in linguistics and language learning, vol. 5), pp. 181-199. Belfast: The New University of Ulster.
- Kießling, Roland. 2007. The ‘marked nominative’ in Datooga. *Journal of African Languages and Linguistics* 28.2: 149-191.
- König, Christa. 2008. The marked-nominative languages of eastern Africa. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*, pp. 251-271. Cambridge: Cambridge University Press.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the world, seventeenth edition*. Dallas, Texas: SIL International. <http://www.ethnologue.com> (last accessed 25 September 2014).

Tucker, A. Norman & Margaret A. Bryan. 1956. *The Non-Bantu Languages of North-Eastern Africa*. London/New York/Cape Town: OUP for IAL.

CASE MARKING IN TUGEN

Prisca Jerono

1 Introduction

Case is used to indicate the relationship of the verb and its core noun phrases. It is used to disambiguate various noun phrases that appear in clauses. Dimmendaal (1982: 65) says that case coding refers to the way in which relationships within a sentence between verbs and other categories are indicated. The means by which these categories are coded (marked) indicate how they function syntactically. Languages mark case by use of word order, (structural case) morphological inflection, tone or a combination of any of the above. Case marking is responsible for the different alignment systems in languages. The alignment systems are based on how the subject (S) of an intransitive sentence is marked in comparison to the subject (A) of a transitive sentence and also how the object (O) is marked (Dixon 1994). Specifically the alignments systems are nominative-accusative, ergative-absolutive and marked-nominative. In the nominative-accusative system, S and A are marked in the same way and O is marked differently. In the ergative-absolutive system S and O are marked in the same way while A is marked differently. The marked-nominative system marks A and S in a similar manner while O remains unmarked. The difference therefore between the nominative-accusative system and the marked-nominative system is that in the nominative-accusative O is marked in some way but in the marked-nominative system O remains unmarked. Tugen¹ falls within the marked-nominative alignment system. In Tugen case marking is achieved morphologically and suprasegmentally by the use of tone or a combination of the two. Tone refers to a significant rise or fall in pitch in a word and can be used to signal lexical, morphological, semantic and pragmatic information. Tone is normally analyzed as a property not of individual segments (or sounds) but of syllables. It is always transcribed on the syllable nucleus which is usually a vowel; this masks the fact that tone may be phonetically realized on the voiced sonorant² segments. In this paper, tonal marking is analysed for the various case forms that manifest as a result of the tonal inflection as well as the tone patterns that occur. The analysis of tone is done on the basis of a sequence of H (High) and L (Low) tones. In this paper we borrow from the Autosegmental Theory the principle of feature spreading to describe the successive manifestation of a tonal feature (Goldsmith 1976). For instance, a successive occurrence of H tones is taken to be a result of one H tone spreading to other segments, thus it will be described as one tone but it will be marked separately in the data provided. Association lines will however not be used to represent the data. This analysis also takes into consideration the number of syllables in a noun in determining its tone pattern. Most of the nouns in Tugen have more than two syllables. This paper analyses morphological and tonal case marking.

2 Tonal marking

The basic Tugen word order in isolated sentences is VSO/VOS. This means that it is possible to permute the order of the subject and the object and the meaning of the sentence still remains the same. For example:

- (1) a. sóómán-í láákwéé kitábúú
read-IMP child book
'The child is reading a book.'

¹ Tugen is one of the dialects within the Kalenjin macro language of the Southern Nilotic group of languages. It is spoken by the community of people living in Baringo County of Kenya. Tugen has a basic word order of VSO in isolated sentences.

² Nasals, glides and liquids.

- b. sóómán-í kitábúú láákwéé
 read-IMP book child
 ‘The child is reading a book.’

In order to disambiguate the meaning of the sentence in such a construction, tonal differentiation on the subject and the object occurs. According to the Minimalist Program (Chomsky 1995) all the inflectional properties of a word are found in the lexicon, and in the construction of a sentence the various words are selected and merged and finally the sentence is checked for all features at spell-out before the different logical features and phonological features are directed to the separate domains for pronunciation and interpretation. In line with this argument, all properties of a word including case are inflected/assigned in the lexicon. The morphological features are checked overtly before spell out and because tone is a phonological feature, it is checked covertly before spell out. The tonal features that are assigned to a noun phrase depend on the function of the NP in the sentence structure. In Tugen NPs indeed are assigned lexical tonal features that are said to be in the absolutive. The absolutive in this case refers to the state of an NP in terms of the tonal markings that the NP bears as a lexical item without reference to any syntactic function.

Absolutive

Tugen is a marked-nominative language because S and A are marked while O remains unmarked. The absolutive is used to refer to the unmarked case to differentiate it from the marked accusative of the nominative-accusative alignment. It is used with NPs that function as direct objects and indirect objects as well as NPs that are used in isolation. It is also used in noun phrases that are moved out of the core sentences for reasons of emphasis or topicalization. Dimmendaal (1982: 66) says that for Turkana the tone patterns of nouns, pronouns and other categories occurring in an NP in the absolutive case is taken to be the basic pattern because:

- i. in this form these categories have their widest range of occurrence,
- ii. maximal tonal differentiation occurs, whereby the tone pattern itself is not predictable,
- iii. the tonal forms of the categories in question when occurring in other cases can be predicted or derived from the corresponding forms in the absolutive case.

Tones in Tugen are usually marked according to the number of syllables in a word and more specifically on the syllable nucleus, which is normally a vowel. Most of the nouns in Tugen have more than one syllable. The tones patterns occurring in nouns with various syllables are shown in the following table:

Table 1: Tone patterns in nouns in nominative and absolutive

(2)	Noun: ABS	Tone pattern	Noun: NOM	Tone pattern	Gloss
	láákwéé	HLH	láákwéé	H	child
	káámèè	HL	káámèè	L	mother
	mürsiik	LH	mürsiik	L	sour milk
	chèègò	LH	chèègò	L	milk
	cháálwóók	HL	cháálwóók	H	wrongdoing
	kwááriik	H	kwááriik	HL	tendon
	sáándé	HLH	sáándé	H	groom
	ííléé	HLH	ííléé	H	lightning
	chèèmòòsí	LH	chèémóósí	H	ogre
	móítá	H	mòità	L	calf
	bèelyóóndé	LHLH	béélyóóndé	H	elephant
	kitábúú	LH	kitábúú	H	book
	làábóótík	LH	làábóótík	H	soddom apple
	káályááng'ík	H	kàályááng'ík	LH	flies

ng'èchèróóník	LHLH	ng'èchèróóník	H	seats
síyòík	LH	síyòík	H	nails
òinòòsyèk	HLH	òinòòsyèk	HL	rivers
kisiryáándé	LH	kisiryáàndè	L	louse
kimútái	LH	kimútái	H	John
kipkòòskéí	LHLH	kípkòóskéí	H	James
kibùkándii	L	kibùkándíí	LH	guitar
kipchóriíréé	LHLH	kípchóriíréé	H	squirrel

From this analysis we can see that these nouns in isolation have the following patterns in the absolutive: HLH, HL, L, H and LHLH. The tone patterns of the nouns cannot be predicted depending on the number of their syllables, though most of them bear LH tones. When they appear as objects in sentences these nouns bear the same patterns as seen in the examples below:

- (3) a. Kà³- ø-rèèrí mòità tèetà (L); (LH)
 PST3-3SG-suck calf:NOM cow:ABS
 'The calf suckled the cow.'
- b. Kii- lú chèègó mòità (LH); (L)
 PST1-drink milk:ABS calf:NOM
 'The calf drank the milk.'
- c. Kii-rfiir-chí⁴ láákwéé káámèè (H); (HL)
 PST1-cry-for child:NOM mother:ABS
 'The child cried for the mother.'
- d. Kòò-ám-éí sáándé kwáárik (H); (H)
 PST2-eat-IMP groom:NOM tendons:ABS
 'The groom was eating hooves.'
- e. Kii-pút íiléé kibùkándii (H); (L)
 PST1-destroy thunder:NOM guitar:ABS
 'The thunder destroyed the guitar.'

In (3)a. the sentence with the subject *moita* 'calf' bears the nominative L tones and is different from when used in isolation as seen in (2) where it bears the absolutive H tones. (3)b. is a sentence with the subject *moita* 'calf' and the object *cheego* 'milk'. The subject bears the marked nominative L tones while the object remains with the absolutive tone patterns of LH. The same absolutive tone patterns are also seen in the tonal patterns of the nouns *kaamee*, 'mother' *kwaarik* 'tendons' and *kibukandii* 'guitar' which take the direct object roles. When these nouns are moved out of their core positions for topicalization or emphasis, they retain the tonal patterns as they appear in the isolated forms. This is seen in the examples below:

- (4) a. Kimútái, kii-ø-wó gáá (LH)
 Kimutai PST1-3SG-go home
 'Kimutai, he went home.'
- b. Ng'èchèróóník ché kà-i-bút-yó (LH)
 Seats that PST3-3PL-fall-IMP
 'It is the seats that fell.'

³ Tugen has three prefixes for tense: *kii-* for the distant past, *koo-* for the recent past and *ka-* for the immediate past.

⁴ The suffix *-chi* can be used for the allative role of towards and the dative role of to/for .

- e. Kii-róch láábóótík kòndá. (H); (LH)
 PST1-enter sodom apples:NOM eye:ABS
 ‘The sodom apples entered the eye.’

In the examples above (6)a. *erene* ‘snake’ changes from a LH tone pattern in the absolutive case as seen in (8)j. to H tone in the nominative. In (6)b. the noun *kaamee* ‘mother’ changes from HL tone in the absolutive to L tone in the nominative. In (6)c. the noun *kwaaryaande* ‘tendon’ changes from LH tone in the absolutive to L tone in the nominative case. In (6)d. the noun *kaalyaang’ik* ‘flies’ also changes from H tone in the absolutive to LH tone in the nominative. For these examples the tone patterns for the nominative are H and L tones. Other tone patterns for the nominative case are seen in the examples below:

- (7) a. Kòò-í-mút Kímútái tòondé (H); (LH)
 PST2-3SG-take Kimutai:NOM guest:ABS
 ‘Kimutai took the guest.’
- b. Kòò-í-bút-yó ng’èchéroóník ééng’ tuiyéé (H)
 PST2-3SG-fall-IMP seats:NOM in meeting
 ‘The seats fell in the meeting.’
- c. Kii-súús-àn kisiryàandé (L)
 PST1-bite-1SG louse:NOM
 ‘A louse bit me.’
- d. Míí-tó kípchóriíré mbàréé (LH); (LH)
 be-IMP squirrel:NOM farm:ABS
 ‘There is a squirrel in the farm.’
- e. Kà-yóón sééséé Kipkóòskéi (HLH); (LHLH)
 PST3-chase dog:NOM Kipkooskei:ABS
 ‘The dog chased Kipkoskei.’

In the examples above, the tone patterns of nouns such as *Kimutai* ‘John’ change from LH in the absolutive to H in the nominative. A noun such as *ng’echeroonik* ‘seats’ has a LH tone pattern in the absolutive and H tone in the nominative while *kipchoriire* ‘squirrel’ changes from LHLH in the absolutive to H tone in the nominative. *Kisiryande* ‘louse’ has a LH tone in the absolutive but a L tone in the nominative while *seesee* ‘dog’ has a LH tone pattern in the absolutive and a HLH tone in the nominative.

In the foregoing analysis most nouns have more than two syllables. The nouns bear various tone patterns in the nominative. Some involve a slight modification of the tone pattern in the absolutive form while others involve a complete change of the tone pattern. For example the noun *moita* ‘calf’ has H tone in the absolutive but L tone in the nominative. A noun such as *kaalyaang’ik* ‘flies’ has a H tone in the absolutive but L tone in the nominative; a noun such as *seesee* ‘dog’ has a HL tone in the absolutive and a HLH in the nominative. From the 22 examples given in 2 above, the distribution of tone patterns for the absolutive are as follows: H-3, L-1, LH-8, HL-2, LHLH-4 and HLH-4. The differences in the distribution of tone patterns between the nominative and the absolutive is as shown in the table below:

Table 2: Tone patterns between nominative and absolutive

(8)	Noun:nom	Tone pattern	Noun:abs	Tone pattern	Gloss
	láákwéé	H	láákwéé	HLH	child
	chéémóósíí	H	chèémóósíí	LH	ogre
	káámèè	L	káámèè	HL	mother
	cháálwóók	H	cháálwòók	HL	wrong
	òinéé	LH	òinèè	L	river
	kibúkándí	LH	kibúkàndi	HL	guitar
	mòità	L	móítá	H	calf
	téétá	H	tèètá	LH	cow
	érené	H	èrené	LH	snake
	kwàryààndè	L	kwàryááandé	LH	tendon
	káályááng'ík	LH	káályááng'ík	H	flies
	kòondá	L	kòondá	LH	eye
	láábóótík	H	làábóótík	LH	soddom apple
	kímútáí	H	kímútáí	LH	John
	séèséé	HLH	séèsèè	HL	dog

From the examples in the table we find seven nouns having a H tone in the nominative and only two with a H tone in the absolutive. On the other hand there are seven nouns with a LH tone in the absolutive and only three with a HL tone in the absolutive. In the nominative four nouns have a L tone and in the absolutive two nouns have a L tone. The nouns have between two to four syllables. Some of the nouns involve a slight modification in the tone pattern between the absolutive while others involve a complete change of tone. For this reason we can conclude that the change in tone pattern is irregular.

3 Morphological marking

In Tugen, gender, the genitive, and double object constructions are marked for case morphologically.

Gender

The case on gender is marked formally by way of gender sensitive prefixes in nouns and modifiers. Tugen marks natural gender⁵ by the masculine prefix *Kip-* and the feminine prefix *Cheep-*⁶. However not all nouns that refer to natural gender have masculine and feminine forms. Other gender forms are marked lexically and used as modifiers. An example is (9)e. below. The gender prefixes change their tone patterns from the absolutive to the nominative when they appear as subjects in a sentence as seen in (8) in the table above and (10) below:

	feminine	masculine	gloss
(9) a.	Chèèróónó	kipróónó	Mary/John
b.	chèèp-tuuge	_____	dove
c.	_____	ki-siryáán-dé	louse
d.	chèèp-íy-wéé	_____	mad person
e.	chèèp-ò	àrááp	daughter of/son of

⁵ The natural gender in Tugen sometimes is referential/semantic in the sense that the female gender prefix may be used to refer to a masculine thing. For example the name *Cheerono* 'Jane' is sometimes used to refer to the masculine gender while *Kip-lekwe* 'hare' bears a masculine prefix.

⁶ The *-p* is sometimes omitted due to some phonological processes. For example in *Ki-maru* 'Peter' and *Cheerono* 'Jane'.

Gender and nominative marking is exemplified in the sentences below:

- (10) a. Kà-biir-ón chéép-iy-wéé (H)
 PST3-beat-1SG F-mad person
 ‘The mad person beat me.’
- b. Kòò-nyó chéép-ò Lùùká ámút (HL); (LH)
 PST2-come daughter of Luke yesterday
 ‘Luke’s daughter came yesterday.’

In (10)a. the tone of the gender prefix changes from L to H when the noun phrase is in the nominative. This is the same case with (10)b., which also has the associative marker *-o* to show possession/genitive case. The tone on the associative marker does not change. Another associative marker for the genitive case is discussed below.

The genitive case

Other than gender, Tugen has another case form that is marked morphologically. This is genitive case⁷. In Tugen the genitive case is marked in noun phrases that function as possessors in an associative construction. These noun phrases are inflected by way of morphology in order to indicate this function. The association between the noun phrases involved is achieved by the suffix *-ap* which is affixed on the possessed noun phrase. For example:

- (11) a. Kà-i-gát-án⁸ láákwéét-áp Chèèróónó (H); (LH)
 PST3-3SG-greet-1SG child:NOM-of Cheerono:ABS
 ‘Cheerono’s child greeted me.’
- b. Kòò-pét káláámít-áp chèèptó (LH); (LH)
 PST2-loose pen:NOM-of girl:ABS
 ‘The girl’s pen got lost.’
- c. Ki-téch-éí kòòt-áp kòònèètíndé (LH); (LH)
 PASS-build-IMP house:ABS-of teacher:ABS
 ‘The teacher’s house is being built.’

In the examples above, the possessor noun phrase remains in its absolutive case form while the possessed noun phrase changes its tone pattern according to whether it is used as a subject or an object. In (11)c. the sentence is in its passive form and the passive subject bears the absolutive case form.

4 Case in double object constructions

Other than the above cases, there are several case forms that are used for the applied object. The applied object occurs in the sentence structure as a result of derivative affixes which introduce them. The presence of the derivative affixes affects the valency of the verb and results in the creation of a double object construction, i.e the direct object and the applied object. Valency refers to the range of syntactic elements either required or specifically permitted by a verb or any other lexical unit (Mathews 1997). The applied object in Tugen is represented by an NP that bears various case forms. The cases involved are the benefactive, the instrumental and the locative (Jerono 2012). The case forms are licensed by the derivative suffixes on the verb and therefore, when a verb is selected, it is

⁷ Also called possessive.

⁸ In some verbs the 3SG prefix is *-i-* which is similar to the 2SG prefix. For example in *ka-i-sup lakwee* (PST3-3SG-follow child) ‘the child followed’. See also (13)b.

In the examples above the subject is marked by H tones and the direct object is marked by LH tones while the applied object is marked by HLH tones. The direct and the applied object are disambiguated tonally and also by the semantics of the lexical items in that the *rootwe* (knife) is an instrument for cutting and therefore is given the instrumental function.

Locative

The locative is another case form that arises in Tugen as a result of the presence of derivative affixes on the verb. In Tugen, the derivative suffix for the locative *-en* is similar to the one for the instrumental NP. This is shown below:

- (14) a. Kòò-pá-én sááng' láàkwéé kààmèè (H); (HLH); (L)
 PST2-feed-LOC outside:ABS child:ABS mother:NOM
 'The mother fed the child outside.'
- b. Kii-ál-én tùùgá síró chéépyóóséé (LH); (H); (H)
 PST1-buy-LOC cows:ABS market:ABS woman:NOM
 'The woman bought the cows at the market.'

In the above examples the subject is marked for nominative case by H tone while the direct and the applied object bear tone patterns that are similar to the tone patterns of the absolutive case. The tone pattern for the locative is H tone and similar to the nominative H tone for the subject. In order to disambiguate the two NPs the semantics of the noun phrases are considered, whereby *siiro* 'market' is a name for a place and thereby is assigned the locative case. This case also applies to the locative noun *saang* 'outside' in (14)a. Another locative case on NPs that show direction is derived by the directional suffix *-u* on the verb. This suffix introduces noun phrases that take the locative case. The locative noun phrase bears a H tone. For example:

- c. Bún-ú gáá Chééroónó (H); (H)
 come-ALL home:ABS cheeroono:NOM
 'Cheeroono is coming from home.'

From these examples it can be seen that the locatives bear H tones and there is no ambiguity in terms of the entities involved because the suffixes introduce adverbial noun phrases of location.

5 Case in loan words

In Tugen, words that are borrowed from other languages are adapted to fit in to the phonology of the language. Case marking is also not an exception on these words. The loan words are assigned lexical tonal markings and absolutive case in the lexicon. In isolation therefore loan words are inflected for absolutive case with various tone patterns as seen in the examples below:

Table 3: Tone patterns for loan words

(15)	Noun:ABS	Tone pattern	Noun:NOM	Tone pattern	Gloss	Source
a.	kòòmpyùtá	LLH	kóòmpyúútá	H	computer	English
b.	kikóòmbéé	LHLH	kíkóòmbéé	HLH	cup	Kiswahili
c.	swéétéé	HLH	swéétéé	H	sweater	English
d.	sííméé	LH	sííméé	H	telephone	Kiswahili
e.	ùsíì	LHL	úsíí	H	thread	Kiswahili
f.	pètèròòl	L	pètèròòl	HL	petrol	English

In the table the tone patterns for the loan words in the absolutive are LH, HLH, L, LHL and LHLH. The tone patterns do not change when the nouns are used as objects in a sentence to reflect the unmarked absolutive case. This is seen in the examples below:

- (16) a. Kà-sir-en kòòmpyùútá ngáléék láákwéé (LLH); (LH); (H)
 PST3-write computer:ABS words:ABS child:NOM
 ‘The child wrote the words with a computer.’
- b. Kà-í-al kíkóòmbéé (LHLH)
 PST3-2SG-buy cup:ABS
 ‘You bought a cup.’
- c. rîir-éí sîiméét-áp Kip-túùm (HLH); (LHL)
 ring-IMP phone-of:ABS M-tuum:ABS
 ‘Kiptum’s phone is ringing.’
- d. tíny-éí Mária swéètéé (H); (HLH)
 have-IMP Maria:NOM sweater:ABS
 ‘Maria has a sweater.’

When the noun phrases are used in sentences as subjects, the tone patterns change to reflect the marked nominative case. This is exemplified below:

- (17) a. Míning’ swéètèè (H)
 small sweater
 ‘The sweater is small.’
- b. sír-éí ngáléék kóòmpyúútá (LH); (H)
 write-IMP words:ABS computer:NOM
 ‘The computer is writing the words.’
- c. tóórét-í pétéròòl kàríí (H); (HL)
 help-IMP Petrol car
 ‘Petrol aids a car.’
- d. náb-éí úsí kómié (H)
 sew-IMP thread:NOM well
 ‘The thread sews well.’

In the examples above, the NP *sweetee* ‘sweater’ changes its tone pattern from the HLH in (15)c. to HL in (17)a. to reflect its marked nominative case. In (16)b. the noun *koompyuuta* ‘computer’ changes its tone pattern from a LH tone in the absolutive to a H tone in the nominative. In (17)c. the noun phrase *petrol* ‘petrol’, which is the subject, has a HL tone pattern and differs from its unmarked absolutive case in (15)f., which is a L tone. The noun *usii* ‘thread’ in (16)d. has a H tone pattern in the nominative and differs from its absolutive counterpart in (15)e. which has a LHL tone pattern.

From the data given it is clear that Tugen does not treat loan words differently from indigenous words. All noun phrases in the language have tone patterns that reflect the unmarked absolutive case either in isolation or as objects in sentences. The noun phrases are marked by different tone patterns for the nominative when they are subjects, irrespective of whether they are foreign or indigenous.

6 Conclusion

This paper set out to investigate how case is marked in Tugen, a dialect of the Kalenjin group of the southern Nilotic branch. In doing so it attempted to find out the various case forms as well as the case marking strategies both morphological and tonal. It emerged that Tugen is a marked-nominative language whereby the subject of an intransitive sentence (S) and the subject of a transitive sentence (A) are marked in a similar manner while the object (O) remains unmarked. This marking is done

suprasegmentally whereby tones that appear in the absolutive form change when an NP is in the nominative. The applied objects are also marked by derivational suffixes which introduce applied objects into the sentence structure. The applied objects are also marked by case forms reflecting the benefactive, the instrumental, and the locative case. Natural gender is marked by the gender prefixes *cheep-* and *kip-* while the genitive case is shown by the associative suffixes *-ap* and *-o* to show possession. The gender prefixes have tonal patterns similar to those of the absolutive form unless when used in the nominative when the tonal patterns change. The tonal patterns of the genitive associative suffix remain unchanged. From this we can see that the tonal patterns for the absolutive case have a wide range of occurrence in that it is used for the direct object, the applied objects, for emphasis/contrast and for nouns in isolation.

Tonally, Tugen marks absolutive case through different tonal patterns that are not predictable in reference to the number of syllables in a noun. The nominative case is marked differently from the absolutive case whereby the tonal pattern of noun in the absolutive case is changed slightly or completely when in the nominative. From the analysis it emerged that there is a preference for H tone patterns for the nominative case, although other tone patterns also exist.

This paper also set out to investigate how case in double object constructions is disambiguated. Double object constructions arise as a result of valency changing operations whereby the derivational suffixes introduce applied objects to the sentence structure besides the direct object. They also involve the benefactive, the locative and the instrumental. These derivational suffixes introduce applied objects which bear the respective case forms. The case forms for the applied objects have the same tone patterns as those for the nouns in their absolutive case. In order to disambiguate the direct object from the applied objects, the tonal patterns and the semantics of the derivational suffixes help in disambiguating the applied from the indirect object. For example in the benefactive, the derivational suffix introduces an animate object which benefits from the action and is disambiguated from the direct object, which is the one being directly affected by the action while for the instrumental, the derivative suffix introduces an inanimate object that is used for performing an action. Tugen does not differentiate loan words from indigenous words in terms of case marking. Loan words in the language are marked with the absolutive case in isolation and when used as objects, they bear the same tone patterns as those of nouns in the absolutive. When the loan words are used as subjects in sentences, they bear the marked nominative case with different tonal patterns to show this.

Abbreviations

A	transitive subject	O	object
ABS	absolutive	PL	plural
H	high tone	PST	past
IMP	imperfective	S	intransitive subject
L	low tone	SG	singular
NOM	nominative		

References

- Creider, Chet A. & Jane T. Creider. 1989a. *Grammar of Nandi*. Hamburg: Buske.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Dimmendaal Gerrit J. 1982. *The Turkana Language*. PhD thesis, University of Leiden.
- Dixon, Robert M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Duanmu, San. 1996. Tone: an overview. *Glott International* 4.2: 3–8.
- Goldsmith, John. 1976. *Autosegmental Phonology*. PhD thesis, Massachusetts Institute of Technology.
- Jerono, Prisca. 2012. *Tugen Word Order: A Minimalist Perspective*. Macedonia, EGALITE.
- Matthews, Peter H. 1997. *Oxford Concise Dictionary of Linguistics*. Oxford/New York: Oxford University Press

CASE AS A META-CATEGORIAL HEURISTIC IN IK GRAMMATICOGRAPHY

Terrill Schrock

1 Introduction

Ik, a Kuliak (Nilo-Saharan) language spoken in northeast Uganda, has a nominal case system with eight cases. Not only must all nouns be inflected for case, but case seems to have nearly the entire grammar in its grip. Case suffixes have been grammaticalized as verbal suffixes and possibly as markers of tense on demonstratives, relative pronouns, and tense clitics. As many as seven of the eight cases have formal parallels in other grammatical categories. For example, while the suffix *-ka* marks the accusative case on nouns, it also marks the present perfect tense on verbs and recent past tense on demonstratives. And the suffix *-kɔ*, though it marks the copulative case on nouns, also marks the subsecutive mood on verbs and the remote past tense on demonstratives.

Such widespread morphological homophony was partially explored in König's study of Ik case (2002), which included proposed case-suffix chains of grammaticalization. This paper expands on those proposals and shows how 'case' can act as a meta-categorial heuristic in ongoing Ik analysis and grammaticography. It also reveals how a growing understanding of nominal case suffixes has shed light on other grammatical systems. By presenting the Ik nominal cases and their parallels across morphological categories, this paper confirms that case is one of the Ik language's most central grammatical characteristics. Ik is spoken by roughly 7,000 people in northeast Uganda. Its grammar has been described to various degrees over the last few decades (e.g. in Crazzolara 1967, Tucker 1971/1972/1973, Heine 1983, Serzisko 1992, and König 2002), but it is the language's case system that has garnered the most attention among linguists. Recent treatments of case include König 2002, an in-depth study of case and case grammaticalizations; König 2008, with an updated chapter on Ik case; and Schrock 2012, a paper introducing the previously overlooked instrumental case. Despite all that has already been written, the topic of case in Ik is far from being exhausted.

In particular, the question of case markers grammaticalized into other grammatical categories continues to yield new insights. With eight nominal case markers, the Ik case system is already complex by African standards. But as others have pointed out, Ik case markers are found well beyond the noun word class. König states that "[Case in Ik] is ... the most productive morphological tool in the language ... As in other languages, case in Ik is a category which is connected with nouns and pronouns, but case is not restricted to these word classes. Function words like conjunctions, postpositions, prepositions, adverbs, and even verbs are also inflected for case" (2008: 81).

But unless we revise the traditional understanding of 'case' as a characteristic of nominal systems, then strictly speaking, Ik case *per se* must be restricted to nominals. The near ubiquity of morphemes identical or similar to nominal cases in Ik grammar suggests that case markers have been grammaticalized over time into other grammatical systems. In other words, from a diachronic point of view, 'case' in Ik can be thought of as a meta-categorial property of the grammar while from a synchronic point of view, one must consider the different functions and meanings of the grammaticalized case markers in each sub-system of the grammar.

This paper is a further exploration of some of the places in Ik grammar where morphemes similar or identical to Ik nominal case markers have so far been found. It begins with a summary of the nominal case markers as a basis for the subsequent discussion. Then it examines grammaticalized case markers in verbs, demonstratives, relative pronouns, and tense clitics. Each section gives examples of how 'case' has become a key heuristic in the analysis and description of all aspects of Ik grammar.

2 Case suffixes in the nominal system

The syntax of Ik case has been thoroughly dealt with elsewhere (e.g. König 2002, 2008; Serzisko 1992), so only a few comments are made here. Ik has a split-accusative system. It is accusative in that when core clausal arguments are marked differently from each other, S and A take the nominative case, and O takes the accusative case. And it is split in that this accusativity only occurs when the subject (A) is third person. Ik's basic clausal constituent order is VSO. In most instances, when a subject (A/S) is preposed before the verb, for example in subordinate clauses, it takes the accusative case.

Ik has eight morphological case markers. Seven of them are suffixes, and one of them is a zero morpheme. Table 1 below presents these cases and their morphological markers. One important feature to note is that the suffixes differ with regard to their 'suffixation strategy'. The nominative and instrumental suffixes first delete the final vowel of the nominal stem to which they affix. By contrast, the other six case suffixes affix directly to the stem-final vowel, allowing it to remain:

Table 1: Ik case markers

Case	Abbreviation	Marker	Suffixation strategy
Nominative	NOM	-a	Deletes stem-final vowel
Instrumental	INS	-ɔ	
Ablative	ABL	-ɔ	Attaches to stem-final vowel
Genitive	GEN	-ε	
Accusative	ACC	-ka	
Dative	DAT	-kε	
Copulative	COP	-kɔ	
Oblique	OBL	-∅	

The seven overt case morphemes and one zero case morpheme (oblique) are subject to a variety of morphological alternations. Schrock (2012) covers these in more detail, but the table below serves to summarize them for the sake of the following discussion:

Table 2: Ik case allomorphs

Case	Underlying form	Vowel harmony	Vowel assimilation	Non-final form	Final form
NOM	-a			-a	- ^a
INS	-ɔ	-o		-ɔ/o	- ^ɔ /o
ABL	-ɔ	-o		-ɔ/o	- ^ɔ /o
GEN	-ε	-e	-i/i	-ε/e	- ^ε /e ⁱ /i ⁱ
ACC	-ka			-a	-k ^a
DAT	-kε	-ke		-ε/e	-k ^ε /k ^e
COP	-kɔ	-ko		-ɔ/o	-k ^ɔ /k ^o

As Table 2 shows, Ik case markers can take on a variety of surface forms. Besides [ATR]-based vowel harmony, the most pervasive case allomorphs are those affected by their position in the clause, namely clause-medially or clause-finally. The accusative, dative, and copulative case markers all lose the segment /k/ in their non-final forms. All case markers have their final vowel reduced (devoiced or deleted) in their final forms.

3 Case markers grammaticalized as verbal TAM suffixes

As König (2002) amply demonstrated, Ik has grammaticalized nominal case markers into use as verbal TAM suffixes. The most well-known examples of this are 1) the dative case turned marker of the subjunctive mood (SUBJ) and 2) the copulative case turned marker of the subsecutive mood (SUBS):

(1) Dative > Subjunctive

ηókí-k ^e	dog-DAT	ats-íí-k ^e	come-1SG-SUBJ
sísí-k ^e	mead-DAT	ats-ídi-k ^e	come-2SG-SUBJ
bóré-k ^e	corral-DAT	àts-i-k ^e	come-3SG-SUBJ
ηèkè-k ^e	hunger-DAT	ats-ísini-k ^e	come-1PL.INCL-SUBJ
ηorá-k ^e	cane.rat-DAT	ats-ími-k ^e	come-1PL.EXCL-SUBJ
		ats-ítí-k ^e	come-2PL-SUBJ
		ats-áti-k ^e	come-3PL-SUBJ

(2) Copulative > Subsecutive²⁰

ηókú-k ^o	dog-COP	ats-íà-k ^o	come-1SG-SUBS
sísó-k ^o	mead-COP	ats-ídù-k ^o	come-2SG-SUBS
bóré-k ^o	corral-COP	àts-ù-k ^o	come-3SG-SUBS
ηèkè-kò	hunger-COP	ats-ísinù-k ^o	come-1PL.INCL-SUBS
ηorá-k ^o	cane.rat-COP	ats-ímà-k ^o	come-1PL.EXCL-SUBS
		ats-ítú-k ^o	come-2PL-SUBS
		àts-in	come-3PL-SUBS

The data in (1) show that the dative case suffix and the subjunctive verb suffix are identical in form (-*k^e* and its various allomorphs). And as shown in (2), the same formal identity is found between the copulative case suffix and the subsecutive suffix. According to König (2002: 318), the subjunctive is formed by a combination of the optative suffix /-i/ and the dative suffix /-*k^e*/. In my analysis, however, the /i/ is not the optative but rather the final vowel of the preceding pronominal subject-agreement markers. Also, König seems to have overlooked the fact that the subsecutive ('narrative' in her terminology) is formed in exactly the same way as the participial, only with the copulative case instead of the dative. Many of the allomorphs König posits for the subsecutive (see 2002: 409) actually consist of the copulative plus the final vowel of the subject-agreement suffixes (whose interaction is subject to a variety of morphophonological changes). But even though my analysis departs from König's on a number of morphological details, I concur with her overall argument that the dative and copulative cases have been grammaticalized as verbal TAM suffixes.

According to König, the following points support the argument that the subjunctive suffix is a grammaticalization of the dative case suffix (2002: 318-319):

1. Their forms are identical.
2. Their morphophonological behavior is identical.
3. Their identity is suggested by the similar identity between the subsecutive suffix and the copulative case suffix.
4. Their position in the word (i.e. always at the end) is identical.
5. The participial and subsecutive exhibit similar syntactic behavior (e.g. they are both used in subordinate clauses).
6. In other languages, there are examples of morphemes from the nominal domain are used for clausal embedding (like the subjunctive).

²⁰ The subsecutive verb paradigm is irregular with respect to the 1SG suffix and the 3PL suffix, the latter of which is a portmanteau morpheme fusing number, person, and mood.

Likewise, König states the following reasons for positing the subsecutive suffix as a grammaticalization of the Ik copulative case suffix (2002: 414):

1. Their forms are identical.
2. Their morphophonological behavior is identical.
3. Their identity is suggested by the similar identity between the participial suffix and the dative case suffix.
4. Their position in the word (always at the end) is identical.
5. The participial and subsecutive exhibit similar characteristics in terms of:
 - Their position in the word (as final or penultimate suffixes).
 - Their non-final/final forms.
 - Their form consisting of /k/ plus a silent vowel.
 - Their deletion of /a/ in the preceding stem.
 - Their exclusion from negated verb forms.

On the basis of these points of comparison, König makes a strong argument that Ik has grammaticalized case suffixes as verbal TAM suffixes. Following König's example, I would like to compare a few other morphemes in the language which show signs of possibly having arisen from some of the other nominal case suffixes. (3) below names these verbal morphemes (right column) and the case suffixes they may have come from:

(3)	Nominative	>	Realis
	Oblique	>	Irrealis
	Accusative	>	Present perfect
	Genitive	>	Singular imperative
	Ablative	>	Plural imperative

The first case > verbal suffix pair to be discussed is nominative > realis. It was Heine who noted that the nominative case and what he called the 'aorist' tense shared something in common: the vowel /a/, which he calls the 'thematic final a' (1983: §2.5.3.1.1). But neither he nor anyone else, to my knowledge, has suggested that the 'aorist' (what I call 'realis') marker /-a/ is a grammaticalization of the nominative case marker. The nominative case marker /-a/ and the realis marker /-a/ share one important characteristic: they both delete the final vowel of the stem or affix to which they attach.

To illustrate this, in (4), the first column shows the underlying forms, first of noun roots and the nominative suffix, and then of verb roots, pronominal subject markers, and the realis suffix. The second column then shows how /-a/ deletes the preceding vowel in both nouns and verbs.

(4) Nominative > Realis

ɲókí-a	→	ɲók ^{-a}	dog-NOM
sísí-a	→	sís ^{-a} ²¹	mead-NOM
bóré-a	→	bór ^{-a}	corral-NOM
ɲèkè-a	→	ɲèk ^{-a}	hunger-NOM
ɲorá-a	→	ɲor ^{-a}	cane.rat-NOM
ats-íí-a	→	ats-í ^{-a}	come-1SG-REAL
ats-ídi-a	→	ats-íd ^{-a}	come-2SG-REAL
ats-i-a	→	àts ^{-a}	come-3SG-REAL
ats-ísini-a	→	ats-ísin ^{-a}	come-1PL.INCL-REAL
ats-ími-a	→	ats-ím ^{-a}	come-1PL.EXCL-REAL
ats-ítí-a	→	ats-ít ^{-a}	come-2PL-REAL
ats-átí-a	→	ats-át ^{-a}	come-3PL-REAL

²¹ Devoiced vowels are often not audibly heard. They are all written here for morphological clarity.

Just from a morphological point of view, it seems unlikely that the shared vowel-deleting behavior should be a coincidence. And from a semantic point of view, just as the nominative case is the default ‘naming’ case for nouns, particularly the main arguments of clauses, so the realis is the default verb form in Ik. For these reasons, I propose that the realis is a grammaticalization of the nominative case.

A comparison of the nominative case suffix and the realis suffix yields these points:

1. Their forms are identical (-a).
2. Their morphophonological behavior is identical.
 - They both delete the final vowel of the preceding morpheme.
 - They both are devoiced/deleted before a pause.
 - They both are toneless but are assigned tone by the stem.
3. Their semantics are similar in that they both mark default forms.
4. Their possible relationship is suggested by the dative > participial and copulative > subsecutive grammaticalizations.

On the basis of these four points, the argument that the realis suffix is a grammaticalization of the nominative case suffix seems as strong as for the dative > subjunctive and copulative > subsecutive pairs.

Next, the oblique case suffix and the verbal irrealis suffix exhibit similarities that cannot be written off as chance given the other grammaticalizations already discussed.

The oblique case in Ik is zero-marked, allowing the basic, underlying form of the noun stem to surface. Since every nominal morpheme in Ik ends in a vowel, any noun stem in the oblique case ends in a vowel (which is devoiced before a pause, as in (5) below). Likewise, the irrealis in Ik is zero-marked. A verb conjugated in the irrealis mood allows the underlying form of the final suffix to surface. In (5), the final vowel of the pronominal subject-agreement markers is present, though devoiced before the pause:

(5) Oblique > Irrealis

ŋók ⁱ	dog[OBL]	ats-í ⁱ	come-1SG[IRR]
sís ^t	mead[OBL]	ats-íd ⁱ	come-2SG[IRR]
ḃór ^c	corral[OBL]	ats- ⁱ	come-3SG[IRR]
ɲèk ^e	hunger[OBL]	ats-ísín ⁱ	come-1PL.INCL[IRR]
ŋúr ^a	cane.rat[OBL]	ats-ím ⁱ	come-1PL.EXCL[IRR]
		ats-ít ⁱ	come-2PL[IRR]
		ats-át ⁱ	come-3PL[IRR]

Besides the morphophonological similarity, the oblique and irrealis also exhibit a semantic similarity. The oblique case has an ‘oblique’ function of marking left-over case relations not already handled by the other cases. Likewise, the irrealis mood has an ‘oblique’ function of predicating that which has not happened (as opposed to predicates encoded by the realis suffix). It seems noteworthy that the counterparts of the oblique and irrealis – the nominative case and the realis mood – are also identical in form and similar in semantics. This suggests that the grammaticalization of the nominative into the realis coincides somehow with that of the oblique into the irrealis. A comparison of the oblique case zero-suffix and the irrealis mood zero-suffix yields the following points:

1. Their forms are identical (∅).
2. Their morphophonological behavior is identical.
3. Their semantics are similar.
4. Their possible relationship is suggested by the dative > subjunctive, copulative > subsecutive, and nominative > realis grammaticalizations.

On the basis of these four points, the argument in favor of the irrealis being a grammaticalization of the oblique comes out reasonably strong.

The next pair to be discussed is the accusative > present perfect pair, as shown here:

(6) Accusative > Present perfect

ɲóki-k ^a	dog-ACC	ats-í-à-k ^a	come-1SG-REAL-PRF
sísí-k ^a	mead-ACC	ats-íd-à-k ^a	come-2SG-REAL-PRF
bóré-k ^a	corral-ACC	ats-á-k ^a	come-3SG-REAL-PRF
ɲèkè-k ^a	hunger-ACC	ats-ísin-à-k ^a	come-1PL.INCL-REAL-PRF
ɲórá-k ^a	cane.rat-ACC	ats-ím-à-k ^a	come-1PL.EXCL-REAL-PRF
		ats-ít-à-k ^a	come-2PL-REAL-PRF
		ats-át-à-k ^a	come-3PL-REAL-PRF

The data in (6) show that the accusative case suffix and the present perfect tense suffix both consist of *-ka*. Both are toneless, and both have their vowel devoiced before a pause. But one important difference is that while the accusative case suffix attaches to the final vowel of the noun stem, the present perfect suffix attaches to the realis suffix. Since, as I have argued above, the realis suffix is itself a grammaticalization of the nominative case, this would constitute the only known situation in Ik where two case suffixes (grammaticalized or not) cooccur in the same word. This could be an argument against the present perfect suffix being a grammaticalization of the accusative case suffix. So although the morphophonological similarities point toward possible grammaticalization, the semantic dissimilarities and the morpho-phonological anomaly of cooccurring case suffixes are a strike against it.

The following points of comparison can be made about the accusative case suffix and the present perfect suffix:

1. Their forms are identical (*-ka*).
2. Their morphophonological behavior is identical:
 - They both lose their /k/ in their non-final forms.
 - They are both toneless.
 - Their vowels are devoiced before a pause.
3. Their strictly morphological behavior is not identical:
 - The accusative affixes to the stem-final vowel, while the present perfect affixes to the realis suffix.
4. Their semantics are not similar.
5. Their possible relationship is suggested by other case > verbal suffix pairs.

On the basis of these five points, it is not conclusive whether this constitutes an instance of grammaticalization from the accusative case suffix to the present perfect tense suffix. Their morphological shape, morphophonological behavior, and existence of other case grammaticalizations indeed suggest a relation between them. But their semantic dissimilarities and their morphological behavior (with regard to concatenation) do not.

The next case-verbal suffix pair to be discussed is the genitive > singular imperative pair, as shown in (7):

(7) Genitive > Singular imperative

ɲóki- ^e	dog-GEN	ka- ^e	go-SG.IMP
sísí- ^e	mead-GEN	ats- ^e	come-SG.IMP
bóré- ^e	corral-GEN	ce- ^e	kill-SG.IMP
ɲèkè- ^e	hunger-GEN	ógò- ^e	leave-SG.IMP
ɲórá- ^e	cane.rat-GEN	tsuwa- ^e	run-SG.IMP

As shown in (7), the genitive case suffix and the singular imperative suffix both have the form of /-e/, making them segmentally identical. They are also both toneless. But one important behavioral difference between the two is that the genitive case marker attaches to the final vowel of the preceding stem, while the singular imperative suffix attaches directly to the verb root without the default intervening vowel /i, i/. This different ‘suffixation strategy’ is a strike against the argument that the

singular imperative suffix arose from the genitive case suffix. Also, semantically, there is nothing immediately obvious linking genitivity to singular imperativity.

A comparison between the genitive and singular imperative yields the following points:

1. Their forms are identical (-*ε*).
2. Their morphophonological behavior is only similar:
 - They both are devoiced/deleted before a pause.
 - They are both toneless.
 - Their suffixation strategy is different: the genitive attaches to the stem-final vowel, while the singular imperative deletes it.
3. Their semantics are not similar.
4. Their possible relationship is suggested by other case > verbal suffix pairs.

On the basis of these four points alone, the analysis of the genitive > singular imperative pair must remain inconclusive for the time being. In particular, their different suffixation strategies and dissimilar semantics decrease the likelihood of grammaticalization having taken place between them.

The final case > verbal suffix pair to be discussed is the ablative > plural imperative:

(8) Ablative > Plural imperative

ɲókú- ^o	dog-ABL	póru- ^o	go.ahead-PL.IMP
sísú- ^o	mead-ABL	rímú- ^o	take.shelter-PL.IMP

As (8) shows, both the ablative case suffix and the plural imperative suffix have the same form (-*ɔ* and its allomorphs). In (8), both the two noun roots and the two verb roots end in a high front vowel (*ɲókí*- ‘dog’, *sísí*- ‘mead’, *póri*- ‘go ahead’, and *rímí*- ‘take shelter’). As elsewhere in the grammar, back vowels tend to partially assimilate preceding high front vowels (*i/i* → *ɔ/u*) as they do in these four examples. Only noun roots ending in high front vowels are used here to facilitate the comparison.

So, segmentally, the ablative and plural imperative suffixes are identical. But they do differ with regard to tonal behavior. The ablative case marker is toneless, while the plural imperative marker has a high tone. If the plural imperative suffix is devoiced before a pause, the preceding vowel takes a high tone. This different tonal behavior is one point against this being an instance of grammaticalization. Also, semantically, there appears to be no obvious connection between ablativity and plural imperativity.

A comparison of the ablative and plural imperative suffixes yields the following points:

1. Their forms are segmentally identical (-*ɔ*).
2. Their non-tonal morphophonological behavior is identical.
 - They both partially assimilate a preceding high front vowel.
 - They both are devoiced before a pause.
3. Their tonal behavior is not identical.
 - The ablative suffix is toneless.
 - The plural imperative has a high tone.
4. Their semantics are not similar.
5. Their possible relationship is suggested by other case > verbal suffix pairs.

On the basis of these five comparative points, it is not conclusive whether the plural imperative suffix originated from the ablative case suffix. Until an explanation for the tonal differences and semantic dissimilarities can be offered, this instance of grammaticalization will have to remain unresolved.

In summary, in addition to the well-established dative > subjunctive and copulative > subsecutive grammaticalization pairs, five other possible pairs have been reviewed. Based on a number of

comparative points, two of these five pairs have strong arguments in favor of a grammaticalization hypothesis. The other three are rather weak:

(9) <u>Stronger</u>	<u>Weaker</u>
Nominative > realis	Accusative > present perfect
Oblique > irrealis	Genitive > singular imperative
	Ablative > plural imperative

4 Case suffixes grammaticalized as tense indicators on demonstratives, relative pronouns, and tense clitics

Ik case suffixes also seem to appear on demonstratives, relative pronouns and tense particles that have themselves been grammaticalized from demonstratives. In Ik, demonstratives and relative pronouns are identical in form and are marked for Ik's three-term past tense system: recent past, removed past, and remote past. These three tense levels are shown in (10) in addition to the present or 'gnomic' tense:

(10) Ik demonstratives/relative pronouns					
na	this/which	ni	these/which	(gnomic)	
nak ^a	that/which	nik ⁱ	those/which	(recent past)	
sin ^a	that/which	sin ⁱ	those/which	(removed past)	
nok ^o	that/which	nuk ^u	those/which	(remote past)	

As can be seen in (10), the gnomic form of the demonstrative/relative pronouns provides the base on which the other tense forms are built. For the singular and plural tensed forms, their respective parts can be dissected as follows:

(11) Singular	Plural	
na	ni	(gnomic)
na-k^a	ni-kⁱ	(recent past)
si-n^a	si-nⁱ	(removed past)
no-k^o	nu-k^u	(remote past)

In (11), the base (gnomic) forms of the demonstratives/relative pronouns are in bold print. For the recent past and remote past, the base form appears to have a suffix resembling a case suffix. Only the removed past shows a prefix instead. The recent past form of the singular demonstrative has *-ka* as a suffix. This is a formal parallel for the accusative case suffix and present perfect marker (see example 8 above). Likewise, in the remote past singular form, the demonstrative has *-ko* as a suffix. This is an obvious reflection of the copulative case and subsecutive mood markers (see example 2 above).

Despite the tantalizing similarities between these suffixes and case suffixes, there are several problems. For example, though the recent past singular form seems to include a grammaticalization of the accusative case, what about the plural form? And in the plural column, how did the copulative case marker (if that is what this is) change from (*-ko*) to (*-ku*)? One possible answer for the latter question comes from the non-final forms of these demonstratives, in which the segment /k/ is deleted. The non-final forms of the data in (12) are presented below:

(12) Singular	Plural	
na	ni	(gnomic)
na-a	ni-i	(recent past)
si-na	si-ni	(removed past)
no-o	nu-u	(remote past)

According to vowel assimilation behavior operating in other parts of the grammar, the vowel in the suffixes for the remote past forms (singular and plural) is likely assimilating the vowels in the base forms. Elsewhere in the grammar, the non-final copulative and ablative suffixes (both *-o*) totally assimilate preceding */a/*:

- | | | | | |
|------|-----------------|---|-----------------|--------------|
| (13) | àsàkâ- <i>o</i> | → | àsàkò- <i>ò</i> | door-COP/ABL |
| | awá- <i>o</i> | → | awó- <i>ó</i> | home-COP/ABL |

However, when the copulative case is grammaticalized as the subsecutive mood, this vowel assimilation does not always take place as expected. In the final forms, the subsecutive *-k^o* does not assimilate a preceding */a/*, and in the non-final forms, it is the subsecutive itself that is assimilated:

- | | | | | |
|------|--------------------------|---|-------------------------|----------------------------|
| (14) | ats-íi-a-k ^o | → | ats-í-à-k ^o | come-1SG-REAL-SUBS.FF |
| | | → | ats-í-à-à | come-1SG-REAL-SUBS.NF |
| | ats-ími-à-k ^o | → | ats-ím-à-k ^o | come-1PL.EXCL-REAL-SUBS.FF |
| | | → | ats-ím-à-à | come-1PL.EXCL-REAL-SUBS.NF |

But the vowel assimilation failures in (14) may have something to do with the irregularities of the subsecutive verb paradigm. If the remote past form of the singular demonstrative/relative pronoun really is an example of the grammaticalization of the copulative case, then it would suggest that the non-final form is older (the */k/* being added later, after the vowel assimilation became the norm). Such a situation would allow the */a/* in the base form */na/* to be assimilated by the copulative suffix to */o/*. Nowhere else do we see an */o/* assimilating an */a/* with an intervening consonant.

Turning back now to the plural remote past form, a slightly different problem arises. Elsewhere in the grammar, the copulative and ablative case suffixes readily partially assimilate a preceding */i/* to */u/*, for example as in:

- | | | | | |
|------|----------------|---|---------------------|----------------|
| (15) | ḡókí- <i>o</i> | → | ḡókú-k ^o | dog-COP.FF |
| | | → | ḡókú- ^o | dog-ABL.FF |
| | | → | ḡókú- <i>o</i> | dog-COP/ABL.NF |

So the copulative suffix *-ko* added to the base plural demonstrative *ni* could easily result in *nuk^o*. This follows observed morphophonological patterns. The question remains, however, as to why the final form of the grammaticalized copulative on the plural remote past demonstrative surfaces as *-k^u* instead of *-k^o*. Perhaps at some level of morphological development, the assimilated base form *nu-* proceeded to fully assimilate the suffix that had assimilated it in the first place.

Finally, if the recent past demonstrative possesses a grammaticalized form of the accusative case suffix, it is not immediately clear why the plural form is *-ki* instead of *-ka* as in the singular. Nowhere else in the grammar does a [+ATR] */i/* assimilate a preceding */a/* (as it would if *ni+a* → *nii*). We have already seen some problems in applying typical morphophonological principles to the demonstratives that hold true in other areas of the grammar (i.e. case markers and their verbal grammaticalizations). So perhaps the demonstratives retain fossilized morphophonology. Or, equally possible, the suffixes observed on the demonstratives did not arise from case markers at all (at least not along the same paths than with verbal suffixes).

Two Ik demonstratives were grammaticalized not only into relative pronouns but also into tense clitics. Along that process, the deictic function of a demonstrative relative to a single noun or a noun phrase was extended to whole clauses. Since clauses are singular, the singular forms of the demonstratives were borrowed to function as tense markers. The tense markers for recent past and remote past are identical to the demonstratives, while the removed past presents a different form of unknown origin:

Table 3: Ik tense clitics

	Non-final form	Final form
Recent past	naa	nak ^a
Removed past	bee	bats ^e
Remote past	noo	nok ^o

The discussion above examining the recent past and remote past singular demonstratives applies equally to the recent and remote past tense clitics. For the removed past tense clitic, the segment /ts/ is missing in the non-final form, leading to vowel assimilation. The origin of the suffix /-tse/ is unknown, though it reminds one of the Proto-Gumuz *tsa ‘body’ that has been grammaticalized into a variety of uses (Ahland 2012: 251).

The parsed versions of the tense clitics in Table 3 are as follows:

Table 4: Ik tensed clitics parsed

	Non-final form	Final form
Recent past	na-a	na-k ^a
Removed past	be-e	ba-ts ^e ²²
Remote past	no-o	no-k ^o

Table 4 shows that just as with the demonstratives and relative pronouns, the suffixes on Ik tense clitics lose their consonantal segment in non-final forms. The suffix on the recent past form closely resembles the accusative case and present perfect tense suffixes, and the suffix on the remote past form closely resembles the copulative case and subsecutive mood suffixes.

In the preceding discussion we have seen how suffixes found on demonstratives, relative pronouns, and tense clitics hint at being grammaticalizations of nominal case suffixes. The situation is less clear-cut than with the verbal suffixes. But still, given the possibility of grammaticalization, one is forced to consider the possibility. Below are points of comparison between case markers and the suffixes found on these deictics:

1. Their forms are identical in the singular, similar in the plural.
2. Their morphophonological behavior is identical in some examples, but only similar in others.
3. Their semantics are linked, but only through the verbal suffixes that are grammaticalized from case markers: a) the present perfect and recent past could easily be linked, and b) the subsecutive mood and the remote past could be linked, especially if the subsecutive used to be a true ‘narrative’.

In summary, the extent of the grammaticalization of case suffixes in the verbal domain has led me to view morphemes in other grammatical domains through the lens of ‘case’. And as it turns out, there are in fact a number of suffixes among demonstratives, relative pronouns, and tense clitics that closely resembles case suffixes.

5 Conclusion

In this paper, we first looked at the Ik language’s nominal case-marking suffixes and their allomorphs. Then it was shown how, based on König’s (2002) arguments, the subjunctive suffix originated from the dative case and the subsecutive suffix from the copulative case. With König’s analysis as a model,

²² The base form for the removed past (*ba/-be-*) may be related to the Teso-Turkana non-present negation marker (*pa/-pe-*; see Dimmendaal 1983: 453).

we then looked at five other potential instances of a case suffix being grammaticalized into a verbal suffix. Two of these instances (nominative > realis and oblique > irrealis) seem as strong as the two proposed by König. Finally, we examined suffixes on demonstratives, relative pronouns, and tense clitics that closely resemble the accusative and copulative case suffixes.

The goal of this paper has been to show that once it was established that Ik grammar borrowed case suffixes for the subjunctive and subsecutive suffixes, the door was opened to consider other case grammaticalizations. As such, ‘case’ as a broad, meta-categorical concept, has become a heuristic for exploring previously undescribed areas of the grammar. Since ‘case’ suffixes are found in so many word classes (nouns, verbs, demonstratives, adverbs, etc.), their path into those word classes needs to be considered.

In closing, this paper proposes that the realis mood suffix in Ik is a grammaticalization of the nominative case and the irrealis mood suffix a grammaticalization of the oblique case. Beyond that, the other proposals are made with more hesitation since they encounter a number of problems as explained above. Nevertheless, there seems to be a connection between the copulative case suffix and markers for the remote past tense in demonstratives, relative pronouns, and tense clitics. If this is true, it gives more evidence for Ik being a language in which nominal case suffixes have been widely exploited source for morphemes in other grammatical categories. Or alternatively, all these homophonous morphemes called ‘case’ have evolved independently from some other morphemes now lost in the sands of linguistic time.

Abbreviations

ABL	ablative	NF	non-final
ACC	accusative	NOM	nominative
COP	copulative	OBL	oblique
DAT	dative	PL	plural
EXCL	exclusive	PRF	perfective
FF	final form	REAL	realis
GEN	genitive	SG	singular
IMP	imperfective	SUBJ	subjunctive
INCL	inclusive	SUBS	subsecutive
INS	instrumental	TAM	tense-aspect-mood
IRR	irrealis	VSO	verb-subject-object

References

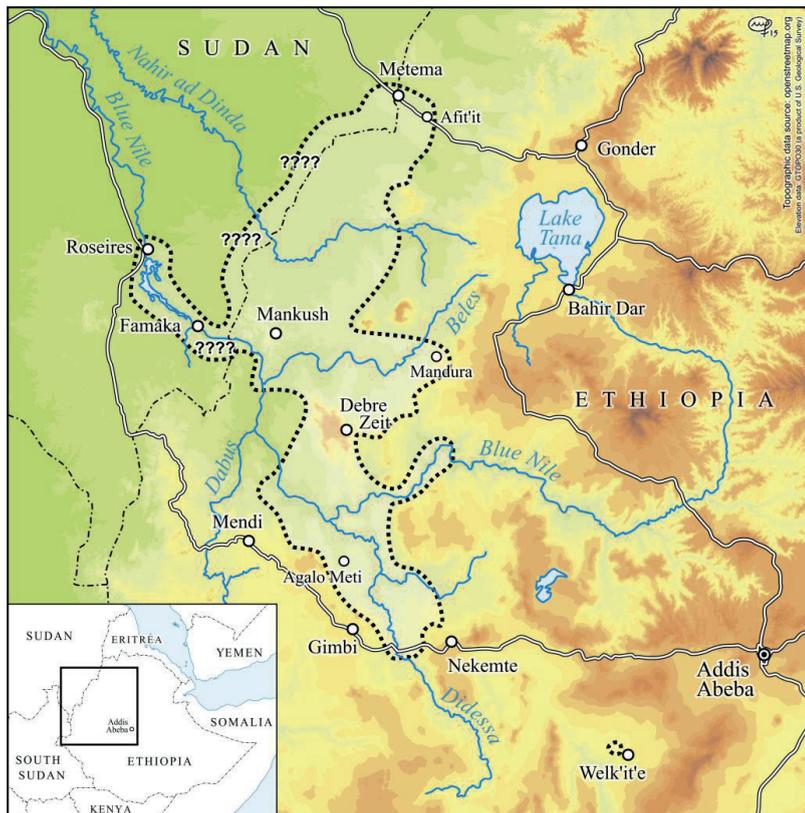
- Ahland, Colleen. 2012. A grammar of Northern and Southern Gumuz. PhD dissertation, University of Oregon.
- Crazzolara, J.P. 1967. General Sketch: Grammar and Vocabulary of the Language of the IK (Alias Teuzo). Typescript, Kang'ole, Uganda.
- Dimmendaal, Gerrit Jan. 1983. *The Turkana Language*. (Publications in African Languages and Linguistics, 2). Dordrecht & Cinnaminson: Foris Publications.
- Heine, Bernd. 1983. The Ik language. Typescript. Cologne: University of Cologne.
- König, Christa. 2002. *Kasus im Ik*. (Nilo-Saharan. Linguistics Analyses and Documentation, 17). Cologne: Köppe.
- König, Christa. 2008. *Case in Africa*. Oxford University Press.
- Schrock, Terrill. 2012. Unlocking the Ik Instrumental Case. Paper presented at the 42nd Colloquium on African Languages and Linguistics in Leiden, The Netherlands.
- Serzisko, Fritz. 1992. *Sprechhandlungen und Pausen: Diskursorientierte Sprachbeschreibung am Beispiel des Ik*. (Linguistische Arbeiten, 282.) Tübingen: Niemeyer.
- Tucker, A.N. 1971–73. Notes on Ik. *African Studies*, 30:341–54; 31:183–201; 32:33–48.

CLASSIFIER, REFLEXIVE, AND BEYOND: A SYNCHRONIC AND DIACHRONIC EXPLORATION OF THE MORPHEME /TSA/ ‘BODY’ IN GUMUZ

Colleen Ahland

1 Introduction

Gumuz is a Nilo-Saharan (NS) language/dialect cluster spoken in the river valleys of western Ethiopia and southeastern Sudan (Republic of Sudan). It is considered an isolate within the NS family (Bender 1997) but is likely most closely related to the Koman languages (Ahland 2013). This article focuses on two varieties of Gumuz spoken in Ethiopia: Northern Gumuz (NoG) and Southern Gumuz (SoG). In general, NoG is spoken north of Debre Zeit and Mankush and SoG is spoken to the south of these areas (see Map 1). Other languages spoken in the area are Berta (NS), Boro (Omoti), Oromo (Cushitic) and Awgni (Cushitic) (Ahland 2004).



Map 1: Gumuz location

Gumuz verbs are highly agglutinative, many of which have incorporated nouns (INs). The majority of these INs are body part terms which are either part of a lexicalized compound, an external possession (EP) construction, or have grammaticalized as verbal classifiers (CLs) (Ahland 2010, 2012). In this paper I explore the various functions and a possible historical source of the IN/CL */tsa/* in Gumuz which is the only incorporated body part term in Gumuz that is bound and has no corresponding free form. In Section 2, I provide an overview of Gumuz verbal morphology and discuss the various functions of */(á)ts(a)/* as part of the verb stem. In Section 3, I briefly describe the occurrence of */tsa/* in nominal compounds. In Section 4, I discuss the historical source of */tsa/* looking both at evidence internal to Gumuz as well as external comparative evidence found in the Koman languages. Lastly, in Section 5, I summarize my conclusions regarding the multiple functions of */tsa/* and its likely historical source as an old term for ‘body’ in Gumuz which has undergone lexical replacement.

2 Functions of */(á)ts(a)/* as part of the verb stem

Gumuz verb stems are comprised of a single verb root or a verb root plus an IN/CL as part of a complex stem. These verbs stems are highly agglutinative with up to fourteen position classes in NoG (Figure 1) and thirteen in SoG (Figure 2) for non-future verbs.¹ The complex verb stems are discontinuous, allowing inflectional (and derivational) morphemes to intervene.

The morpheme */tsa/* is one of many body part terms that can fill the IN/CL position class and this morpheme serves several functions. It can 1) form a verb-noun lexical compound, 2) participate in an external possession construction, 3) can serve as a verbal classifier which can also function as part of a reflexive construction and 4) can index complement clauses such as perception-cognition-utterance (PCU) complements.

Figure 1: Position class chart for NoG non-future verbs

Mood	Person (S / A)	Person (O or object of preposition)	Uncertainty	Instrumental & Dative; (-5) for INTR	Reciprocal	Pluractional	Main Verb Root	Greater Plural	Middle Voice	Directional	Incorporated Noun / Classifier	Perfect	Locative
(-7)	(-6)	(-5)	(-4)	(-3)	(-2)	(-1)	(0)	(+1)	(+2)	(+3)	(+4)	(+5)	(+6)

¹ Gumuz follows a different position class chart for future verb stems (cf. Ahland 2012).

Figure 2: Position class chart for SoG non-future verbs

Mood	Remote Past	Person (S / A)	Instrumental & Dative	Uncertainty	Reciprocal	Pluractional	Main Verb Root	Nonfuture Tense	Middle Voice	Directional	Incorporated Noun / Classifier	Locative / Past Perfect
(-7)	(-6)	(-5)	(-4)	(-3)	(-2)	(-1)	(0)	(+1)	(+2)	(+3)	(+4)	(+5)

2.1 Verb-noun lexical compounds

Gumuz has a limited number of lexicalized verb-noun compounds in which the IN is distinct from the IN of an external possession construction (Section 2.2) or a verbal classifier construction (Section 2.3). Of these, there are only a few lexical verb-noun compounds in Gumuz involving /*tsa*/ ‘body’. Unlike Mithun’s Type I noun incorporation (1984, 1986), the IN in these verb-noun compounds does not decrease the valence of the verb; rather, the valence remains the same (1a-b, 2a-b). The only instance in which valence is decreased with the addition of a body part term like /-(*á*)*ts(a)*/ is in a reflexive construction (see Section 2.3.2). Some of these lexical verb-noun compounds have compositional semantics in which the meaning of the verb-noun pair more or less equals the sum of its parts; e.g. ‘rise’ + ‘body’ = ‘get up, stand up’ (1b).² Other lexical compounds exhibit more idiosyncratic meanings. For example, the verb root /*gam*/ ‘know, find’ (2a) combines with /-(*á*)*ts(a)*/ ‘body’ to form the complex verb stem ‘see’ (2b).

- | | | |
|-----|---|---|
| | NoG | |
| (1) | a. fád
rise
‘Rise!’ | b. fád-áts
rise-body
‘Get up!’ |
| (2) | a. áxó saŋa d-a-gam
3SG salt AFF-3SG.TR-find
‘He found salt.’ | b. á-gam-átsa=ŋgó óminza
2SG.TR-find-body=NEG leopard
‘You didn’t see the leopard.’ |

2.2 External possession constructions

Arguably, the majority of incorporated nouns in Gumuz form part of an external possession (EP) construction. EP constructions involve an additional argument of the clause which serves as a possessor. With noun incorporation, the external S or O argument serves as possessor and the IN as the possessed entity (Payne and Barshi 1999). This is often referred to as “possessor raising” or “manipulation of case” as with Mithun’s Type II noun incorporation (1984). For example in (3),

² Example (1b) is likely related to the intransitive reflexive construction discussed in 2.3.2. However, it differs in that the verb root ‘rise’ cannot have an additional argument when /-(*á*)*ts(a)*/ ‘body’ is incorporated. That is, there exists no transitive reflexive counterpart as with other complex verb stems in reflexive constructions. The construct in (1b), however, could be an example of how /-(*á*)*ts(a)*/ ‘body’ has grammaticalized as a reflexive marker (in certain constructs) whereas other incorporated body part terms have not.

classifier /-(á)ts(a)/ ‘body’ classifies the human body, entities associated with the human body or those that are functionally or physically body-like (objects longer than they are wide). Examples (7-8) demonstrate /-(á)ts(a)/ ‘body’ as a classifier and examples (9-10) demonstrate other body part term classifiers that contrast with /-(á)ts(a)/. For instance, /-iíl(á)/ ‘belly’ classifies entities that have a concave surface or encompass a large area (9) and /-(á)cá/ ‘eye’ classifies clothing (10) among other entities (cf. Ahland 2012: 270).

NoG

- (7) ára d-ár-ʔáf-áts eʔâ-m
 1SG AFF-1SG.TR-wash-CL:body hand-1SG.POSS
 ‘I washed my hand(s).’
- (8) ʔaf-áts ɓagá-máts’á
 wash-CL:body body-house
 ‘Wash the wall(s)!’
- (9) ʔaf-iíl sánéá
 wash-CL:belly plate
 ‘Wash the plate(s)!’
- (10) aɣwa ʔaf-ác
 clothes wash-CL:eye
 ‘Wash the clothes!’

2.3.2 Reflexive constructions in Gumuz

In Gumuz, one can express a reflexive by means of a transitive or intransitive construction. The transitive reflexive construction involves a possessed body part term as the O argument or simply a possessed form of the word /baga/ ‘body’ in NoG (11) or /boga/ ‘self’ in SoG (13)⁴. The main difference between a more general transitive construction (12, 14) and the transitive reflexive construction (11, 13) is that in the latter, the possessor of the O argument is the same referent as the A argument. In both transitive constructions, the classifier ‘body’ is used if the O argument is a body part term or is body-like (7-8, 11-14), and the bound subject pronominal exhibits a transitive tonal pattern.

NoG

- (11) ára d-ára-ʔaf-áts ɓagá-ma
 1SG AFF-1SG.TR-wash-CL:body body-1SG.POSS
 ‘I washed my body (myself).’
- (12) ára d-ára-ʔaf-áts tʃogwá-dua
 1SG AFF-1SG.TR-wash-CL:body foot-child
 ‘I washed the child’s foot/feet.’

SoG

- (13) k’óá b-a-t’ar-ag-áts bogo-má
 dog AFF-3SG.TR-lick-NFUT-CL:body self-3SG.POSS
 ‘The dog licks himself.’
- (14) k’óá b-a-t’ar-ag-áts ɓaga
 dog AFF-3SG.TR-lick-NFUT-CL:body person
 ‘The dog licks the person.’

⁴ In SoG, the historical source for ‘self’ /boga/ is the lexeme /baga/ ‘body, person’.

Gumuz also has an intransitive reflexive construction. This construction involves an incorporated body part term as part of a complex verb stem and an intransitive tonal melody on the bound subject pronominal. The classifier /-(á)ts(a)/ ‘body’ can be found in both the transitive and intransitive reflexive constructions. In most instances in which the free NP lexemes for ‘body’ or ‘self’ are used in a transitive reflexive construction, one can simply eliminate this O argument and use an intransitive tonal melody to create an intransitive reflexive construct (15). Alternatively, it is also possible to have any incorporated body part term in a complex verb stem with an intransitive verbal melody (16). The transitive counterpart of this more general reflexive construction is nearly always an EP construct (17).⁵

NoG

- (15) ára d-ára-ʔaf-áts
 1SG AFF-1SG.INTR-wash-CL:body
 ‘I washed my body (myself).’
- (16) d-ára-ʔaf-átʃogw
 AFF-1SG.INTR-wash-foot
 ‘I washed my feet.’ (lit: ‘I foot-washed.’)
- (17) d-ára-ʃʔaf-átʃogw dua
 AFF-1SG.TR-wash-foot child
 ‘I washed the child’s feet.’ (lit: ‘I foot-washed the child.’)

Therefore, while /-(á)ts(a)/ ‘body’ serves a special function within reflexive constructions, it is not a required element for reflexives, and does not function as a reflexive marker per se. Rather, /-(á)ts(a)/ ‘body’ maintains its designated function as verbal classifier, and as people and body part terms are classified as “bodies” in Gumuz, this verbal classifier regularly participates in reflexive constructions.

2.4 Valence increaser and the indexing of complement clauses

Beyond verbal classifier, the IN/CL /-(á)ts(a)/ ‘body’ has further grammaticalized as a valence increaser which can index certain types of clausal complements, PCU complements in particular. This valence-increasing construction with /-(á)ts(a)/ is similar to that of a (transitive) EP or verbal classifier construction in that an O argument or clausal complement is required. The main difference is that in EP or verbal classifier constructions the verb root is transitive or labile whereas with valence-increasing constructions, the verb root is intransitive. This construction is most commonly used for adding and indexing PCU complements. For example, the PCU verb roots /ŋgaf/ ‘speak, talk’ and /sáánz/ ‘think’ are intransitive (18, 20). When /-(á)ts(a)/ is added, the resulting complex verb stems take a non-finite clausal complement with /ŋgaf-áts/ meaning ‘say (that), tell’ (19) and /sáánz-áts/ meaning ‘think about’ (21).

NoG

- (18) baab-éa zialá d-ée-ŋgaf-á
 father-1SG.POSS now AFF-FUT-speak-3SG.INTR
 ‘My father will speak now.’

NoG

- (19) norága ma-kód-a-má d-a-ŋgaf-áts
 book NMLZ-buy-NM-IP:O AFF-3SG.TR-speak-BODY
 ‘He said that he bought a book.’

⁵ Also note that there exists a corresponding internal possession construction (possessor and possessed expressed within a noun phrase) given in (12). The IN/CL /-(á)ts(a)/ ‘body’ in this construct is not part of an EP construction.

SoG

- (20) b-á-sáanz-agá
AFF-3SG.INTR-think-NFUT
'He thought.'

SoG

- (21) b-a-sáanz-agá-ts ma-tsá
AFF-3SG.TR-think-NFUT-BODY NMLZ-go
'He thought about going.'

It is also possible for verbs in this construction to take non-clausal complements. In (22), for example, when the intransitive verb root 'sing' forms part of a valence-increasing construction with */(á)ts(a)/*, the verb stem takes the O argument 'game' qualified by the relative clause 'which they play when their relatives go to the fields during the day'.

SoG

- (22) [ga-así-iida et-íi-asíi-ga
NMLZ2-play-children REL-3PL.TR-play-NFUT
nó↓ká=b-ii-tsá-gá dá-gwa-é↓bá-máám
TEMP=AFF-3PL.INTR-go-NFUT PRO.PL-PLACE-home-3PL.POSS
kiílámíts'a níilóka ?á] káma-e?é-íla-ts ká=áca
to.field in.day MED FUT-sing-1PL.EXCL.TR-BODY BEN=2PL
'We will sing for you all a game (children's play) which they play when their relatives go to the fields during the day.'

For transitive/labile PCU verb roots, */(á)ts(a)/* is not added for complement clauses. In (23), the transitive verb root */gay/* 'know, find' (NoG: */gam/*) does not require the addition of */(á)ts(a)/* in order to express a complement clause or an O argument for that matter (see example 2a).⁶

SoG

- (23) mazá ?á ma-u-é-á alá-má b-ár-gay-gá
guy MED NMLZ-go-TWRD-NM GEN-3SG.POSS AFF-1SG.TR-know-NFUT
'I knew that the guy would come.' (lit: 'I knew the guy's coming.')

The valence-increaser */(á)ts(a)/* can also be used to index other types of non-finite verbal complements. For example, the intransitive verb root */kól/* means 'return' (24) but when */(á)ts(a)/* is added to the verb stem, a non-finite clausal complement must be added and the meaning changes to 'do again' (25).

SoG

- (24) b-ár-↓kól-agá
AFF-1SG-return-NFUT
'I returned.'
- (25) b-íi-kól-agá-ts ma-sá-ŋga
AFF-3PL.TR-return-NFUT-BODY NMLZ-eat-food
'They ate again.'

⁶ In fact, if */(á)ts(a)/* 'body' were added to the verb root */gay/* 'know' (NoG: */gam/*), the resulting verb stem would mean 'see' (see example 2b). In any case, it is clear that */(á)ts(a)/* 'body' functions as a valence increaser of sorts for intransitive verb roots and not merely a suffix that indexes certain types of complement clauses.

Another incorporated body part term /-(á)k'w(á)/ 'head' can also serve as a valence increaser on intransitive verbal roots. However, this IN often functions as a causative⁷ and typically cross-references an O argument which is not a nominalized verb (i.e. a non-derived noun). For example, in (25) the verb stem with the IN /-(á)ts(a)/ takes the nominalized (infinitival) verbal complement 'to eat food' whereas in (26) the verb stem with the IN /-(á)k'w(á)/ (which has the same intransitive verb root /kól/) takes the O argument 'book'.

SoG

- (26) kól-ak'w norága
 return-head book
 'Return the book.'

3 The morpheme /-tsá-/ found in nouns

The morpheme /tsa/ 'body' also appears in nominal compounds. Its occurrence on nouns parallels to some degree the distribution of class morphemes within class morpheme compounds in Gumuz (cf. Ahland 2012). The Gumuz class morphemes are classifying morphemes that form compounds with certain nouns, often functioning in a hypernym-hyponym relationship. The set of class morphemes in Gumuz is similar to the set of verbal classifiers with similar categorization: /k'wá-/ 'head', /cá-/ 'eye', /k'ós-/ 'tooth', /ts'è-/ 'ear', and /lil-/ 'belly'. The morpheme /-tsá-/ very rarely occurs in this sort of nominal compound and when it does, it is often preceded and followed by another noun. Class morpheme compounds, on the other hand, are typically comprised of two nouns – the class morpheme and the noun it classifies.⁸ In (27), /k'wá-/ 'head' classifies containers and is therefore used in a compound with 'gourd'. The morpheme /tsa/ appears to form a part of a class morpheme compound in (28) and (29), always co-occurring with the otherwise free noun form /baga/ 'body' and its verbal classifier counterpart /-(á)ts(a)/. This type of compound with /-tsá-/ is only known to occur in NoG.

NoG

- (27) k'wá-↓báχágá
 head-(specific type of)gourd
 'gourd for drinking'

NoG

- (28) χókwa-ka-ts bágá-tsá-↓máts'á ka-tjéraka
 clean-INSTR-body body-body-house INSTR-cloth
 'Clean the wall with a cloth.'

- (29) wíd-áts maχabá-ts-bágá-ma
 see-CL:body wound-body-body-1SG.POSS
 'Look at my (body) wound.'

4 Historical source of /tsa/

The morpheme /tsa/ can be internally reconstructed as a noun-like morpheme meaning 'body' based on constructions like example (30) below (Ahland 2010: 165-66, 2012: 251). In (30), several body part terms can be used in a prepositional construction denoting 'by oneself'. However, /tsa/ is the only body part term used in this construction that cannot be uttered on its own (31).

⁷ Other examples of /-(á)k'w(á)/ as a causative include but are not limited to the following stems formed with intransitive verb roots: /jâ-k'w/ (die-head) 'kill', /bâts'-ák'w/ (hide-head) 'hide something', /bit-ák'w/ (descend-head) 'lower something'.

⁸ There also exist endocentric class morpheme compounds in Gumuz in which the dependent noun does not appear to be classified by the class morpheme. Rather, the whole compound belongs to the class.

- SoG**
- (30) a. dua b-a-damb-agá-ts ma-n-tsa ka-**tsá**-má
 child AFF-3SG.TR-try-NFUT-body NMLZ-PL-go INSTR-body-3SG.POSS
 b. “ “ “ ka-**bii**-má
 INSTR-neck-3SG.POSS
 c. “ “ “ ka-**cá**-má
 INSTR-eye-3SG.POSS
 ‘The child tried to crawl by himself/herself.’ (lit: ‘...with his/her body/neck/eye’)
- (31) a. *tsá-má b. bii-má c. (il)cá-má
 body-3SG.POSS neck-3SG.POSS eye-3SG.POSS
 ‘his/her body’ ‘his/her neck’ ‘his/her eye’

As mentioned previously, the free noun form for ‘body’ in Gumuz is /baga/ which is also the lexeme for ‘person/people’ (see also examples 11 and 14). The polysemy of this lexeme is so commonplace in Gumuz that the compound for ‘human body’ uses /baga/ as both the head and dependent noun: first as the head noun meaning ‘body’ and second as the dependent noun meaning ‘person’ (32).

NoG and SoG

- (32) bagá-baga
 body-person
 ‘person’s body’ or ‘human body’

Considering the polysemy discussed above and the comparative evidence to be presented below, I propose that the bound morpheme /tsa/ ‘body’ was once a free noun form in Gumuz that existed alongside the free noun form /baga/ ‘person/people’. Over time, *tsa was incorporated into the verb stem and became phonologically reduced. As *tsa became more grammaticalized and more phonologically reduced, Gumuz speakers began to use /baga/ ‘person/people’ to refer to ‘body’ in place of *tsa. In Koman, the languages most closely related to Gumuz, apparent cognates of both noun forms co-exist, namely in Komo and Uduk (Table 1). These likely co-existed as free noun forms in Gumuz as well.

Table 1: Lexemes for ‘body’, ‘people’, and ‘person’ in Gumuz and Koman

	Koman Languages				Gumuz
	Gwama	Komo	Opo	Uduk	Gumuz (all)
	(Kievet & Robertson 2011)	(Manuel Otero, p.c.)	(Lemi 2010)	(Don Killian, p.c.)	
‘body/self’	yís ‘body’	ɪʃ ‘body’ (-ɪʃ) REFL	e:s ‘body’	īs ‘self, body’	-ts, -tsa -tsá-
‘body/ people’	-	giba ‘people’	-	bùngwār ‘body’	baga ‘body’, people/person’
‘person’		yiba ‘person’			baaha ‘person’

I claim that these “cognates” from Table 1, namely those for /tsa/, are “apparent” due to the lack of regular sound correspondences. However, there is a somewhat regular sound correspondence between Koman languages and Gumuz which suggests that the correspondence in Gumuz should be /s/ instead

of /ts/ (Table 2).⁹ The protoforms for all of these lexemes appear to have the proto-segment (most likely *s) in non-word-final position. It is therefore possible that *s became /ts/ word-finally in Gumuz and the earlier proto-form in Gumuz may have been *ats or *its.

Table 2: s:f Regular sound correspondence in Koman plus Gumuz

	‘eat’	‘stomach, intestine’	‘tooth’	‘god, sky, up’	‘body’
Gwama	ʃa	bùʃiʔ	ʃiʔ	wúús	yís
Komo	ʃá	bùʃ	ʃē	míʃ	íʃ
Opo	saa	-	see	wusi	e:s
Uduk	ʃwá	bùʃ	ʃē	mís	īs
Gumuz	sá	boosa	k’ó-sa	mísá (NoG) músá (SoG)	*its

5 Conclusion

The bound morpheme /tsa/ is found in both verbs and nouns in Gumuz. In verbs, the IN /-(á)ts(a)/ ‘body’ can form part of a lexicalized verbal compound, an EP construction, a verbal classifier construction, and a valence-increasing construction in which nominalized complements (most typically PCU complements) can be added to an intransitive verb root. The IN/CL /-(á)ts(a)/ ‘body’ can also be found in reflexive constructions but its designated function in such constructions is that of verbal classifier.

This bound morpheme /tsa/ is a relic of a free form that once co-existed with the synchronic free form for ‘people’ /baga/. Over time, /baga/ eventually replaced *tsa as the free noun form for ‘body’. Both internal evidence from Gumuz and external (comparative) evidence from the Koman languages suggest that /tsa/ ‘body’ in Gumuz is very old and not likely borrowed.

Abbreviations

1PL	first person plural	NEG	negative
1SG	first person singular	NFUT	nonfuture tense
2PL	second person plural	NMLZ	nominalizer, verbal noun
A	most agent-like argument of a verb	NMLZ2	(product) nominalizer, derives nominals which retain no verbal arguments
AFF	affirmative mood	NoG	Northern Gumuz
BEN	benefactive	O	object of a verb
CL	classifier, either verbal or the head of a class morpheme compound	REFL	reflexive
EP	external possession	REL	relativizer
EXCL	exclusive	S	single argument of a verb
FUT	future tense	SoG	Southern Gumuz
IN	incorporated noun	TR	transitive
INSTR	instrumental	TWRD	action directed towards speaker, also used for an action taking place in a different location from the speaker
INTR	intransitive		
MED	medial demonstrative		

⁹ The sources for Gwama, Komo, and Opo in Table 2 are the same as those in Table 1. For Uduk, I consulted Beam and Cridland (1979).

References

- Ahland, Colleen. 2004. Linguistic variation in Gumuz: A study of the relationship between historical change and intelligibility. Master's thesis, University of Texas at Arlington.
- Ahland, Colleen. 2010. Noun incorporation and predicate classifiers in Gumuz. *Journal of African Languages and Linguistics* 31: 159–203.
- Ahland, Colleen. 2012. A grammar of Northern and Southern Gumuz. PhD dissertation, University of Oregon.
- Ahland, Colleen. 2013. The status of Gumuz as a language isolate (LSA Meeting Extended Abstracts 2013). <http://elanguage.net/journals/lsameeting/issue/view/374> (25.09.2014).
- Beam, Mary S. & Elizabeth A. Cridland. 1979. *Uduk-English Dictionary*. (Linguistics Monograph Series, 4). Khartoum: Sudan Research Unit, University of Khartoum.
- Bender, M. Lionel. 1997. *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom.
- Grinevald, Colette. 2000. A morphosyntactic typology of classifiers. In Gunter Senft (ed.), *Systems of Nominal Classification*, pp. 50–92. New York: Cambridge University Press.
- Heine, Bernd. 2011. Areas of grammaticalization and geographical typology. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical Typology and Linguistic Areas*, pp. 41–66. Amsterdam/Philadelphia: Benjamins.
- Kievit, Dirk & Erika Robertson. 2012. Notes on Gwama grammar. *Studies in African Linguistics* 41.1: 39–59.
- Kebebew, Lemi. 2010. A grammatical description of Opo. Master's thesis, Addis Ababa University.
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60: 847–94.
- Mithun, Marianne. 1986. The convergence of noun classification systems. In Colette Craig (ed.), *Noun Classes and Categorization*, pp. 379–397. Amsterdam/Philadelphia: Benjamins.
- Payne, Doris L. & Immanuel Barshi. 1999. Introduction. In Doris L. Payne & Immanuel Barshi (eds.), *External Possession: What, How, Where and Why?*, pp. 3–29. Amsterdam/Philadelphia: Benjamins.
- Schladt, Mathias. 2000. The typology and grammaticalization of reflexives. In Zygmunt Frajzyngier (ed.), *Reflexives: Forms and Functions*, pp. 103–124. Amsterdam/Philadelphia: Benjamins.
- Tucker, Archibald N. & Margaret A. Bryan. 1966. *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press.

LOCATIONAL NOUNS IN DATOOGA¹ (SOUTHERN NILOTIC, TANZANIA)

Stefan Bruckhaus

1 Introduction

Body parts serving as a structural template for the conceptualization of spatial relations is attested in many languages of the world and even assumed to be a universal cognitive strategy (Heine 1997: 143). The argument put forward in this article is that the human body is a very significant source object for the expression of spatial relations in Datooga. This becomes evident because in the greater context, i.e. with different animate and inanimate centers of orientation, most expressions for spatial relations obviously derive from expressions for parts of the human body. The major objective of this article is to present a structured overview of Datooga's locational nouns and to provide analyses regarding their origins as well as their context dependencies, in order to fill in partially the descriptive gap for Datooga in the study of the spatial grammars of Nilotic languages (e.g. Andersen 2012; von Heyking & Storch 2007; Mietzner 2012, 2009, 2007; Payne 2013). The second and minor aim is to identify the stage of grammaticalization of locational nouns, which, according to Heine (1997, 1989), could lose nominal properties and develop into adverbs or prepositions. Therefore locational nouns are compared against their potential target categories and instances which may speak for a categorial change will be discussed.

In other recent works (Heine 2011, Kießling et al. 2008), the term “relational nouns” is used to associate spatial relator nouns with the conceptual category relational² which, according to Sapir (1921), comprises also other non-spatial concepts.³ By the term locational nouns I refer to a class of polysemic nouns in Datooga which typically have the following three characteristics: First, they serve as spatial relators to localize objects in a geometric relation to one another or to add a locational specification to the predication. Thus they fulfill the task which is accomplished by spatial prepositions and adverbs in some other languages. A second feature of locational nouns in Datooga is that, in most cases, they are homophoneous with expressions for body parts. The following two examples illustrate a quasi-prepositional (1) and a quasi-adverbial (2) usage of two body part nouns.

- | | | | | |
|-----|--------------|---------------|-----------|------------------------------------|
| (1) | gáwây | daqátá | qeeda | ‘They went in front of the house.’ |
| | ga-way | dáqád-da-’ | qee-da | |
| | S3-go.PL | chest-SPC-AS | house-SPC | |
| (2) | qabadáadá | dngáanêeda | | ‘He followed (them) sideways.’ |
| | qa-badaad-a | dngáan-ée-da | | |
| | S3-follow-IS | cheek-NUM-SPC | | |

¹ The Southern Nilotic language cluster Datooga is spoken by 87.800 individuals (Lewis et al., citing a 2000 census) in northern and central Tanzania. Along with extinct Omotic they constitute one branch of the Southern Nilotic languages next to the Kanlenjin languages which form the other branch. The basic tenets of the phonology and the morphology of Datooga are sketched in Rottland (1982). This study is based on narrative texts of the Gisamjanga variety of Datooga collected by Paul Berger in 1935–36 which have been converted into electronic format by Roland Kießling. In addition I have conducted fieldwork on this topic in Tanzania in June/July 2013. Examples are taken from the Berger corpus (BC), if not indicated otherwise. The phonetic transcription follows the IPA, except for *j* which represents the voiced palatal plosive [j] and *y* which stands for the palatal approximant [j]. For their contributions to this article, I am indebted to Rev. Festo Basso (Barabaiga), Herman Malleyeck (Gisamjanga), Rev. Josephat Masaka (Rotigenga) and Rev. Emmanuel Dudiyeck (Gisamjanga). I express my gratitude to my supervisor Roland Kießling for reviewing an earlier draft of this paper.

² Note that there exists a different notion of “relational” versus “free” nouns by Welmers (1973: 212), which is based on morphosyntactic distinctions in Mande languages.

³ Sapir (ibid.: 9) himself makes a threefold distinction of concrete, abstract and purely relational (e.g. ‘of’, ‘by’ or ‘and’) concepts.

In the first example the noun *daqátá* ‘chest’ is used to express the concept ‘in front of’ and in (2) the noun *dugáanéeda* ‘cheek’ conveys the meaning ‘sideways’. Both sentences feature peripheral locational arguments of a predicate which is filled by a motion verb. The most important difference between the adverbial and the prepositional usage is that in the former the locational noun stands alone and makes a locational specification of the verbal activity, while locational nouns in prepositional function form a genitival construction⁴ (Kießling 2007b: 178). In such a construction a possessor noun constitutes the center of orientation and follows the locational noun which adds a geometric specification to the center of orientation, which is ‘house’ in (1). Morphologically, most of the nouns in Datooga can be parsed into a root which is followed by a number suffix and a specifying suffix (Kießling 2000, 2007b: 154). Typically of Nilo-Saharan languages (Dimmendaal 2000), there are plenty of allomorphic choices for the number categories in Datooga (see Kießling 2000), but the specifiers are limited to *-da* for singular and *-ga* for plural. The third characteristic aspect of locational nouns in Datooga is their reference to abstract geometric sublocations of objects, e.g. ‘top’, ‘bottom’, ‘front’, ‘back’, ‘side’, ‘middle’, ‘edge’, ‘inside’, ‘outside’. These sublocations are conceptually related to specific geometric features, like the horizontal axis in the case of top and bottom or centrality in the case of ‘middle’ and ‘edge’. All of these sublocational notions also include a three-dimensional conception. In a nutshell, locational nouns in Datooga are nouns which, depending on their contextual usage, express a) spatial relations, b) geometric sublocations and frequently but not always c) body parts.

A challenge during the inquiry was posed by cases, where more than one possibility to express a certain relation occurred. For example, there exist at least four different lexical possibilities to say ‘in front of the house’. Another impression which arose in course of the investigation of the Berger corpus was that, for one and the same spatial concept, different relator nouns are used with different centers of orientation and that their uses might depend on the features animate vs. inanimate. My primary research goal has therefore been to clarify the semantic differences of multi-option cases and to check whether different types of centers of orientation trigger specific occurrences of locational nouns. Especially animals with four legs have been included in order to check the existence of “boviformic” (Kießling et al. 2008: 216) concepts.⁵ Such semantic extensions have been represented only poorly in the Berger corpus and, as the Datooga are pastoralists, it has been suggesting to investigate the range of the boviformic model in the language.

The remainder of this article is organized in the following way: The succeeding section provides an overview of Datooga’s locational nouns. Their stage of grammaticalization is discussed thereafter and conclusions are drawn in the final section.

2 Overview of locational nouns

According to the category and the semantics of their source lexemes, Datooga’s locational nouns can be divided into nominal semantic extensions (§2.1), deverbal locational nouns (§2.2) and opaque cases (§2.3). Distributional factors and semantic differences of overlapping expressions will be presented in section 2.4.

2.1 Locational nouns derived from other nouns by semantic extension

Locational nouns in Datooga derive prevalently from other nouns which express concrete concepts, namely body parts. In general, a development from concrete to abstract concepts is assumed, but in some cases the polysemic situation can also lead to an opposite conclusion.

Before turning to the semantic extension of body part nouns, I will commence the discussion with a rather extraordinary case. The Berger corpus features one noun referring to a cardinal direction which

⁴ In this construction the locational noun possessee proceeds the possessor noun and bears a final associative high tone (Kießling 2007b: 174) which I parse as a suffix in the underlying representation, assuming that this tonal morpheme historically derived from a high toned segmental suffix.

⁵ The term boviformic assumes that the physiognomy of cattle serves as a structural template for the conceptualization of spatial relations. In its assumptions on metaphorical extensions it follows the “zoomorphic” model, as outlined in Heine et al. (1991: 121ff), Svorou (1994: 74) and Heine (1997).

is also used in a relational way. One meaning of *dírbáyda* is the cardinal direction ‘west’ but it also translates as the sublocation ‘frontside’ and in relational function as ‘in front of’.

(3) *dírbáyda* ‘west(side)’, ‘frontside’

gaway	dírbáydá	qeed ^a	‘They went in front of the house.’ ⁶
ga-way	dírbáy-da-	qee-da	
S3-go	frontside-SPC-AS	house-SPC	

Why is the westside the frontside and why is the westside of the house the frontside of the house? The answer to the first question is that *dírbáyda* denotes the direction of the sun’s movement, which in the area close to the equator is, roughly speaking, nothing else but westwards. The answer to the second question, why the westside denotes the frontside of the house, is rooted in the tradition of the Gisamjanga Datooga to build their houses with their entrances facing to the west. This correlates with traditional religious beliefs, because the name for ‘sun’ *aséeta* is also the name for ‘God’. Mietzner (2012: 167) reports that the main axis for cardinal orientation in Nilotic is constituted by the rising and the setting of the sun and that these directions are also connected to religious practices.⁷ Unlike in other Nilotic languages (ibid.), where ‘right’ is used to refer to ‘north’ and ‘left’ to refer to ‘south’, there are neither such extensions in Gisamjanga and Barabaiga, nor are there any alternative expressions for ‘north’ and ‘south’.⁸ The words for ‘east’ in Gisamjanga are *qáyawa* or *gayéew*. In relation to a noun the form with the specifying suffix is used, e.g. *gayéewdá geet^a* ‘east of the tree’.

The situation differs in the Rotigenga variety of Datooga – a 400 km north-western offshoot with Bantu contact, where one finds expressions for all four cardinal directions.

(4) Cardinal directions in Rotigenga

lagidéewi	‘west’	bajúuti	‘north’
gayéewi	‘east’	máseeba	‘south’

The item *lagidéewi* is linked to the verb *lajt* ‘go west’ which is cognate with Gisamjanga *rakt*. This verb translates also as ‘roll down’, which is probably the basic meaning that has been metaphorically extended to the setting of the sun.⁹ The item *gayéewi* has the cognatum *gayéew* ‘east’ in Gisamjanga. It looks like this form contains the 3SG verbal prefix *ga-*. The most resembling verb stem in the BC is *yawd* ‘cry, weep, bray, low, bewail, mourn for’ and there is indeed the convention to perform such activities in the direction of the sunrise. Thus, it seems that the verb form became lexicalized as ‘east’ via metonymic extension of the culturally significant direction of the mourning activity.¹⁰ For the noun *bajúuti*, it seems plausible that it is connected to the ethnonym of the Bajuuta subgroup of the Datooga. Maseeba resembles the name of a large volcanic mountain at the Kenya-Uganda-border, i.e. Mt. Masaba. The quirk is that nowadays the Masaba mountain is north of the Rotigenga and that the Bajuuta are settling south of them. A similar situation can be found in the Bisanjida variety of Datooga (Mietzner 2012:166), where *suqumeera* expresses ‘north’ and *taqameera* ‘south.’ These expressions are linked to the ethnonyms Sukuma and Takama, respectively, which are currently

⁶ Note that the noun *qeeda* refers not only to a single house but also to a fenced-in dwelling with several small huts or houses (Tomikawa 1978: 5).

⁷ For the ‘setting of the sun’ an alternative lexeme – *miktáyda* – is used which derives from the verbal root **mikt* ‘set, sink, be finished’ plus the suffix *-ay* which has a range of functions, such as habitual (Rottland 1982: 186) or pluractional (Kießling 1998: 179). There are more action nominals containing the suffix *-ay* and I assume that it is also featured in *dírbáyda*.

⁸ The expressions for ‘right hand’ are *geetindá dugwa* ‘hand of the cattle’ (HM) or *gaatindá yút* ‘hand of the spear’ (BC). The ‘left hand’ is *geetindá xaadaam* (HM), which means ‘difficult’ or ‘wrong hand’. Compare also the verb form

gi-wadaam-jaan
S2SG-give.wrong.way-1SG
‘you are giving me the difficult/wrong way (of receiving)’ (HM).

⁹ Such a link does not exist not for *gayéewi* ‘east’ and the verb *arosa* ‘go east!’ (Gisamjanga *ádɔɔsa* ‘go east!’).

¹⁰ Thanks to Angelika Mietzner for this suggestion.

neither in direct regional contact nor settling in the derived cardinal directions of the Biyanjida (*ibid.*). It is assumed that these terms lexicalized during a period in which the Biyanjida settled north of the Takama and south of the Sukuma (*ibid.*: 167). This could also be the case with ‘north’ and ‘south’ in Rotigenga, if the above mentioned lexemes have the subsumed etymology.

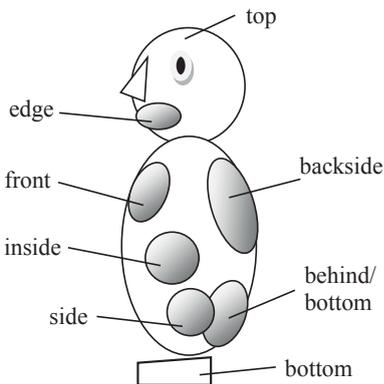
The extension of expressions for body parts is rather straight forward. The following table provides a summary of the body part extensions from the Berger corpus.

(5) Overview of locational nouns derived from body parts

Source	Extension		
	sublocation	prepositional	adverbial
<i>apéega</i> ‘buttocks’	‘behind’, ‘bottom’		temporal: ‘afterwards’
<i>badáyda</i> ‘back’	‘backside’	‘behind’	
<i>dáqáta</i> ‘chest’	‘front’	‘in front of’	
<i>dngáanéeda</i> ‘cheek’			‘sideways’
<i>geefta</i> ‘foot, leg’	‘bottom’		
<i>gujáwéeda</i> ‘horn’	‘corner’		
<i>jeeda</i> ‘belly’	‘inside’	‘in’	
<i>murdéeda</i> ‘hip’	‘side (of body)’	‘near (somebody)’	
<i>quuta</i> ‘mouth’	‘opening’, ‘edge’	‘at the edge of’, ‘close to’	
<i>uhuuda</i> ‘head’	‘top’	‘on’	

The table features two semantic overlappings. Both *apéega* and *geefta* express the sublocational concept ‘bottom’ and *murdéeda* and *quuta* express ‘near’ and ‘close to’ in prepositional usage. In general, most of these extensions apparently follow the anthropomorphic model, while there are only two extensions which clearly orient by the bovimorphic model. The subsequent figure illustrates the scheme of sublocational extensions of human body parts.

(6) Sublocational concepts derived from body parts



Anthropomorphic extensions		
<i>uhuuda</i>	‘head’	> ‘top’
<i>dáqáta</i>	‘chest’	> ‘front’
<i>badáyda</i>	‘back’	> ‘backside’
<i>apéega</i>	‘buttocks’	> ‘bottom’
<i>dngáanéeda</i>	‘cheek’	> ‘sideways’
<i>quuta</i>	‘mouth’	> ‘edge’
<i>jeeda</i>	‘belly’	> ‘inside’
<i>geefta</i>	‘foot’	> ‘bottom’
<i>murdéeda</i>	‘hip’	> ‘side’
Bovimorphic extensions		
<i>gujáwéeda</i>	‘horn’	> ‘corner’
<i>apéega</i>	‘buttocks’	> ‘afterwards’

The metaphoric extensions of *apéega* ‘buttocks’ to ‘bottom’ follow the anthropomorphic model, because if it was following the bovimorphic model, it would be extended to ‘backside’. Interestingly,

the corpus features also one case of a bovimorphic extension of the same source item which is used temporal-adverbially to express ‘afterwards’.¹¹

- (7) gwaéeɟa gadéɛmga ayáaga náa
 gwa-yeɟf-a gadéem-ga á-yaag-a náa
 S3-say-IS women-SPC IMP.PL.1-eat-IMP.PL since
- gaygwááábú apéega sɔuki
 gay-gwa- laab-un apée-ga suuki
 3SG.FUT.AFF- S3-come-CP buttocks-SPC other.things
- ‘He told his wives “Eat, since afterwards other things will come!”’

One speaker I worked with (HM) also used the noun *apéega* to express the concept ‘behind’ in relation to ‘giraffe’. This gives some support for the derivation of ‘afterwards’ via the bovimorphic extension ‘behind’. A case of the use of *apéega* as ‘bottom’ in connection with the inanimate concept *qadéeda* ‘calabash’ is presented in (8).

- (8) gwaéeɟa búunéedá gájeeŋɔólá eedáahá
 gwa-yeɟf-a búun-ee-da gájee-ŋool-a ee-daah-a
 S3-say-IS people-NUM-SPC 1PL.FUT-stir-1PL 1PL.SJN-see-1PL
- apéekadéeda eedáahá gíduurjⁱ.
 apée-ga qadée-da ee-daah-a gí-da wuurj-i
 buttocks-SPC calabash-SPC 1PL.SJN-see-1PL manner-SPC be.like-IS
- ‘The people said: “we want to stir (the honey) in order to see the bottom of the calabash that we may see the (honey’s) quality.”’

The other less common term for ‘bottom’ which derives from *geefta* ‘foot, leg’ is also used to express ‘under’ (see §2.4) and also means ‘base’¹² in modern varieties. In the BC the only instance of *geefta* as ‘base’ or ‘bottom’ occurs in a genitival construction with *geeta* ‘tree’.

- (9) íŋjat geeta qwalac geeftá geeta midéæwá
 ii-ŋad geed-da qwa-laj geeɟ-da-’ geed-da midéæwá
 TEMP-fell tree-SPC S3-cut foot-SPC-AS tree-SPC together.with
- dámáalgáascí sɛɛn
 dámáal-ga-sci seen
 branches-SPC-3SG.POSS.PL all
- ‘When he fells a tree, he cuts it at the bottom of the tree, together with all its branches.’

Clearly following the anthropomorphic model, the body part noun *badáyda* ‘back’ has been extended to render the spatial concept ‘behind’, as the following example illustrates.

- (10) gwanda búunéedá badáydá hílgwaand
 gwa-nd-a búun-ee-da badáy-da-’ hílgw-áan-da
 S3-be.there-IS people-NUM-SPC back-SPC-AS cattle.enclosure-NUM-SPC
- ‘There were people behind the cattle enclosure.’

¹¹ My Barabaiga assistant reports that this reference of *apéega* is uncommon in his variety and that he uses *iiduw* for ‘afterwards’. The BC features *iduu* which translates as ‘maybe, probably, but then’. Both items are related to the adjective *duuw* ‘black, dark, full, rich’ and seem to bear the temporal relative prefix *ii-*.

¹² Thanks to Terril Schrock for the suggestion to look for this meaning.

It was shown in (1) that *dáqáta* ‘chest’ expresses the relational concept ‘in front of’ in relation to the inanimate concept ‘house’. In the following instance *dáqáta* is related to *qaamata* ‘mother’, i.e. a human being.

- (11) ákaŋusá jéépta daqátá qaamat.
 ag-ga-ŋus-a jéép-da dáqád-da-´ qaamad-da
 SEQ- S3-pierce- IS child-SPC chest-SPC-AS mother-SPC
 ‘He pierced the child in front of the mother.’

The relational extension of *dugáanéeda* ‘cheek’ to ‘sideways’ seems to be confined to the adverbial usage illustrated in (2). There is no evidence that this item denotes also ‘side’ or ‘beside’. These latter concepts are expressed by other nouns. One of them is *murdéeda* ‘hip’ which is obviously also metonymically extended to ‘side’. In the following example from the narrative which is named after its hero Mondeya, the noun *murdéeda* is drawn in to express the relational concept ‘near’.

- (12) gawa gágáandá nùtkáaká mida murdéedá qaamat.
 ga-way ga-gaan-d-a nùtkáaC-ga mida murd-ée-da-´ qaamad-da
 S3-go S3-take-CF-IS spears-SPC LOC.REL hip-NUM-SPC-AS mother-SPC
 ‘He went to take the spears near his mother.’

It is actually not only *murdéeda* but *murdéeda* together with the locative-relative copula *mida* which translate as ‘near’ (Swahili ‘karibu’) in (12). A more literal translation would be ‘...spears to (the place) where there was the side of the mother’. Another case will demonstrate below that in modern varieties *murdéeda* can also express ‘beside’ – without *mida* preceding the locational noun.

Just like *pembe* ‘horn’ is extended to express the sublocational concept ‘corner’ in Swahili, the same extension can be found in Datooga with *gujávéeda*.

- (13) úna gujávéedá dúgwa hóowwá
 un-a gujáv-ée-da-´ dú-gwa hóowá
 go-IMP.SG horn-NUM-SPC-AS SPC there
 ‘Go to the corner of the cattle there!’

The context of this passage is that the hero Mondeya tells his companions to take position around a herd of cattle. Since they are four, everyone is assigned to a corner so that they form a rectangle.

The Berger corpus features a lot of cases like the following two, where the noun *jeeda*, which in its original meaning refers to ‘belly’, denotes the ‘inside’ of something or expresses the spatial concept ‘in’.

- (14) goobiidá gádéemgá jeedá béega
 goo-biid-an gadéem-ga jee-da-´ bée-ga
 S3.col-stay-MP women-SPC belly-SPC-AS water-SPC
 ‘The women stayed in the water.’
- (15) gaway qáamátá aa jééptááni jeedá geet
 ga-way qaamad-da aa jéép-da-ni jee-da-´ geed-da
 S3-go mother-SPC and child-SPC-3SG.POSS.SG belly-SPC-AS tree-SPC
 ‘The mother and her child went into the tree.’

The goal *jeedá geet* in (15) really translates as ‘into the tree’ and not ‘under the tree’, because the tree referred to in this story is a big tree with branches reaching the ground, like a bubble tree, for example. Thus, this tree is conceived as a container in (15) into which the mother and her child enter.

The noun *quuta* ‘mouth’ has been extended in two different ways. One target concept is ‘opening’ or ‘entrance’ and the other one is ‘edge’. It also has the potential to express the relational concept ‘close to’, like in the following two examples.

- (16) *náa gwanda qee maháróostá quutá dîmd^a*
náa gwa-nd-a qee mahároosta quud-da-’ dîm-da
 And S3-be.LOC-IS house Maharosta mouth-SPC-AS forest-SPC
 ‘and Maharosta’s house was close to the forest.’

- (17) *naa síilûl jéépta aba quutá baast^a*
náa sí-luul jéep-da aba quud-da-’ bææs-da
 and 3PL.PF-make.sleep child-SPC at mouth-SPC-AS fire-SPC
 ‘And they had made the child sleep close to the fire.’

Note that the preposition *aba* is obligatory in (17) and that this clause would be ungrammatical without *aba*. The preposition *aba* occurs several times before spatial relator nouns in the data corpus. In these contexts it seems that its function is to add a locational argument to a predicate that is filled by a non-motion or non-locational verb.

The final locational noun of the table in (5) is *uhuuda* ‘head’, which has been extended to the sublocational concept ‘top’ in Datooga. The concepts which *uhuuda* relates to in the corpus are *geeda* ‘house’, *geeta* ‘tree’, *gijjeeda* ‘mountain’, *dilgwajáanda* ‘termite hill’ and *habardáanga*. The latter are branches, which are traditionally put in front of the entrance of a homestead, in order to signal that the place is occupied and that everything is fine in that dwelling.

- (18) *gágáandá biyéeydá qwápa iin gwáargwéedá uhuudá habardáang^a*
ga-gaand-a biyéeyda qwápawa yim gwáargwee-da uhuh-da-’ habardáanga
 S3-take-CF-IS piece.of.meat before put man-SPC head-SPC-AS habardáanga
 ‘He took the meat which the man had put on the *habardáanga* before.’

Evidence from modern varieties will demonstrate below in §2.4 that *uhuuda* is also capable of expressing the relational concept ‘above’.

This section has illustrated the metaphoric extensions of ten different body part nouns to spatial concepts. Two of these extensions, i.e. *apéega* ‘buttocks’ > ‘afterwards’ and *gujávéeeda* ‘horn’ > ‘corner’, orient by the bovimorphic model, while the remaining either orient by the anthropomorphic model or are ambiguous. The concept ‘behind’ derives from ‘back’ according to the anthropomorphic model and from ‘buttocks’ according to the bovimorphic model. In the same way, the derivation of the concept ‘in front of’ from ‘chest’ follows the anthropomorphic model, since this concept typically derives from ‘head’ in the bovimorphic model. The derivation of ‘bottom’ from ‘buttocks’ is also an anthropomorphic extension. Hence, clearly anthropomorphic extensions feature the items *apéega* ‘buttocks’, *badáyda* ‘back’, *dáqáta* ‘chest’ and *geesta* ‘foot’ in Datooga. The argument for *geesta* ‘foot’ to be oriented by the anthropomorphic model is the presence of *biùléedà* ‘hoof’ which would be the bovimorphic source.

2.2 Deverbal locational nouns

Beyond nominal semantic extensions Datooga features also locational nouns which derive from spatial verbs. Consider the noun *gwéenda* in the following two examples, which expresses the sublocational concept ‘middle’ and the relational concepts ‘between’ or ‘among’.

- (19) *gáalaabú hòo bææ gwéendá qeedaan*
gaa-laab-un hoo baa gwéen-da-’ qee-da-n
 1SG-come-CP there POSS middle-SPC-AS settlement-SPC-DEM.POSS
 ‘I come from there among the settlement.’

- (20) qeedídá baa gwééndá qeedaan
 qee-da -ida baa gwéen-da-’ qee-da-n
 house-SPC-DEM.FAR.SG POSS middle-SPC-AS settlement-SPC-DEM.POSS
 ‘That house in the middle of the settlement.’

According to Rottland (1982: 167) it is possible that verbs convert to nouns merely by suffixation of the nominal suffix *-da* (SG) or *-ga* (PL) to the verbal root or stem. In the case of *gwéen-da*, it appears plausible that the noun derives from the verbal root *gween* ‘be in the middle, be at the zenith’, which is exemplified in (21).

- (21) áamá agwéená awee-da iiyeṇa qwájáat qeet.
 ama a-gween-a awee-da iiyeṇa qwa-jaad qee-da
 when 3SG.REL-be.in.middle-IS night-SPC two S3-open house-SPC
 ‘In the middle of the night he opened the house.’

Paul Berger has noted that *iiyeṇa* ‘two’ is included in (21) to express the concept ‘midnight’ as being in the middle of the two halves of the night.

The following two deverbal forms *hiijáadáyda* and *diiyáadáyda* resemble each other both morphologically and semantically. Both of them contain an identical sequence of three verbal suffixes preceding the nominal suffix *-da*. Furthermore, the two nouns are used to express a vertical relation to an object in a slope situation. One of my Gisamjanga assistants (HM) reported that if you instruct someone on the way and you have a slope situation with a tree as center of orientation, you can tell him that something is located *hiijáadáyda geet^a* ‘downward from the tree’ or *diiyáadáyda geet^a* ‘upward from the tree’. The noun *hiijáadáyda* literally expresses a ‘state of going down’, i.e. an activity. Two other expressions in the Berger corpus which built on the verbal root **hiij* are featured in the following passage from the narrative *heydeesh* ‘the white bullock’.

- (22) qwáhiijáadá niirjá, gwaéeḟa niirjá hiijaana nimi
 qwa-hiijaa-d-a níirjá gwa-yeef-a níirjá hiijaa-n-a ni-meew-i
 S3-descend-CF-IS other.SG S3-say-IS other.SG descend-CP-IMP.SG 3SG.PF-die-IS
 ‘The other one descended, the first one said: “Descend, he is dead!”’

In addition, there also exists the antonymic directional stems **diiyaa-d* ‘ascend thither’ and **diiyaa-n* ‘ascend hither’. The two verb stems **hiijaa-d* and **diiyaa-d* can be further analyzed as containing the mobilitive suffix *-aa* (Kiebling 2007a: 124) which indicates that the verbal action is performed against the background of a motion event, but the semantics of the bare roots is obscure.¹³ In the case of the noun *hiijáadáyda* the verbal base with the centrifugal suffix *-d* is extended further by the suffix *-ay*. The meaning described by my consultant as a “state” resembles that of the verbal suffix *-ay* which has a range of facets like habitual, frequentative (Rottland 1982: 184, 186) or pluractional, durative and intensive (Kiebling 1998). The noun *diiy-áa-d-áy-da* features the same derivational suffixes as *hiij-áa-d-áy-da*, namely mobilitive *-aa*, centrifugal *-d* and durative *-ay*. There exist probably more deverbal locational nouns beyond those mentioned in this section and it will be demonstrated below that Datooga features also prepositions which historically have derived from verbs.

2.3 Opaque locational nouns

The following table sums up the few locational nouns which are not transparent with respect to their source concepts in the data collected by Berger. The rightmost column lists potential sources or links of these nouns.

¹³ The stem **hiijaa-d* is linked to the stem **hiij* ‘bend, stoop’, which can be parsed as **hiij-j*, i.e. the root with the causative extension *-j* (Kiebling 2007a: 124).

(23) Overview of non-transparent locational nouns in the Berger corpus

Dat. noun	Abstract reference	Relational reference	Potential source/link
<i>bugústa</i>	'middle', 'center'	'among', 'in the middle of'	Barabaiga: 'waist'
<i>góomda</i>	'bottom'	'at the bottom of', 'under'	'shadow (of tree)'?
<i>héeda</i>	'place'	'at', 'to'	?
<i>járooda</i>	'outside'	'outside'	<i>jaróoneeda</i> 'shadow'
<i>máatkwéeda</i>	'side'	'beside'	Proto-Southern-Nilotic * <i>matay/mataay</i> 'cheek' (Rottland 1982: 302) <i>mádáæŋda</i> 'cheek'; more exactly: 'fleshly part between the cheek bone and the lower jar'; plural suffix - <i>gwee</i> ?
<i>madeeda</i>	'frontside'	'in front of'	Common Datooga * <i>mad</i> 'first' (ibid.: 487)
<i>úuta</i>	'backside'	'behind'	Barabaiga: 'buttock'

The previous section has presented the deverbal noun *gwéenda* 'middle' which seems to have overlapping semantics with *bugústa* that also expresses 'middle' and 'center'. My Barabaiga assistant told me that *bugústa* also refers to 'waist' in his variety. Measured by the average human body, the waist is approximately located at the middle of the vertical head-to-foot-axis. Thus, the extension of *bugústa* can be counted as a further metaphoric transfer which follows the anthropomorphic model.

The noun *góomda* frequently expresses 'bottom' and 'under'. In addition, there is one case where it translates as 'shadow'.

- (24) qooq̄iis sáwá góomdá ḡeet^a
 qoo-q̄iis saawa góom-da- geed-da
 S3SG-take.a.place 3PL shadow-SPC-AS tree-SPC
 'They entered into the shadow of the tree.'

It would make some sense that 'shadow' has been recruited to express the concept 'under' since shadow is perceived predominantly on the ground and in certain cases also under objects, when a light source is shining on them from above. However, there is only this one case in the corpus and my Barabaiga consultant, who is the most advanced with respect to age, has confidently denied that *góomda* means 'shadow'. According to him the only word for 'shadow' is *jaróoneeda* which also has the variant *jiróoneeda* in the BC. This noun strongly resembles the Datooga expression for 'outside' – *járooda*, which occurs exclusively adverbially and never as possessee in a genitival construction in the Berger corpus. Kießling (2000: 353) states that nasal deletion applies to a root-final alveolar nasal in collision with the nominal suffix *-da*. Hence, I assume that *járooda* has the underlying shape *jároon-da*. The form *jaróoneeda* contains the singular suffix *-ee* (ibid.: 352), which can be hold responsible for the tonal and maybe also the semantic difference between the two forms. The semantic link is very tentative, but one could imagine that the suffix *-ee* has antonymic function in this case and that 'shadow' was also used to refer to 'inside'.

It has been demonstrated in (1) that *dingáanéeda* 'cheek' can express the concept 'sideways'. In a similar fashion the noun *máatkwéeda* which refers to 'side' could derive from the partially matching noun *mádáæŋda* 'cheek', which is composed of the root **mád* and the individual suffix *-aan* (ibid.) and the specifying singular suffix *-da*. One problem with this derivation is to account for the length difference of the root vowel. Another oddity is that the form *máatkwéeda* would be analyzed as containing the plural suffix *-gwee* (ibid.) while it occurs with the singular suffix *-da*.

In other languages the term for 'frontside' has been derived from 'face', but this is apparently not the situation with *madeeda* 'frontside' in Datooga. According to my language helpers, *madeeda* also refers to the 'genital part' or the 'genital area'. Furthermore, Rottland (1982: 487) lists the Common

Datooga root **mad* which expresses the ordinal concept ‘first’. If we assume again a development from the concrete to the abstract, the extension would be from ‘genital part’ to ‘frontside’ to ‘first’. However, in this case it seems more convincing that ‘frontside’ has been extended to ‘genital part’ in order to avoid a more vulgar term.¹⁴ The extension from ‘frontside’ to ‘first’ seems sound, though.

The final item from the table in (23) is *úuta* which expresses ‘backside’ or ‘behind’. This noun occurs four times and exclusively in relation to *qeeda* ‘house’ in the Berger corpus, as in the following example.

- | | | | | |
|------|------------|-----------|------------------|-------------------|
| (25) | qafka | bárbayíǵá | úutá | qeed ^a |
| | qa-fuk-a | barbayíǵa | úud-da-’ | qee-da |
| | S3-come-IS | Barabaiga | back.side-SPC-AS | house-SPC |
- ‘The Barabaiga came to the backside of the house.’

In contemporary Barabaiga and Gisamjanga, *úuta* also refers to ‘buttock’. This concept is expressed by *ápeeda* in the Berger data, which refers to ‘bum cheek’ or ‘buttock’. It is not uncommon to have more than one expression referring to this part of the body. For example, in German there are various expressions for ‘buttock’, such as *Hinterbacke* ‘cheek of behind’, *Po(dex)*, which is a loan word from Latin, or *Gesäßbacke* ‘cheek of the thing which you sit on’. The extension of *úuta* apparently orients by the bovimorphic model, since, according to the anthropomorphic model, ‘buttock’ would be extended to the concept ‘bottom’. Whether *úuta* can be related to more locations is a question that leads us to the next section of this paper, which treats governing factors that determine the choices of locational nouns.

2.4 Distribution and semantic differences of overlapping expressions

This section takes a look at the different concepts which co-occur with locational nouns in relational function. The semantics of the figure or the object which is related to the locational center does not influence the choice of the relator noun. The basic question is: Can any noun specify any concept or do certain locational nouns only occur with specific centers of orientation? I will begin with the human being as the first center of orientation, then proceed with animals with four legs and end with different other concepts, including ‘tree’, ‘mountain’, ‘house’ and other artifacts. As the base of comparison, a static situation with the relational concepts ‘in front of’, ‘behind’, ‘beside’, ‘under’ and ‘above’ is generally taken into account. It is assumed that the relation of the concerned objects is always a detached one.

For the human body I elicited objects in relation to the speaker as the center of orientation. Thus, the relator nouns acquire a possessive suffix of the first person singular for singular possessors, i.e. -*ju*. This lengthens the vowel of the preceding specifying suffix *-da* and changes its quality to [ɛ]. The following Gisamjanga examples have been contributed by Emmanuel Dudiyeck. Note that they are not marked for tone.

- | | | | |
|------|----|---|-------------------------------|
| (26) | a. | <i>duuga gwanda madee-dɛɛ-ju</i> | ‘the cows are in front of me’ |
| | b. | <i>deeda gwanda baday-dɛɛ-ju</i> | ‘the cow is behind me’ |
| | c. | <i>deeda gwanda murdee-dɛɛ-ju</i> | ‘the cow is beside me’ |
| | d. | <i>haqweega gwanda uhuu-dɛɛ-ju</i> | ‘the fruits are above me’ |
| | e. | <i>horgeeda gwanda hiijɛɛday-dɛɛ-ju</i> | ‘the chair is below me’ |

It is not natural to use *daqatɛɛ-ju* or *quutɛɛ-ju* to convey the meaning ‘in front of me’ in (26a). These items would only express the literal meaning ‘at my chest’ or ‘at my mouth’, respectively. Similarly, *baday-dɛɛ-ju* in (26b) is not interchangeable with the reflexes of the nouns *uuta*, *apeeda* or *huydeenda*; the latter means ‘end’ and is also recruited to express ‘behind’ in relation to (a) quadruped animal(s) and ‘tree’ in Rotigenga. The usage of these nouns would be vulgar in this context. In the third sentence in (26) it is possible to replace *murdee-dɛɛ-ju* by *meetkwee-dɛɛ-ju*. The

¹⁴ I am thankful to Mechthild Reh for this comment on my presentation at the department of African languages and Ethiopian Studies of the University of Hamburg.

difference is that *murdee-dee-ju* has a connotation of proximity to the speaker and *meetkwee-dee-ju* not. Finally, if the possessor concept is in plural, it is not grammatical to use the plural forms of the relator nouns. So, for ‘in front of/behind/beside us’ you would say *madee-dee-ja/baday-dee-ja/murdee-dee-ja*. The loss of the potential to pluralize in these cases suggests that desemanticization has happened to these relator nouns.

It is interesting that a reflex of the deverbal noun *hiinjáadáyda* ‘state of going down, downward’, which has been analyzed more thoroughly in §2.2, is used in (26e) to express ‘below’. It remains to be investigated whether the antonym *diyáadáyda* could also be drawn in to say ‘above’ in (26d).¹⁵

Turning to the quadrupedic concepts as related centers of orientation, the following table summarizes the different co-occurrences.

(27) Locational nouns related to animals with four legs

deictic reference	Datooga noun	original meaning	concept related to	source
‘in front of’	<i>daqáta</i>	‘chest’	cows	HM
	<i>madeeda</i>	?	lion, cows	HM, BC
	<i>mareera</i>	?	lion	JM
‘behind’	<i>apéega</i>	‘buttocks’	giraffe	HM
	<i>huydeenda</i>	‘end’, ?	giraffe(s)	JM
‘beside’	<i>meedugweera</i>	?	cattle	JM
	<i>murdeeda</i>	‘hip’	cattle	FB
‘above’	<i>uhuuda</i>	‘head’	cow	HM
	<i>badáyda</i>	‘back’	cows	HM
‘under’	<i>jeeda</i>	‘belly’	cow	HM
	<i>geefia</i>	‘foot’	cow	JM

There is a slight semantic difference between the two ways to express ‘in front of’. The noun *daqáta* expresses more proximity to the center of orientation than the noun *madeeda*, which has the shape *mareera* in Rotigenga. It is not grammatical to use *daqáta* ‘chest’ with the singular concepts ‘lion’ and ‘cow’. Neither is there a regular use of *uhuuda* ‘head’ in order to express ‘in front of’ in these two contexts.

Two alternative expressions for the concept ‘behind’ can be found. The first expression derives from the body part *apéega* ‘buttocks’ which has been treated above. The item *huydeenda* is obscure with respect to its origin, but in modern Gisamjanga *huydeenda* encodes also the concept ‘end’.

The difference between *murdeeda* and *meedugweera* (reflex of *máatkwéeda* in Rotigenga) may be that – analogous to human and inanimate centers of orientation – the former expresses proximity, i.e. ‘close to’, while the latter just denotes ‘at a certain side’ or ‘beside’ without a connotation of relative distance. Again, it is important to note that a singular noun – *murdeeda* ‘hip’ – co-occurs with a plural concept *dugwa* ‘cattle’, which illustrates the desemanticization of *murdeeda*.

It has been demonstrated in (26d) that *uhuuda* can also express ‘above (a human being)’. In the same way it can also denote ‘above’ an animal with four legs, as in the following example.

- (28) *daridáænda* *gwánda* *uhuudá* *dêed^a*
darid-ææn-da *gwa-nd-a* *uhuu-da-’* *dee-da*
bird-NUM-SPC 3SG-be.LOC-IS head-SPC-AS cow-SPC
‘The bird is above the cow.’ (HM)

¹⁵ My assistant notes that the sentence in (26d) can be translated alternatively as *haqweega qwalabwanani hetiyæeda*. Likewise, (26e) translates also as *horgeeda qwalabwanani jeenyi*. I assume these sentences mean something like ‘the fruits are past me up(ward)’ and ‘the chair is past me down(ward)’, respectively.

An alternative expression for the concept ‘above’ is achieved by the body part *badayda* ‘back’ in relation to a plural concept.

- (29) udóowaŋg χάηετα badáydá dugw^a
 udóowaŋ-ga qa-ηεεt-a baday-da-’ du-gwa
 crane.bird-SPC 3SG-fly-IS back-SPC-AS cows-SPC
 ‘A crane bird flies above the cows.’

Another instance where the body part ‘back’ is used to express ‘above’ can be found in the domain of inanimate centers of orientation. There, the ‘back’ of certain objects which have greater length than height can express the top part of them.

In addition to the locational nouns *hiijáadáyda* and *goomda* which are used to refer to the relational concept ‘under’, there exist two further items in relation to animals with four legs expressing ‘under’. The first noun is *jeeda* ‘belly’ which is normally extended to the concept ‘inside’. In the following instance it is recruited to express ‘under’.

- (30) jéepta gwánda jéeteed^a
 jéeb-da gwa-nd-a jee-da-’ =dee-da
 child-SPC 3SG-be.LOC-IS belly-SPC-AS=cow-SPC
 ‘The child is under the cow’ (HM). Note that this sentence can also translate as ‘the child is in the cow’

When I asked my Rotigenga consultant to translate some similar sentences, he used the noun *geefia* ‘foot, leg’ to express ‘under’ or ‘below’.

- (31) a. *riibígá¹ gwándá geefiá rêer* ‘The children are below the cow.’
 b. *riibígá¹ gwándá geefiá rúgw^a* ‘The children are below the cows.’
 c. *riibígá¹ gwándá geefanjá rúgw^a* ‘The children are below the cows.’

Evidence for the desemanticization and expression of ‘under’ through *geefia* comes from (31b), where the locational noun is in singular, while the related concept *rúgw^a* ‘cows’ is pluralized. Note, however, the alternative sentence in (31c) where the locational noun is pluralized. This indicates a retainment of a nominal property and a transitional stage of this noun. We will deal with more of such cases in the next section.

Finally, I will examine various, mostly inanimate concepts as centers of orientation and the relator nouns which are used together with them. The following table provides an overview. I have elicited sentences with ‘house’, ‘tree’ and ‘mountain’ as centers of orientation. Additional concepts from the Berger corpus are listed as well.

There are four different ways of saying ‘in front of the house’ in (32). The differences of the individual expressions are as follows: The noun for ‘chest’, *daqáta*, expresses, again, greater proximity than *madeeda*, which seems to be unspecified with respect to distance. The cardinal direction *dirbáyda* ‘west’ is related to an alternative frame of reference, but can denote the ‘frontside’ of the house due to a conventionalized construction with entrances facing to the west among certain Datooga groups. The noun *quta* ‘mouth’ is also drawn in to express ‘in front of the house’. It differs from the other expressions in the way that it has the connotation of the space close to the door.

Concerning the three different expressions for ‘behind’, *úuta* occurs only with ‘house’.¹⁶ Like in relation to animals with four legs, the formulation of my Rotigenga assistant differs in his use of *huydeenda*. It seems to be an expression which is unspecified for proximity and, in this regard, analogous to *madeeda* in the case of ‘in front of’ and *méetkwéeda* for expressing ‘beside’.

Both *méetkwéeda* and *murdeeda* can express ‘beside the house’. The latter one has a proximal connotation. Unfortunately, I am not aware whether these expressions can be related to ‘tree’ as well,

¹⁶ According to ED it is not correct to use *apéega* ‘buttocks’ in relation to house in order to express the concept ‘behind’.

but, according to ED, ‘beside the mountain’ is *murdeeda gijeeda* and *méetkwéeda gijeeda* is not common speech.

(32) Locational nouns related to inanimate concepts

Deictic reference	Datooga noun	Original meaning	Concept related to	Source
‘in front of’	<i>daqáta</i>	‘chest’	house	BC, FB
	<i>dayaráyda</i>	‘chest’	mountain	JM
	<i>dirbáyda</i>	‘west’	house	BC, FB
	<i>madeeda</i>	?	house	ED, HM
		?	pleura	BC
	<i>mareeri</i>	?	mountain, tree	JM
	<i>quta</i>	‘mouth’	house	BC, FB
‘behind’	<i>badáyda</i>	‘back’	mountain, house	HM, JM, ED
	<i>huydeenda</i>	‘end’ ?	tree	JM
	<i>úuta</i>	‘buttock’	house	FB, HM, BC
‘beside’	<i>máatkwéeda</i>	?	house	BC
	<i>méetkwéeda</i>	?	house	ED
	<i>murdeeda</i>	‘hip’	house	FB
	<i>murdeeda</i>	‘hip’	mountain	ED
‘above’	<i>uhuuda</i>	‘head’	tree	HM
	<i>uhuura</i>	‘head’	tree, mountain	JM
	<i>badáyda</i>	‘back’	house/mountain of special shape	HM
‘under’	<i>gumida</i>	?	tree	JM
	<i>guumda</i>	?	tree, mountain	ED
	<i>góomda</i>	?	tree, sticks	BC

The most frequently used expression for ‘above’ is *uhuuda*, which has the concrete meaning ‘head’. Herman Malleyeck reports that the noun *badáyda* ‘back’ can also refer to the top part of a building or a mountain when the object has a greater length or width than height. In such cases *badáyda* is apparently also used to refer to an area which is detached from the object and thus indicates ‘above’.

Unlike in all of the other above cases of spatial relations to inanimate concepts, where more than one choice of noun is possible, the only expression for ‘under’ in relation to inanimate concepts is a reflex of the opaque noun *góomda*.

To summarize the patterns in the data, the following observations can be made: There are more opaque grammaticalized locational nouns, such as *madeeda* ‘in front of’, *méetkwéeda* ‘beside’ and *huydeenda* ‘behind’, which are unspecified with respect to relative distance. These nouns can be used in relation to almost every type of center of orientation, i.e. humans or bipedes, animals with four legs and inanimate concepts. In addition, there are expressions which semantically overlap with the former expressions. However, they seem to have a greater granularity and they are further specified for proximity, as they have the connotation of being close to a specific sublocation of the related center of orientation. These are *daqáta* for ‘in front of’, *murdeeda* for ‘beside’, and *badáyda* for ‘behind’. These expressions are also more constrained, as they cannot be related to any conceptual type. The noun *daqáta* can be related to all concepts, except quadrupedic singular and ‘tree’. The item *murdeeda* can be related to all types of concepts but needs further investigation for quadrupedic singular and ‘tree’, and *badáyda* can only be related to bipedic and certain inanimate concepts but not to animals with four legs to express ‘behind’. The most constrained and least grammaticalized items are *úuta* ‘behind’, which only occurs in relation to ‘house’, and *apéega* ‘behind’, which only relates to an animal with four legs. Finally, I have pointed out that the bovimorphic extensions of *jeeda* ‘belly’ and *geefta* ‘foot, leg’ to ‘under’ or *badáyda* ‘back’ to ‘above’ occur only with quadrupedic centers of orientation. The only potential contextual transfer of these extension has been found with *badáyda* ‘back’ to ‘above’ in

relations to inanimate objects, which resemble quadrupedic concepts in the way that they have greater horizontal than vertical range.

3 Stage of grammaticalization

In order to determine whether there are indications that locational nouns are developing further towards prepositions and adverbs, I will first take a look at typical representatives of the two grammatical target categories and demonstrate their morphological differences from nouns. Thereafter, I will discuss cases which speak for and against category shifts of locational nouns.

According to their source categories Heine (1997: 49) distinguishes between N-adpositions (deriving from nouns), V-adpositions (from verbs) and A-adpositions (deriving from adjectives). The fact that Datooga's spatial prepositions differ in their morphological make-up from locational nouns is due to their verbal origin.

(33) Overview of spatial V-prepositions

Preposition	Potential source verb
<i>aba</i> 'at', 'from', 'among' ('because of', 'with')	<i>baa</i> 'have'
<i>waa</i> 'leading to', 'in the direction of', 'towards'	<i>way</i> 'go'
<i>neekéewá</i> 'near'	<i>neek</i> 'draw near', 'approach'

The preposition *aba* is multifunctional and expresses the meanings 'at', 'from', 'because of', 'with', 'among'. The lexical source of *aba* is likely to be the resembling the verb *baa* 'have'. The most resembling verb form in the BC is *g-a-ba* 'he/she/it has', whereas the initial consonant is analyzed by Rottland (1982: 176) as the affirmative marker *g-*, which is followed by the third person singular subject prefix *a-* and the reduced root. The form *aba* – without the affirmative prefix – looks like what Rottland describes as the dependent verb form which surfaces after modal verbs (ibid.: 192).

The following example illustrates the use of the preposition *waa* 'leading to', 'in the direction of', 'towards'.

- (34) *gábá súu sillá náweedá waa qeedá mida síjána.*
gaba súuwa si-laab náweeda waa qeeda midá síi-pawá
 every 3PL 3PL.PF-walk path towards house LOC.REL person-3PL.POSS.SG
 'All of them had walked the path towards the home where their people were.'

The verb *way* 'go' is likely to be the source of this preposition because it has the reflex *waa* as a temporal relative form of the third person singular ('when he/she/it goes'). Classifying *waa* as a preposition on the base of this sentence is problematic, though, because at the moment it can not be excluded that *waa* is a relative clause form of the verb.

The third preposition *neekéewá* 'near (to)' obviously derives from the verbal root *neek* 'draw near', 'approach', which is extended by a long reflex of the purposive suffix *-aw* (Kießling 2007a: 124). The rationale of categorizing *neekéewa* as a preposition and not a verb in the following example is simply that it is located in a non-relative clause and that it surfaces without inflectional verbal prefixes.

- (35) *náa neekéewá héedá gwáci im biyéydaap qwanaawa habardáanga*
náa neekáawá héeda gwáci yim biyéyda-ji qwa-naw-a habardáanga
 and near.to place before put meat-3SG.POSS.SG S3-pass-IS habardanga
 'And near to the place, where he had put on his meat before, he passed the habardanga.'

Turning to the spatial adverbs of Datooga, the Berger corpus features the following items.

(36) Overview of spatial adverbs

Adverb	Potential source
<i>qoo</i> ‘home(wards)’	<i>qeeda</i> ‘homestead’
<i>gaa</i> ‘inside (the house)’	<i>geeda</i> ‘tent’
<i>badaw</i> ‘backwards’	<i>badayda</i> ‘back’
<i>ɲɛæɲi</i> ‘down’	<i>ɲɛæɲiida</i> ‘earth’, ‘ground’, ‘land’, ‘world’

The adverb *qoo* most probably derives from *qeeda* ‘homestead, house’, but the development of the back vowel in *qoo* lacks an explanation at this moment. The item *gaa* resembles *geeda* ‘tent’. Rottland (1982: 172) reports that in the Rotigenga variety the indefinite or unspecified form of the noun can also express a locative meaning, i.e. *gahi* ‘any house’ and ‘at home’. The phonetic and semantic differences between *gaa* and *gahi* are only small. This could indicate that *gaa* derives also from the unspecified form of *geeda* in Gisamjanga.

The adverbial form *badaw* is related to the item *badáyda* ‘back’. A labio-velar glide regularly becomes palatal before *-da*. Hence, *badaw* seems to be the nominal root.

The Berger data feature the noun *ɲɛæɲiida* ‘earth’, ‘ground’, ‘land’, ‘world’. The adverbial *ɲɛæɲi* expresses the meanings ‘below’, ‘down’, ‘on the ground’. This semantic extension of ‘earth’ or ‘ground’ to ‘down’ is also reported to be common among Bantu languages, whereas in Western Nilotic it is ‘buttocks’ which is extended to the deictic reference point ‘down’ (Heine 1997: 49). In the case of *ɲɛæɲi* it is likely that this is the unspecific form of *ɲɛæɲiida*, which is built by the suffixation of *-i* to the primary form of the noun (Rottland 1982: 169; Creider & Rottland 1997: 97), that is the stem plus an eventual number suffix. In the form *ɲɛæɲi* the final vowel would have been shortened then.

What can be observed from all the adverbs and prepositions is that they don’t have the nominal suffixes *-da/-ga*. For the prepositions, this is not surprising, since they derive from verbs historically.

A further difference between prepositions and nouns is that prepositions do not change tonally but surface with their underlying melody when they introduce locations.

- (37) qoo sáawá fú-gá hees aba hée-dá fú-gá duuwas
 qoo saawa fú-ga hees aba hée-da-’ fú-ga-’ duuwas
 become 3PL people-SPC big at place-SPC-AS peoples-SPC-AS black.PL
 ‘They became a great people among the black peoples.’

In (37) *aba* surfaces with its underlying low-low melody. However, in genitival constructions of locational nouns, a high tone regularly associates with the final vowel of the locational noun, as can be observed from previous examples.

If locational nouns were developing to adverbs and prepositions, one would expect them to lose nominal properties like the bearing of nominal suffixes and the tonal marking within the possessive construction. Tonally, locational nouns follow the associative pattern very regularly. However, with respect to the nominal suffixes there are in fact cases of locational nouns without the specifying suffixes *-da/-ga*. In the following example you can note that *jeeda* ‘belly’ has the shape *jée*.¹⁷

- (38) qoobíta jée foomandá bicâand^a
 qoo-bid-d-a jee-da foom-aan-da bic-aan-da
 S3SG-enter-CF-IS belly-SPC pit-NUM-SPC wild.boar-NUM-SPC
 ‘She entered into the pit of a wild boar.’

According to my Barabaiga language helper the sentence in (38) is grammatical and the dropping of *-da* is typical of fast speech and applies not only to locational nouns. There is no evidence that the dropping of the suffix is conditioned by other factors. Other situations of locational nouns without

¹⁷ It could be argued here that the associative terminal high tone links to the remaining root syllable (Kießling, p.c.).

specifying suffixes do exist. However, they can not be interpreted as traces of categorial shifting because there the dropping of the specifying suffixes does not only apply to locational but to all nouns. The specifying suffix is, for example, generally dropped if the noun is a) modified by a proximal demonstrative suffix or b) possessee of a possessor noun modified by a numeral.

Creider & Rottland (1997: 79) found that the proximal demonstrative suffixes *-(n)i* (SG) and *-su* (PL) are attached to the primary form of the noun. This finding is confirmed by the Berger material, e.g. in the following sample.

- (39) óo-piidán-á jee-ni
 IMP.PL-stay.at-IS inside-DEM.NEAR
 ‘Stay in there!’

The data in (40), which were contributed by my Barabaiga consultant, demonstrate that in constructions with possessor nouns modified by numerals the clitic *mba* is used and the specifying suffix does not surface.

- (40) a. jee-dá hílgwééndá χeed^a ‘in the enclosure of the house’
 b. jee-mbá χáawâakⁱ ‘in one house’
 c. jee-mbá gééliyeɛɲa ‘in two houses’
 d. jee-dá-bá hílgwéenjⁱ ‘in these enclosures’
 e. jee-dá-bá géɛlsú asiiyeɛɲaan ‘in these two enclosures’

The phrase in (40e) indicates that the demonstrative rules out the numeral construction, because *jee-dá=bá* is used instead of *jee=mbá*. The occurrence of *jee-mbá* is governed by the numeral and **jeeda mba* is ungrammatical.

To conclude, there are no cases speaking clearly for a category shift of locational nouns. It seems that they are rather very noun-like. They are capable of taking possessive and demonstrative suffixes, as exemplified in (26) and (39), respectively. In addition, many cases in the data show that they still have the potential to be pluralized in relation to plural possessors, like *jéesingá* – the plural of *jeeda* ‘belly’ – in (41) and *quusigá* – the plural of *quuta* ‘mouth’ – in (42).

- (41) qwáyɪn jéesingá mʃóodaká digéeniɲaasc
 qwa-yɪn jéesiŋ-ga-’ mʃóodaC-ga-’ digéeni-ga-sci
 S3-put bellies-SPC-of leather.sacks-SPC-AS donkeys-SPC-3SG.POSS.PL
 ‘He put it into the leather sacks of his donkeys.’

- (42) gwaduuléesá quusigá báascéeka sêen
 gwa-duulees-a quusi-ga-’ báas-jéeC-ga seen
 S3-gather-3 mouths-SPC-AS fires-NUM-SPC all
 ‘They gathered near all the fires.’

According to my Barabaiga helper, the locational nouns in the preceding two examples can be replaced by their singular equivalents *jeeda* and *quuta*. Such an optionality has been demonstrated before in (31) with *geefanga* and *geefita* and applies probably also to more locational nouns. The fact that locational nouns in singular relate to centers of orientation in plural is an indication of their desemantization to abstract spatial concepts. The retainment of nominal properties such as the potential to pluralize and to acquire possessive and demonstrative suffixes, speaks, however, for an early stage in their grammaticalization.

Concerning the development of locational nouns to adverbs, category shift may happen more easily than in the case of prepositions via the unspecified or primary forms of the nouns, as it seems to be the case with some of the adverbs presented in this section.

4 Conclusion

Locational nouns – defined as nouns which express geometric sublocations and function as spatial relators – most frequently derive from human body parts in Datooga. In addition, a few nouns exist which are of verbal origin. Furthermore, some nouns of obscure origin have been presented. Items of this type, such as *madeeda* ‘in front of’, *huydeenda* ‘behind’, and *méetkwéeda* ‘beside’, do not have a specification for relative distance of their related objects, nor are they very restricted in their distribution. They can be related to almost any type of center of orientation, i.e. human, non-human animates and various inanimate concepts. On the other hand, their transparent correlates *daqáta* ‘in front of’ (< ‘chest’), *badáyda* ‘behind’ (< ‘back’) and *murdeeda* ‘beside’ (< ‘hip’) often have a connotation of proximity of their related objects. The nouns *daqáta* and *badáyda* can also not be related to any center of orientation. They can not be associated with the category quadrupedic singular, for instance, to express ‘in front of’ and ‘behind’, respectively. Other transparent locational nouns are even more limited in their co-occurrence to one type of center of orientation. For example, the noun *uuta* ‘buttock’ expresses the concept ‘behind’, but only in relation to the concept ‘house’. The bovimorphic extension of *jeeda* ‘belly’ to ‘under’ only occurs with quadrupedic singular and *badáyda* ‘back’ to ‘above’ seems to be limited to quadrupedic concepts and certain inanimate concepts with greater horizontal than vertical range.

The presented data give the impression that some locational nouns have been grammaticalized further than others and that the anthropomorphic extensions are much more prevalent than the bovimorphic. The different conceptual types of the centers of orientation govern the choices of relator nouns, whereas the category number seems also to be a determining factor.

A shared feature of locational nouns is their bearing of the specifying suffixes *-da* (SG) and *-ga* (PL). This distinguishes them from the categories adverbs and prepositions. Certain adverbs seem to have developed from nouns historically, probably via the unspecific forms of these nouns. There are, however, no convincing indications of category shifts to prepositions. The few prepositions in Datooga appear to derive from verbs historically. Instances of locational nouns in prepositional function, where the nouns surface without specifying suffixes, have been presented. The droppings of the suffixes in these situations have been identified as resulting from fast speech or as grammatically conditioned droppings.

For further comprehensive research, I would suggest to include a greater variety of orientational center concepts of different shapes and to investigate number oppositions of all kinds of concepts in order to understand the principles underlying the assignment of locational nouns better. This study has mainly focussed on static detached situations, and spatial relations of moving objects as well as attached relations could also be taken into account.

Abbreviations

AFF	affirmative	MP	multipurpose
AS	associative	NUM	number
CF	centrifugal	PF	perfective
CP	centripetal	PL	plural
DEM	demonstrative	POSS	possessive
ED	Emmanuel Dudiyeck (Gisamjanga)	REL	relative
FB	Festo Basso (Barabaiga)	S	subject
FUT	future	SEQ	sequential
HM	Herman Malleyeck (Gisamjanga)	SG	singular
IMP	imperative	SJN	subjunctive
IS	inflectional suffix (of uncertain function)	SPC	specifier
JM	Josephat Masaka (Rotigenga)	TEMP	temporal
LOC	locative	1/2/3	first/second/third person

References

- Andersen, Torben. 2012. Verbal directionality and argument alternation in Dinka. In Angelika Mietzner & Ulrike Claudi (eds.), *Directionality in Grammar and Discourse: Case Studies from Africa*, pp. 35–54. Cologne: Köppe.
- Dimmendaal, Gerrit J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42: 214–261.
- Heine, Bernd. 1989. Adpositions in African Languages. *Linguistique Africaine* 2: 77–127.
- Heine, Bernd. 1997. *Cognitive Foundations of Grammar*. New York/Oxford: Oxford University Press.
- Heine, Bernd. 2011. Areas of Grammaticalization and Geographical Typology. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical Typology and Linguistic Areas: With Special Reference to Africa*, pp. 44–66. Amsterdam/Philadelphia: Benjamins.
- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago/London: University of Chicago Press.
- von Heyking, Beatrix & Anne Storch. 2007. Motion in Northern Lwoo: Luwo and Boor comparative case studies. In Angelika Mietzner & Yvonne Treis (eds.), *Encoding Motion: Case Studies from Africa (in memory of Ursula Drolc)*. (Annual Publication in African Linguistics, No. 5.), pp. 89–112. Cologne: Köppe.
- Kießling, Roland. 1998. Der Pluraktionalis im Datooga. In Ines Fiedler, Catherine Griefenow-Mewis & Brigitte Reineke (eds.), *Afrikanische Sprachen im Brennpunkt der Forschung*, pp. 179–196. Cologne: Köppe.
- Kießling, Roland. 2000. Number Marking in Datooga Nouns. In Rainer Voßen, Angelika Mietzner & Antje Meißner (eds.), *Mehr als nur Worte: Afrikanistische Beiträge zum 65. Geburtstag von Franz Rottland*, pp. 349–366. Cologne: Köppe.
- Kießling, Roland. 2007a. Space and Reference in Datooga Verbal Morphosyntax. In Doris Payne & Mechthild Reh (eds.), *Advances in Nilo-Saharan Linguistics: Proceedings of the 8th Nilo-Saharan Linguistics Colloquium. Hamburg, August 22–25, 2001* (Nilo-Saharan, 22), pp. 123–142. Cologne: Köppe.
- Kießling, Roland. 2007b. The "marked nominative" in Datooga. *Journal of African Languages and Linguistics* 28.2: 149–191.
- Kießling, Roland, Derek Nurse & Maarten Mous. 2008. The Tanzanian Rift Valley area. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*, pp. 186–227. Cambridge: Cambridge University Press.
- Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the World, Seventeenth edition*. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com> (25.09.2014).
- Mietzner, Angelika. 2007. Location, motion and movement in Keyo. In Angelika Mietzner & Yvonne Treis (eds.), *Encoding Motion: Case Studies from Africa (in memory of Ursula Drolc)*. (Annual Publication in African Linguistics, 5.), pp. 145–164. Cologne: Köppe.
- Mietzner, Angelika. 2009. *Räumliche Orientierung in nilotischen Sprachen: Raumkonzepte – Direktionalität – Perspektiven*. Cologne: Köppe.
- Mietzner, Angelika. 2012. Spatial orientation in Nilotic languages and the forces of innovation. In Angelika Mietzner & Ulrike Claudi (eds.), *Directionality in Grammar and Discourse: Case Studies from Africa*, pp. 165–175. Cologne: Köppe.
- Payne, Doris L. 2013. The Challenge of Maa 'Away'. In Tim Thornes, Erik Andvik, Gwendolyn Hyslop & Joana Jansen (eds.), *Functional-Historical Approaches to Explanation: In Honor of Scott DeLancey*, pp. 259–282. Amsterdam/Philadelphia: Benjamins.
- Rottland, Franz. 1982. *Die südnilotischen Sprachen: Beschreibung, Vergleich und Rekonstruktion*. Berlin: Reimer.

- Rottland, Franz & Chet Creider. 1997. Noun classification in Southern Nilotic: Datooga. *Afrika und Übersee* 80: 71–93.
- Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace & Co.
- Svorou, Soteria. 1994. *The Grammar of Space*. Amsterdam/Philadelphia: Benjamins.
- Tomikawa, Morimichi. 1978. Family and daily life: An ethnography of the Datoga pastoralists in Mangola (1). *Senri Ethnological Studies* 1: 1–36.
- Welmers, William E. 1973. *African Language Structures*. Berkeley: University of California Press.

THE USE OF BODY PART LEXEMES IN BERTA

Susanne Neudorf

1 Introduction

Berta is a Nilo-Saharan language spoken by the Berta people, also called Beni-Shangul in the western part of Ethiopia and in eastern Sudan. Greenberg (1963) classifies Berta as part of the Chari-Nile sub-family, forming a branch of its own. Bender (1997) classifies it as a branch with no subfamilies, belonging to the satellite group of the Nilo-Saharan language phylum, together with Central Sudanic, Fur, Maban, and Kunama.

The purpose of this paper is to introduce an interesting verbal construction, namely a verb in close combination with a body part noun. After briefly showing the body part lexemes in their basic meaning and function (section 1) and as relational nouns (section 2), the rest of the paper will present cases where they are used in combination with verbs. It will be seen that they are sometimes a fixed part of the verb or have valence changing functions. They are also used to express middle voice, can function as verbal classifiers and may cause semantic changes to the verbs.

As this is a vast area of exploration, the focus is primarily on the lexeme *alú* ‘head’, occurring in various places and functions more or less grammaticalized (Heine & Kuteva 2002)¹. Other body part lexemes will be treated in less detail.

Some possible labels will be suggested in order to find predictable rules when a body part lexeme would be used and what it would then accomplish. As a body part lexeme in combination with a verb is usually a discontinuous part of a verb + particle construction², yet at the same time also has some nominal characteristics of an argument (it may change case), it will be henceforth referred to as a *partargument* (particle + argument).

Research on similar verb constructions in related Nilo-Saharan languages is not vast. There is, however, a somewhat comparable verb construction in Uduk (Killian 2013). The use of Gumuz body part lexemes as incorporated nouns is a little less similar with Berta, but still some structures are comparable (Ahland 2010). The data used here are mainly from the Mayu dialect spoken around Asosa, Ethiopia. Grammatical terms on cases and other grammatical labels are taken from and used according to Andersen’s (1995, 1993a and 1993b) works on Berta verb morphology and case marking. One exception is the modified noun marker, which is called anti-genitive in Andersen’s articles. Another exception is the verbal suffix *-a*, which he calls *perfect*. While there are examples where this suffix is used in a *perfect* sense, in my data it usually shows up in a *perfective* sense, and is thus called *perfective* in this paper.

If not mentioned otherwise, nouns are presented in their citation form, the absolutive case.

Examples are presented in the current official Berta orthography used in Ethiopia, Benishangul-Gumuz Region, which may require the following explanations:

¹ Heine and Kuteva (2002) mention four main processes in the grammaticalization of a word, which are interlinked with each other: (a) desemantization, (b) extension, (c) decategorialization and (d) erosion. Although it would be interesting to look at the different stages of grammaticalization of the Berta body part lexemes, this topic will have to be explored in another paper.

² Similar to a verbal particle or a phrasal verb construction like in English (pick+up: he picks it up) or German: *um* ‘around’ + *kommen* ‘come’ = *umkommen* ‘die, perish’

Er **kam** bei dem Unfall **um**.
he came at the accident around
‘He died/perished in the accident.’

/á, é/	vowel with high tone H
/a, e/	vowel with low tone L
/â, ê/	vowel with glide HL
/a, e, i, o, u/	the five basic vowels in Berta vary in their pronunciation, depending on their surrounding consonants and length
/q/	[ʔ] glottal stop
/tq, pq, sq, kq/	ejectives
/dq/	[d] implosive
/ñ/	[ɲ] palatal or [ŋ] velar nasal, depending on the following sound

Voiced and unvoiced stops and fricatives are in free variation in Berta. The voiced symbol is always used, except for Arabic loan words, where the distinction is marked by the respective symbols.

2 Introducing body parts and the Berta clause structure

According to Andersen (1993a, 1993b), in Berta the order of constituents in the clause is as follows:

(Topic) V (S) (IO) (DO) (Adverbial)

No matter what the topic is, the verb occurs as the second constituent in a clause. Phrases usually have a ‘head first’ construction.

Berta is a language with a marked nominative case system, which seems to be a common feature for many languages in the Nilo-Saharan family, particularly in the border areas of the Sudans, Uganda, Kenya and Ethiopia (König 2008). However, nominative is only marked when the subject is postverbal. This would be the case if, for example, another constituent is in the topic position.

Subject and object are unmarked by default (absolutive case), and the subject, if postverbal, is marked by tone (nominative case), as can be seen in example (1).

- (1) a. Maaba gagi borid. Borid gagi **máába**.
 man.ABS chase lion.ABS lion.ABS chase man.NOM
 ‘The man chases the lion.’ ‘The man chases the lion.’
- b. Maaba gagi borid. Maaba gagi **bórid**.
 man.ABS chase lion.ABS man.ABS chase lion.NOM
 ‘The man chases the lion.’ ‘The lion chases the man.’

The following example presents the word *alú* ‘head’ in its basic meaning and function:

- (2) a. Ñine gíd-i alú dqááñí.
 3.SG have-PRES head big.PARTICIP
 ‘He has a big head.’
- b. Maaba lí-a alú thá or.
 man wrap-PFCV head by/with cloth
 ‘The man wrapped his head with cloth.’
- c. Aa-kqól álu.
 1.SG-eat.PRES head.NOM
 ‘I have a headache.’

In example (2c), *alú* is in the nominative case, showing that it is the postverbal subject of the clause, but not the topic. Example (2d) is a common construction for expressing experiences and feelings, and

it can be used for all body parts, and also with other nouns (2f), with other transitive verbs (2e, 2f) and intransitive verbs (2g, 2h).

- d. Aa-kqól álu / íyu / ńdu / áre
 1.SG-eat.PRES head.NOM belly.NOM tooth.NOM eye.NOM
 ‘I have a head/ tummy/... ache.’
- e. Maabá-lé tqiikq-á íle.
 man-DEM stuff-PFCV ear.NOM
 ‘This man is deaf.’ (lit.: ‘the ear has stuffed the man’)
- f. Maabále fii húlañ.
 man-DEM beat.PRES hunger.NOM
 ‘This man is hungry.’ (lit.: ‘hunger beats this man’)
- g. Maabá-lé míl-a áre.
 man-DEM be.black-PFCV eye.NOM
 ‘This man is blind.’ (lit.: ‘the eyes have blackened this man’)
- h. Maabá-lé dáñ álu.
 man-DEM be.big.PRES head.NOM
 ‘This man has a big head.’ (lit.: ‘the head is big at this man’)

As in the above examples, many idiomatic expressions use body part lexemes, which are then filling an argument slot of the verb. Many of the partargument constructions seen below quite likely have developed from such idiomatic expressions, and it may not always be possible to clearly distinguish between the two. Other examples for idiomatic expressions with body part lexemes can be seen in (3a) and (3b). In (3a), *ile gundi* ‘ear-back’ is a compound word made up of two body part lexemes. This compound holds the place of the O-argument³ in this sentence. In (3b) the word *hu* ‘foot’ is the NP of an adpositional phrase. As will be seen in section 4, the partargument does not normally function as an argument of the verb.

- (3) a. Niñé dáá idé-kqedqé ile gundi.
 wife give.PFCV husband-her.DAB ear back
 ‘The wife ignored her husband.’ (lit.: ‘the wife gave her husband the ear-back’)
- b. Boñgór fiá abú-né thá hu.
 boy beat.PFCV father-his at foot
 ‘The boy paid respect to his father.’ (lit.: ‘The boy hit his father at the foot’)

3 Adpositional usage to indicate spatial relations

Berta, like many other languages (Heine 1995), uses body part lexemes to express spatial relations. In this use, the body part lexeme is partially grammaticalized as a relational noun, but the body part origin remains transparent. Example (4) demonstrates the typical use of Berta body part lexemes to express spatial relations:

- (4) a. thá tqarepquesa **alú**
 at table **head**
 ‘**on** the table’

³ For the sake of clarity, here and in the following we will make use of the A (agent, transitive subject), S (intransitive subject) and O (object) categories as introduced by Dixon (1994).

This construction is used with many body part lexemes:

- b. thá tqarepqesa **are**
 at table **eye/face**
 ‘above the table’
- c. thá tqarepqesa **hu**
 at table **foot**
 ‘under the table’
- d. thá buli **ndu**
 at river **mouth**
 ‘at the river/along the river’
- e. thá buli **-yú**
 at river **belly**
 ‘in the river’
- f. thá shúl **gundi**
 at house **back**
 ‘behind the house’

4 Partarguments in combinations with verb stems

From now on, the focus will be on the use of body part lexemes in combination with a verb. In these constructions the following characteristics apply: the partargument has lost much of its semantic content, and sometimes also some phonological substance. Furthermore, partarguments may have grammatical functions and can cause semantic changes to the verb.

4.1 Partarguments with a bound verb stem

A partargument in combination with a verb stem may be a complex lexeme in which the verb stem has lost its independent status. Thus, the partargument can never disappear, no matter what sentence structure the verb is used in.

The verb *liikqa alú* in example (5) is a transitive verb, although it can also occur without an overt O-argument (5b).

- (5) *liikqa alú* ‘wrap up’
- a. Maaba liikq-á oorá-kqedqe **alú**.
 man wrap-PFCV cloth-his head
 ‘The man wrapped up his clothes.’
- b. Aa-liikq-á **alú** min dqishe.
 1.SG-wrap.up-PFCV head from cold
 ‘I wrapped (myself) up because of the cold.’

No matter what the usage, the verb stem *liikqa* ‘wrap’ is always used with *alú* ‘head’ and has no meaning without it.

Another verb not existing without the partargument *alú* is found in example (6), where the verb stem has a causative suffix:

- (6) *shúkqathiña alú* ‘scare someone’
 Gadí-lé shúkqa-thiñ-óó-gí **alú**.
 child-DEM scare-CAUS-PAST-me head
 ‘This child scared me.’

So far, hardly any verb stems have been found with a fixed partargument other than *alú* ‘head’. The one exception is shown in example (7). It needs a partargument, but can have two different ones:

- (7) *bisqa -yú* ‘ignore’
 ignore -belly
- a. Ñgó bisq-óó gíñ kqál-óó-né-**yú**-á?
 you ignore-PAST thing.MOD say-PAST-he.NOM-belly-Q
 ‘Have you not listened to what he said?’

bisqa ile ‘be disobedient’
 ignore ear.NOM

- b. Attóm-á-athá maré Nasira bisq-í **ile-gú**⁴.
 Attóm-MOD-POSS.PL and Nasira be.inobedient-PRES ear.NOM-PL
 u maa agorthé-gú.
 and be.PRES thief-PL
 ‘Attom and Nasira are disobedient and (they) are thieves.’

In (7b), the S-argument is an experiencer (Payne 1997: 48–51) and the partargument is in the nominative case, as in example (2) above. The verb *bisqa* without a partargument has not been encountered in the data.

4.2 Partarguments as verbal classifiers

In some constructions, the choice of partargument is dependent on the O-argument of the verb. In this usage, the partargument serves as a classifier for the verb with respect to the O-argument. The basic meaning of the verb does not change with different partarguments.

Example (8a) is intransitive, showing the intransitive marker. *fir* ‘water’ is the S-argument of *kqóóda* ‘finish’. In (8b), *alú* ‘head’ has the function of pointing to a particular way of finishing, in the sense of disappearing. In contrast to this, the use of the partargument *are* ‘eye’ shows the completion of work resulting in a product.

- (8) *kqóóda -qí*⁵ ‘be finished’
kqóóda alú ‘finish food/eating’
kqóóda are ‘finish work’
- a. Fir kqóóda-qí thá albiír-ú.
 Water.ABS finish.PFCV-ITR at well-GEN
 ‘The water in the well is finished.’

⁴ Note that the plural marker *-gú* is actually a phrasal clitic that attaches to the last constituent of a phrase. Here it does not pluralize the partargument *ile* ‘ear’, but makes the whole verb plural to agree with the plural A-argument. Partarguments do not take plural marking.

⁵ One function of the suffix *-qí* (with a high tone) in many verb constructions is that of an intransitive marker, i.e. it can be used as a placeholder for a non-overt O-argument, or when a verb is used intransitively. However, it has other functions as well, which still need to be investigated. At this point of the research, though, we use Andersen’s terminology and call it an intransitive marker.

- b. [Aakqísqóó thabáñkqo] shambá kqóód-óó-lí thiñthiñ **alú**.
 [I washed my hands] after finish-PAST-1.SG.NOM food head
 ‘[I washed my hand] after I had finished eating.’
- c. [Lía diiñílé] ma ñgó kqóóda **are**.
 [Twine the rope] being 2.SG.ABS finish.PFCV eye
 ‘Twine the rope when you have finished (your work)!’

In the following example (9a), although not specified more closely, one can assume that the whole bamboo is peeled. In (9b), a round object is peeled, so *alú* ‘head’ can be used in place of the unspecific *-qi*. However, *alú* ‘head’ could not be used in (9a). In (9c), a long, thin object is peeled on one end only, so the partargument *ndu* ‘mouth’ is used. In (9d), a round object is peeled on one side only, so *gundi* ‘back’ is used.

- (9) *féétqa -qí*⁶ ‘peel a hard object’
féétqa alú
féétqa ndu
féétqa gundi
- a. Féétqa gaagú-qi!
 peel.IMP bamboo-PARTARG
 ‘Peel the (whole) bamboo cane!’
- b. Féétqa albambé -qi/**alú**!
 peel.IMP potato -PARTARG/-head
 ‘Peel the (whole) potato!’
- c. Féétqa-gé abandu **ndu** [áákátaba tháñ tha attawáya].
 peel.IMP-1.SG.DAB reed mouth [so I can write with it]
 ‘Sharpen a reed for me, [so I can write with it]!’
- d. Féétqa alpapayá **gundi**.
 peel.IMP papaya back
 ‘Peel the papaya on one side!’

One last example of this function of the partargument is with the verb *hérmiña* ‘roast’. When being used with *-qi* it refers to a moist mass being roasted, but when referring to an object made up of many little similar pieces, the partargument *alú* is used.

- (10) *hérmiña -qi* ‘roast meat, soaked mashed grains, moist things’
hérmiña alú ‘roast coffee, grains, nuts’
- a. Hérmiña uqûñ-i. / baaza-qi.
 roast.IMP meat-PARTARG local.beer-PARTARG
 ‘Roast the meat/beer!’
- b. Daa-gé atqáwa áá-hérmiña albún **alú** tháñ.
 give-me pan 1.SG.JUS-roast.IMP coffee head with.it
 ‘Give me a pan so I can roast coffee (beans)!’

⁶ The morpheme *-qi* (low tone) is different from *-qí* (intransitive marker, always high tone). It functions just like a partargument. According to Andersen (p.c.), it is derived from an old Berta word, meaning ‘body’. The new word for ‘body’ (*pqsuqikqe*) is unrelated.

Summarizing these examples, it is apparent that partarguments in combination with some verbs can function as verbal classifiers. Their choice is dependent on the O-argument or indicates its state or shape.

4.3 Semantic changes to the verb caused by partarguments

In the examples so far, partarguments have been either bound to a verb root or have been used to classify a verb according to its O-argument. They can also have an influence on the verb's semantics and valence. This section will show that partarguments have their own specific functions and provide additional meaning; yet these functions sometimes overlap.

4.3.1 Partargument *alú* 'head'

The bulk of the examples presented in this paper make use of this partargument. As seen, other body part lexemes are also used as partarguments, but for reasons of space we will only show a handful of examples including these other body parts.

4.3.1.1 Change of valence or clause structure

The partargument can cause a verb to become transitive, as happens with *dáara* 'burn' in the following examples. (11a) has 'fire' as the subject, in the semantic role of FORCE (Payne 1997: 49). The verb itself is intransitive (as indicated by the intransitive marker *-i*). In (11b), the verb has become transitive, the *-i* has disappeared (only the present tense *-i* is there) and *alú* needs to be present because of the overt O-argument, as seen in (11c). (11d) has a postverbal A-argument with the semantic role AGENT (1.SG). 'Grass' is the O-argument in this clause.

- (11) a. Mú dár-i thá guugu.
fire burn.PRES-ITR at opposite
'The fire is burning on the other side.'
- b. Mú dár- ñeera alú thá guugu.
fire burn-PRES grass head at opposite
'The fire is burning the grass on the other side.'
- c. *Mú dár- ñeera thá guugu.
fire burn-PRES grass at opposite
'The fire is burning the grass on the other side.'
- d. Ñeera dáar-óó-lí alú.
grass burn-PAST-1.SG.NOM head
'I burned down the grass.'

A similar use of *alú* is found in example (12), where *alú* appears in combination with the causative form of the verb. With *baná* 'hide', the transitive form does not exist without the partargument *alú*, nor can *alú* be used without the causative form. Here, the causative changes the verb from being intransitive to transitive:

- (12) *baná* 'hide oneself'
ábana alú 'hide something'
- a. Baná thále bakqá múf-i ñgó álpolis.
hide.IMP here not catch-PRES 2.SG police.NOM
'Hide (yourself) here so the policeman doesn't catch you!'
- b. Á-bana fuuda alú thá agorthé adóó-qí.
CAUS.-hide.IMP money head because thief come.PAST-ITR
'Hide the money because the thief has come!'

- c. *Á-bana* *ñalú-qa* **alú** *sha* *borid* *ad-óó-qí.*
 CAUS-hide.IMP neck.MOD-your head because lion come-PAST-ITR
 ‘Hide yourself because a lion has come!’

Examples (12b) and (12c) show that the partargument *alú* is part of the verb and not an argument. The 2.SG implied in the imperative is the A-argument, *fiuda* ‘money’ is the O-argument. The verb *baná* ‘hide’ in example (12a) by itself is implicitly reflexive and structurally intransitive. (12a) can be made explicitly reflexive by adding the reflexive pronoun *ñalu* ‘neck’, as in (12c). This reflexive pronoun then takes the place of the O-argument; *alú* is still needed.

Like in (11) before, the partargument must be used as soon as there is an overt O-argument. (13) shows another example of structural change in the verb phrase. In (13a), the verb has a clausal complement in place of the second argument. No *alú* is needed there. In (13b) however, there is the O-argument *thiñthiñ* ‘food’. Here, the presence of *alú* is dependent upon whether the complement of the verb is an O-argument or a clausal complement; *alú* is only used with O-arguments.

4.3.1.2 Totality of action

With another set of verbs, the partargument *alú* ‘head’ changes neither the lexical meaning of the verb nor the valence, but rather gives it an aspect of something *definite*, implying *completeness* or (leaning on Payne 1997: 222) *totality of action*. This means that the verb’s action is not reversible anymore or is completed to the very end. Semantically, these constructions are related to linguistic aspect; the partargument accomplishes perfective action, or telicity. As will be seen, most examples in this section use either imperative, present, or past verb forms. Once, a perfective verb form is used (example 19). This suggests that the aspectual component of the partargument can actually be used with any verb form; still the exact relationship between the use of the different tenses/aspects and the use of *alú* as a marker of totality of action needs to be investigated further.

- (14) *kqátqasa* ‘immerse, penetrate’
kqátqasa alú ‘immerse completely’
- a. *Kqátqasa* *almusmár* *thá* *sqisqía-yú-é-qí.*
 immerse.IMP nail at wood-belly-DIR-ITR
 ‘Hammer the nail into the wood!’ (how deep is not specified)
- b. *Kqátqasa* *almusmár* **alú** *thá* *sqisqía-yú-é-qí.*
 immerse.IMP nail head at wood-belly-DIR-ITR
 ‘Hammer the nail into the wood completely!’

The following example may at first sight suggest a directional function (a motion away) of *alú*. However, example (15c) shows that even with an obvious direction, *alú* is not present in the sentence. The reason for this is that the ball is wanted back and is not thrown away for good.

- (15) *dqaakqa* ‘throw’
dqaakqa alú ‘throw away’
- a. *Dqaakqa* *hodia!*
 throw.IMP rod
 ‘Throw the rod!’ (out of your hand, you can pick it up later)
- b. *Aa-dqaakq-óó* *basqaro* **alú** *aañá* *squruñga-áñi!*
 I.SG.-throw-PAST stick head place.MOD long-place.NOM
 ‘I threw my stick away in a far place.’ (because I do not need it again)

- c. Dqaakqo-gé algúra!
 throw.CP.IMP-1.SG.DAT ball
 ‘Throw the ball to me!’

In example (16a) and (16b), *alú* is used similarly. It adds a component of completeness or telicity (Dik 1989). In (16c), this component is missing.

- (16) *gagá* ‘chase’
gagá alú ‘chase away’
- a. Fámiilí gag-í gal **alú**.
 people chase-PRES dog head
 ‘People chase dogs away.’
- b. Maabá ma zozo gagí-lá **alú** mithil gali.
 person.MOD being lazy chase-PASS head like dog
 ‘A lazy person is chased away like a dog.’
- c. Gali gag-í mia.
 dog chase-PRES goat
 ‘The dog chases a goat.’

In example (17a), the *gíndidqí* ‘refuse’ is dumped for good, while in (17b) it is just outside. So in (17b), the notion of totality of the action is missing.

- (17) *feetha* ‘sweep’
feetha alú ‘sweep away’
- a. Feetha gíndidqí **alú** thá fuulú-é-qí.
 sweep.IMP refuse head at hole.GEN-DIR-ITR
 ‘Sweep the refuse into the hole!’ (so it is gone completely)
- b. Feetha gíndidqí thuuth-é!
 sweep.IMP refuse outside-DIR
 ‘Sweep the garbage which is outside!’ (it is still there, then)

In example (18a) the giving is a general action, e.g. giving for any reason, while in (18b) the person may give something for some time, but then wants it back. *alú* is used to state that the given object will be out of the agent’s control. That *daa alú* is also used for giving a daughter away in marriage confirms this function. Once the daughter is given to someone, she is no longer the father’s responsibility and belongs to the husband.

- (18) *daa* ‘give’
daa alú ‘give away for good’
- a. Walá daa gíñmáñ.
 NEG give.PRES.3.SG something
 ‘He does not give to anybody.’ (neither for borrowing, nor for keeping)
- b. Walá daa gíñmáñ **alú**.
 NEG give.PRES.3.SG something head
 ‘He does not give anything away.’ (for the other person to keep)

In example (19a) the water has been poured out of the bottle (or container), but it is still preserved. The process is reversible. In (19b), the water is not retrievable.

- (19) *pquudqa* ‘pour’
pquudqa alú ‘pour out’
- a. Pquudqa fir thálé!
 pour.IMP water there
 ‘Pour the water there!’ (in a container, because we need it again)
- b. Niñé pquudq-á fir **alú** thá sqakqa-yúéqí.
 woman pour-PFCV water head at ground-into
 The woman has poured water on the ground.

The words for ‘buy’ and ‘sell’ also seem to fit this pattern:

- (20) *shibila* ‘buy’
shibila alú ‘sell’
- a. Shibil-o-gé almánga holoñoníñ.
 buy-CP.IMP-I.SG.DAB mango two
 ‘Buy two mangoes for me!’
- b. Aa-shibil kqosh **alú** tha súúgú.
 I.SG.-buy.PRES oil head at market
 ‘I sell oil on the market.’

On first sight, the partargument *alú* seems to give a new meaning to the verb. But that can be explained by translating *shibila* with ‘trade’; *alú* then changes the direction of the trading. The traded item is either in one’s control (*shibila* ‘trade to get something/buy’) or out of one’s area of control (*shibila alú* ‘trade something away/sell’). Thus this example fits into the pattern of *totality of action* seen in earlier examples.

In the following example (21), it may be interesting to compare (21c) with the sentences in section 4.4. Here, the sentence has only one overt argument, in contrast to (21a) and (21b). The partargument stays in the absolutive case and does not change to nominative. The reason is that here it is the AGENT of the verb, and in the examples of paragraph 4.4 the only overt argument of the clause is the O-argument. Because of this, (21c) is not a middle construction.

- (21) *buura* ‘collect, gather’
buura alú ‘collect, gather’
- a. Buura fuuda ñgó fээdqa niñé tháñ!
 gather.IMP money 2.SG.JUS seek.IMP woman with.it
 ‘Save money, so you may get married.’
- b. Buura fuudá-qa **alú**!
 collect.IMP money.MOD-your head
 ‘Collect your money!’ (from the different places)
- c. [Aahaalóo añiyú tha] maabí bur **alú** ma añsqaré.
 [I heard that] people gather.PRES head in the morning
 ‘[I heard that] people will gather/meet tomorrow.’

The following three sections give a brief overview over the functions of some other partarguments. This is by no means exhaustive, but rather meant to give a little impression of the complexity of the matter.

4.3.2 Partargument *iyú* ‘belly’

4.3.2.1 Change of valence

There are not many instances in which the partargument *iyú* ‘belly’ changes the valence of the verb. In the following example (22), the verb changes from having one argument slot (S) to two argument slots. However, the O-argument is not in the absolutive case, but is marked for DAB case. This is a rather exceptional case, as other examples of this kind have not been encountered.

- (22) *sqikqá* ‘happen, fall’
sqikqá -yú ‘happen to someone’
- a. *Áné sqikq-a baró shambá kqal-a na abbá.*
 TOP happen-PFCV then like say-PFCV by my.father
 ‘It then happened just like my father said.’
- b. *Ñgó gadarí ñgó roothu-qi shíñ thá maabí-le*
 you can you discuss-PARTARG then with people-ASSOC
thá roothá sqikq-í maabí-é-yú.
 at matter.MOD happen-PRES people-DAB-belly
 ‘You can discuss with other people the things happening to them.’

4.3.2.2 Change of meaning

With stative verbs

A much more widespread use of *iyú* ‘belly’ is with stative verbs that are semantically intransitive and describe a state rather than a process or action. These are changed slightly in meaning by the addition of the partargument, which is always in nominative case (like in the experiencing structure mentioned in section 2, example (2d)).

- (23) *dqáñí* ‘be big’ ‘be big, old’
dqáñiyu ‘be big belly.NOM’ ‘be wide’
- a. *Maaba dáñ-i*
 man be.old.PRES-ITR
 ‘The man is old.’
- Thoñor dqáñ-íyu*
 road be.big.PRES-belly.NOM
 ‘The road is wide.’
- pqíllí* ‘be angry’ ‘be angry’
pqíllíyu ‘be angry belly.NOM’ ‘be deep’
- b. *Maaba pqíl-i*
 man be.angry.PRES-ITR
 ‘The man is angry.’
- Gaama pqíl-íyu*
 road be.angry.PRES-belly.NOM
 ‘The lake is deep.’

kqosqí ‘be small’ ‘be small’
kqosqíyu ‘be small belly.NOM’ ‘be narrow’

- c. Gadi *kqosq-í*
 child be.small.PRES-ITR
 ‘The child is small.’

Thoñor *kqosq-íyu*
 road be.small.PRES-belly.NOM
 ‘The road is narrow.’

It is tempting to analyze this construction as an external possession, with the partargument as the possessed noun expressing the (often metaphorical) location of the state. Thus ‘belly’ could be the metaphorical location of the S-argument. However, while it is only a small leap of imagination that sees the belly of the road as big (= wide, see (23a)) or small (= narrow, see (23c)), for (23b) it is not quite so convincing to assume that the belly of the lake is angry (= deep).

Example (23d) has no valence change, but the intransitive base word has a suffix that looks like the reciprocal marker *-u* (glossed as such in the following examples). This ending together with the partargument results in a semantic change. The subject of *miilu íyu* needs to be a specific thing (23e).

d. *miila* ‘be black’ ‘be black’
miil-u íyu ‘be black.REC belly.NOM’ ‘be dark’

- e. Náñ *miil-u-ga* *shor íyu* *shókqóñ* *shúgo?*
 what be.black-REC-for sky belly.NOM today this.way
 ‘Why is the sky so dark today?’

With transitive verbs in reciprocal form

The partargument *-yú* plus a verb in reciprocal form gives the meaning of something happening specifically between two things or resulting in two things.

(24) *maala* ‘change’ change
maalu -yú ‘change-belly’ exchange two things

- a. Maala *oori-qa.*
 change.IMP clothes-your
 ‘Change your clothes!’
- b. Áñ *maal-u* *alradiá-ñkqa-yú!*
 we.JUS change-REC radio.MOD-our-belly
 ‘Let’s exchange our radios!’

similar:

pqesha ‘break’ ‘break off, break in many tiny things’
pqeshu -yú ‘break-belly’ ‘break in two pieces’

sqoma ‘tear’ ‘pinch, tear off’
sqomu -yú ‘tear-belly’ ‘divide in two’

4.3.2.3 Totality of action

Another slight semantic change is found with *-yú* in example (25). Note that in both (25a) and (25b) the verb is transitive. The addition of *-yú* adds an endpoint to the verb, like *alú* in section 4.3.1.2:

- (25) *goora* ‘surpass’
goora -yú ‘win against’
- a. Moháméd gor geedí il-í-gú kqíllíñ.
 Mohamed surpass.PRES child.PL play-PRES-PL all
 ‘Of all the children who are playing, Mohamed plays best.’
 (Lit: ‘M. surpasses all playing children’)
- b. Tha algúra-yú hatháñ goor-óó maabí-yú gíidi.
 At ball-belly 1.PL win-PAST people-belly yesterday
 ‘Yesterday, we won the football game against the other team.’

4.3.3 Partargument *ndu* ‘mouth’

4.3.3.1 Totality of action

Following is an example with the partargument *ndu* ‘mouth’ where it signals totality of action. It is mostly used when the O-argument is human, or at least animate.

- (26) *fentqetqa* ‘destroy’
fentqetqa ndu ‘destroy completely’
- Ñine fentqetqá maabí **ndu**.
 3.SG eradicate.PFCV people mouth
 ‘He wiped out the people completely.’ (no one was left)

Without *ndu*, someone might have been left out in the destruction.

In example (27), the reciprocal verb stem in connection with the partargument *ndu* renders again the idea of totality of action:

- (27) *kqithá* ‘cut’ (basic meaning)
kqithu ‘cut all communication’
kqithu ndu ‘fight each other till death’
- a. Rooth-á á-kqith-u maab-í m-maa-né.
 matter.MOD CAUS-cut-REC person-PL TOP-be.PRES.EMPH-3.SG
 ‘This matter will cut all communication between people.’
- b. Rooth-á á-kqith-u maab-í **ndu** m-aa-né.
 matter.MOD CAUS-cut-REC person-PL mouth TOP-be.PRES.EMPH-3.SG
 ‘This matter will make people fight till death.’

4.3.4 Partargument *gundi* ‘back’

The last partargument mentioned here is *gundi* ‘back’. It is found mostly in combination with stative verbs. It may change the meaning of a verb stem without changing its valence.

4.3.4.1 Semantic change

Like other partarguments, *gundi* ‘back’ may classify a verb with a specific meaning derived from the O-argument. It generally refers to large round things or areas, or to one side of something round.

The following example (28) shows a transitive verb; the A-argument is, however, rather an experiencer, and the verb describes a state rather than an action or process. The partargument is used in the nominative case, like in the experiencer construction mentioned before in example (2). The verb takes no other argument, and the subject of the verb is considered an S-argument.

- (28) *sholá* ‘know not, be ignorant’
*shól gúndi*⁷ ‘know.not back.NOM’ ‘be insufficient’
- a. Aa-shól algiráya shambá gárra-lá-ne.
 I-know.not.PRES school how read.PRES-PASS-3.SG
 ‘I do not know the teaching how to read’. (= I do not know how to read)
- b. Assúkar-á-lé shól **gúndi**
 sugar-MOD-DEM know.not.PRES back.NOM
 ‘This sugar is not enough.’

The verb *miila* ‘be black’ is stative, and the partargument is in the nominative case.

- (29) *miila* ‘be black’
miilá gúndi ‘be.black back.NOM’ ‘set (for sun, moon)’
- a. Sqakqá-lé mil-í.
 soil.MOD-dem be.black.PRES-ITR
 ‘This soil is black.’
- b. Zigi ma miilá **gúndi** añ miilu-qí.
 moon being be.black.PFCV back.NOM place be.dark.PRES-ITR
 ‘When the moon sets, it gets dark.’

In the following two examples (with transitive verb stems), the partargument, if used, is in the absolutive case. The complex verb in example (30) has different meanings depending on the context.

- (30) *kqithá* ‘cut’
kqithá gundi ‘cut back’ - surround, make fence/barrier around something
 - gossip
- a. Kqithá diini thá hangirá pqílí ndu.
 cut.IMP rope with knife-mod be.sharp.PARTICIP
 ‘Cut the rope with a sharp knife!’
- b. Kqithá-tha shúl-kqedqe **gundi**.
 cut.IMP-PL house.MOD-his back
 ‘Surround his house!’
- c. Niñé-lé kqith-óó-lí **gundi**.
 woman.MOD-DEM cut-PAST-1.SG.NOM back
 ‘I gossiped about this woman.’

As in the previous example, the complex verb in example (31) has two meanings, depending on the context. One meaning refers to an area around something (b). The other renders an aspect of forcefulness to the action (c). Note that the same action of taking away without the forcefulness would be stated by a different verb (*fará* ‘take’).

⁷ The partargument in the nominative case gives the phrasal verb a stative meaning. In these cases, the present tense form is used, instead of the imperative form, the usual citation form of a verb.

- (31) *fia* 'beat'
fia gundi 'beat back' - cut around the house with machete
 - take by force
- a. Maaba *fī-a* *gadí-kqedqe*.
 man beat-PFCV child.MOD-his
 'The man beat his child.'
- b. Maaba *fī-a* *shúl-kqedqe* ***gundi***.
 man beat-PFCV house.MOD-his back
 'The man cleared (the grass) around his house.' (with a machete)
- c. Maaba *fī-a* *alkitabá* *Fátna* ***gundi***.
 man beat-PFCV book.MOD Fatna back
 'The man forcefully took Fatna's book.' (to claim it his own)

4.3.4.2 Deontic mood or totality of action?

The partargument *gundi* 'back' in combination with some verb stems in imperative form results in a meaning of necessity or obligation. The same complex verb in the indicative does not render the urgency of the action as much as its completeness. It may still have some similarity with the notion of totality of action.

Example (32a) can best be analyzed as having a deontic mood, while in (32b) a totality of action interpretation suggests itself.

- (32) *ágena* 'send'
ágena gundi 'must send'
- a. Ñgó *ágen-a* *gadí* ***gundi*** *thá* *alگیرáyá-yú-é-qí*.
 you.JUS send-IMP child back at school.MOD-belly-DIR-ITR
 'You have to send your child to school.'
- b. Maabále *walá* *ágen-i* *gadí-kqedqe* ***gundi*** *tha* *almádarasáyúéqí*.
 man-MOD-DEM not send-PRES child.MOD-his back to school
 'The man does not send his child to school.' (the child never goes at all)

4.4 Partarguments as middle voice markers

Middle voice or middle constructions are neither active (an agent acting upon a patient) nor passive (a patient being acted upon by an agent), but look at a scene rather as a process than as an action (Payne 1997: 216–218). It is a valence decreasing operation, with the patient of a transitive verb becoming the subject of an intransitive clause.

Berta frequently uses partarguments to render middle voice by changing the case of the partargument from absolutive to nominative. Syntactically, it looks like the experiencing constructions of example (2), but the body part lexeme here is a partargument, without any apparent metaphorical locational meaning.

The middle voice is distinct from the Berta passive, which looks as follows:

- (33) a. Ñine *hásqúl-á* *albááka*.
 he take-PFCV container
 'He took the container.'
- b. Maaba *dqááñi* *hásqúl-á-lá* *tha* *alhakim-ú-é-qí*.
 man old.PARTICIP take-PFCV-PASS at doctor-GEN-DIR-ITR
 'The old man was taken to the doctor.'

Very few verbs are found so far which are subjected to both passive and middle voice.

In example (34) the partargument is a fixed part of a complex verb, and the verb stem has no meaning without it. In example (34b), the partargument *alú* is in nominative case, and the grammatical valence of the verb is reduced from two to one.

(34) *dámala alú* ‘cheat’

- a. Náñ dámal-óó-gá-ñó-gi alú?
 What cheat-PAST-for-you.NOM-me head
 ‘Why have you deceived me?’
- b. Gadí-lo da-damal-iñ **álu** kúllu yom.
 child-DEM REP.cheat.REP.PFCV head.NOM at.market always
 ‘That child is always tricked at the market.’

In example (35), *alú* is not a fixed part of the verb stem, but is obligatory as soon as an O-argument is expressed. (Compare 4.3.1.1, example (11)). In (35b), O becomes S, and the partargument is in the nominative case.

(35) *dáára* ‘burn’

- a. Ñeera dáar-óó-lí **alú**.
 grass burn-PAST-1.SG.NOM head
 ‘I burned down the grass.’
- b. Ñeera dáar-a **álu**.
 grass burn-PFCV head.NOM
 ‘The grass is burned.’

The following example shows the partargument *alú* giving an aspect of totality of action (see 4.3.1.2, example (19)).

(36) *pquudqa* ‘pour’

- a. Niñé pquudq-á fir **alú** thá sqakqa-yúéqí.
 woman pour-PFCV water head at ground-into
 ‘The woman has poured water on the ground.’
- b. Fir pquudq-á **álu**.
 water pour-PFCV head.NOM
 ‘The water is spilled.’

Another verb using partarguments to create middle voice is *fia* ‘beat’. Each partargument combined with this verb stem will change its meaning. The examples given in (37) vary in their meaning only in the degree of how much the ‘erasing’ has progressed, i.e. if the process is completed (*alú*) or not yet completed (*are*). Both partarguments in nominative case create a middle voice meaning. *fia are* ‘beat eye’ with the partargument in absolutive case renders a different meaning.

- (37) *fia* ‘beat’
fia alú ‘beat head’ ‘erase completely, arrive’
fia álu ‘beat head.NOM’ ‘be erased completely’
fia áre ‘beat eye.NOM’ ‘be erased (but still recognizable)’
fia are ‘beat eye’ ‘add more to, make smooth’

- a. Thoñor ñia álu.
road beat-PFCV head.NOM
'The road is out of use.' (does not exist anymore)
- b. Thoñor ñia áre.
road beat-PFCV eye.NOM
'The road is out of use.' (but still visible)

In summary, the partargument used here changes the focus of the scene from the verb's action to the result of a process. As the translations suggest, all these examples have a perfect meaning, but use a perfective verb form with a partargument in nominative case. It needs to be further studied if middle voice is restricted to verbs with partarguments or if there are other ways to express it. Other rewarding subjects for further research are the relationship of middle voice to the Berta passive construction, and the question of whether middle voice can be used with other aspectual forms than the perfective.

5 Summary

A partargument is a grammaticalized form of a body part lexeme, which together with a verb root forms a complex verb. Partarguments have differing functions, depending on the verbs they combine with, and on the syntactic relationships. Sometimes a hint of the original meaning of the body part lexeme can still be deduced, but more often, the partargument is devoid of its original lexical meaning and function.

When used to classify a verb in respect to its O-argument, a general, rather vague correlation can be noticed between the choice of the partargument and the form of an O-argument of a verb. So for example, *alú* 'head' tends to refer to a round object or many small things (e.g. coffee beans). Also it may refer to the complete disappearance of the O-argument, in contrast to *are* 'eye', which points to the visibility of the O-argument. *gundi* 'back' also hints at something round or surrounding, but also to one side of some (mostly round) thing. If the O-argument of a verb is long and thin, either as a patient (e.g. leak), or as a result (e.g. a path), then the partargument *ndu* 'mouth' may be used. *ndu* also can refer to an opening or a side of something, as well as to human O-arguments. *iyú* 'belly' usually refers to an internal location or quality of the S-argument. In this function, the partargument construction is somewhat comparable to Gumuz (Ahland 2012: 186–192).

The choice and use of a partargument can also be completely independent from the O-argument. Then it has either a structural function to change the number of arguments of a verb, or it changes its meaning, or it changes its aspect. In one function, the partargument takes the nominative instead of the absolutive case, in order to render a middle voice construction, reducing the valence of the verb. This usage of the partargument is very similar to the complex verbal predicates found in Uduk (Killian 2013).

While both function and meaning of each partargument can be described in terms of general tendencies, the meaning of each partargument with any verb needs to be learned and can hardly be predicted.

It may well be worth to have a closer look and compare the partargument to coverb constructions found in other languages, as well as to phrasal verbs in English or German. Comparisons with these constructions may be helpful to better understand the multiple functions of the partarguments in Berta. It is also apparent that the use of partarguments influences the semantic transitivity of a verb in Berta, and a better understanding of transitivity in this language may shed new light on this subject.

Appendix

I. Some verbs combined with various partarguments

fia ‘beat’

<i>fia</i>	‘beat (hit several times)’	
<i>fia doño</i>	‘beat neck’	‘kill’
<i>fia alú</i>	‘beat head’	‘arrive’
		‘erase completely, not visible anymore’
<i>fia</i> ⁸ <i>álu</i>	‘beat head.NOM’	‘be erased (completely)’
<i>fia áre</i>	‘beat eye.NOM’	‘be erased (but still recognizable)’
		‘be drunk’
<i>fia are</i>	‘beat eye’	‘add more to’
		‘make smooth’
<i>fia ndu</i>	‘beat mouth’	‘blow out’
		‘make a loop’
<i>fia gundi</i>	‘beat back’	‘withhold, take by force’ (not give what is legally someone else’s)
		‘cut around a house or other place, by machete’
<i>fia hu</i>	‘beat foot’	‘beat a foot’ (no indirect meaning)

kqithá ‘cut’

<i>kqithá</i>	‘cut’	‘cut’ (repeated movement)
		‘harvest’ (cut and gather)
<i>kqithá alú</i>	‘cut head’	‘cut down’ (e.g. grass), ‘reduce’ (for liquids)
<i>kqithá gundi</i>	‘cut back’	‘gossip’
		‘make a barrier around an area, surround’
<i>kqithá ñalu</i>	‘cut nape’	‘slaughter’
<i>kqithá hu</i>	‘cut foot/name’	‘guess’
		‘cut grass around e.g. a tree, by machete’
<i>kqithá ndu</i>	‘cut mouth’	‘cut something long, thin’ (result, e.g. cut a path through the grass)
		‘cut something long, thin’ (e.g. onion leak)
		‘cut the top off to make an opening’ (e.g. pumpkin)
<i>kqithá ile</i>	‘cut ear’	‘cut the ear’ (no other, figurative meaning)
<i>kqitha iyú</i>	‘cut belly.NOM’	‘have no hope for something, resign’
		‘be out of breath after working hard’
<i>kqithá iyú</i>	‘cut belly’	‘pour liquid into something’
<i>kqithu ndu</i>	‘cut.REC mouth’	‘fight each other till death’
<i>kqithu are</i>	‘cut.REC eye’	‘cross each other on the road’
<i>walá kqithu áre</i>	‘not cut.REC eye.NOM’	‘not to be trusted, be cunning’

maada ‘know’

<i>maada</i>	‘know’	‘know’
<i>maada -qi</i>	‘know’	‘see, look’
<i>maada -yú</i>	‘know belly’	‘think about/ make up one’s mind about’
<i>maadá iyú</i>	‘know belly.NOM’	‘be fitting’ (e.g. clothes)
<i>maada gundi</i>	‘know back’	‘know about’
<i>maada hu</i>	‘know foot/name’	‘be sure about’
<i>maada alú</i>	‘know head’	‘know exactly about’ (restricted to the knowledge of climbing a tree or crossing a river)

⁸ The imperative citation form for stative verbs would be unnatural. Therefore the perfective verb form will be used for all 5 verbs in this section

daa ‘give’

<i>daa</i>	‘give’	
<i>daa alú</i>	‘give head’	‘give away to someone’
<i>daa ndu</i>	‘give mouth’	‘give instructions, give command’
<i>daa are</i>	‘give eye’	‘not intervene, just look upon a problem/struggle’
<i>daa(gá) ile</i>	‘give ear to’	‘obey’ (no partargument, but O-argument, this is an idiomatic expression)
<i>daa(gá) ile gundi</i>	‘give ear back to’	‘ignore’ (opposite of the above)
<i>daa gundi</i>	‘give back’	‘send someone’
<i>daa hu</i>	‘give foot/name’	‘give a name’ (not a partargument, but O-argument)

adá ‘go’

<i>adá hu</i>	‘go foot’	‘go together with’
<i>ada álu</i>	‘go head.NOM’	‘tilt to one side, be not straight anymore’
<i>ada gúndi</i>	‘go back.NOM’	‘become bigger, stronger, overwhelming’
<i>adá ile</i>	‘go ear’	‘be already informed’

II. List of bodypart lexemes

Most often used lexemes:		Rarely grammaticalized lexemes:		Non-body part lexemes with similar functions:	
<i>alú</i>	head	<i>hu</i>	foot / leg / name	<i>añ</i>	place
<i>gundi</i>	back	<i>are</i>	eye / face	<i>-qi/-i</i>	no meaning on its own, may have derived from an old word for ‘body’
<i>ndu</i>	mouth	<i>ile</i>	ear		
<i>iyú</i>	belly	<i>qĩñ/-iĩñ</i>	buttocks		
<i>-yú</i>		<i>ñalú</i>	nape		
		<i>bolo</i>	windpipe		
		<i>doño</i>	throat		

Abbreviations

ABS	absolutive case (default, unmarked case)
ASSOC	associative marker
CAUS	causative
COP	copula
CP	centripetal
DAB	dative/allative/benefactive case
DEM	demonstrative
DIR	directional marker
EMPH	emphatic form
GEN	genitive case
IMP	imperative
ITR	intransitive marker
JUS	jussive form of a pronoun, accompanies an imperative verb
MOD	modified noun marker (marks the head of a noun phrase, whenever the noun is modified in any way. Andersen calls this the AG [anti-genitive] case)
NEG	negation particle
NOM	nominative case (marked case, used for post verbal subjects)
NP	noun phrase
PARTARG	partargument

PARTICIP	participle of a stative verb, used like an adjective
PASS	passive
PAST	past tense
PFCV	perfective aspect of the verb
PL	plural
PRES	present tense (non-past, can also have future meaning)
Q	question particle
REC	reciprocal verb inflection
RELPRON	relative pronoun
REP	repetitive
SG	singular
TOP	topic reference
VP	verb phrase

References

- Ahland, Colleen. 2010. Noun incorporation and predicate classifiers in Gumuz. *Journal of African Languages and Linguistics* 31: 159–203.
- Ahland, Colleen. 2012. A Grammar of Northern and Southern Gumuz. Ph.D. thesis, University of Oregon.
- Andersen, Torben. 1993a. Verbal roots and verbal inflection in Berta. *African Languages and Cultures* 6.2: 97–119.
- Andersen, Torben. 1993b. Aspects of Berta phonology. *Afrika und Übersee* 76: 41–80.
- Andersen, Torben. 1995. Absolutive and nominative in Berta. In Robert Nicolai & Franz Rottland (eds.), *Proceedings of the Fifth Nilo-Saharan Linguistics Colloquium. Nice, August 24–29 1992*, (Nilo-Saharan, 10), pp. 39–69. Cologne: Köppe.
- Bender, M. Lionel. 1997. *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom.
- Dik, Simon C. 1989. *The Theory of Functional Grammar, Volume I*. Dordrecht: Foris.
- Dixon, Robert M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Greenberg, Joseph. 1963. *The Languages of Africa*. Bloomington: Indiana University Press.
- Heine, Bernd. 1995. Conceptual grammaticalization and prediction. In John R. Taylor (ed.), *Language and the Cognitive Construal of the World*, pp. 119–135. Berlin: DeGruyter.
- Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Killian, Don. 2013. *Complex Verbal Predicates in Uduk*. Paper presented at the Colloquium of Nilo-Saharan Languages and Linguistics 2013, Cologne.
- König, Christa. 2008. The marked-nominative languages of eastern Africa. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*, pp. 251–271. Cambridge: Cambridge University Press.
- Payne, Thomas E. 1997. *Describing Morphosyntax*. Cambridge: Cambridge University Press.

THE AMA DUAL SUFFIX: AN INTERNAL RECONSTRUCTION¹

Russell Norton

1 Summary of the Ama dual

Ama [nyi], also known as “Nyimang” in the literature after Stevenson (1938, 1956–57), is a North Eastern Sudanic language spoken in villages of the north-west Nuba Mountains of Sudan, and in urban diaspora communities in Khartoum and other towns (Rilly 2010). The Ama verb system has a rare dual suffix that occurs with dual participants in a wide range of positions in the clause (Stevenson, Rottland & Jakobi 1992; Norton 2012). This paper considers three possible historical scenarios for the origin of the Ama dual suffix: incorporation, borrowing, or semantic shift. It is argued that the dual arose by semantic shift in an old reciprocal suffix. This internal reconstruction is supported by a comparison of the form of the Ama dual suffix with reciprocal suffixes in other Eastern Sudanic languages.

As described in Norton (2012), the Ama dual suffix is neither a subject marker nor an object marker, rather it occurs in clauses with dual participants in a wide range of positions – subject, direct object, indirect object, copula complement, and reciprocal constructions. The suffix has the functions of dual agreement, dual reference, and dual reciprocal:

Dual agreement

Agreement with a noun phrase containing the numeral ‘two’

- (1) nórđú ārbā bā wòóŋ n-ēn
book two DECL 1SG.POSS be-DU
‘The two books are mine.’

Agreement with a complex argument of two nouns

- (2) fōrāŋ-ō mòr bā ir gīl-i n-ēn
rabbit-GEN friend DECL elephant dog-DAT be-DU
‘The friends of Rabbit were Elephant and Dog.’

In Ama, two nouns in a complex argument are not linked by a conjunction, instead the second noun is in dative case used with a comitative function, expressing accompaniment to the first noun. So example (2) reads literally ‘the friends of Rabbit were Elephant with Dog’.

Dual reference

Dual reference of a personal pronoun

- (3) ē ní đū ƒā kèérn-èi kó ƒwōr-ēn
and 2PL TOPIC now peace-ABL CFOC:DEON come-DU
‘You two must now only come in peace.’

Dual reference of a rational noun

- (4) ài bā dōrēŋ-ì wáđà áǰínāsō-n-ēn
1SG DECL children-DAT story tell-APPL-DU
‘I told a story to two children.’

¹ I am grateful to the Colloquium audience and to Andrew Spencer, Noura Ramli, and John Roberts for their comments.

Unlike dual agreement, dual reference can be assigned only for rational participants. In (3) the subject pronoun is plural, as the Ama pronoun system distinguishes singular and plural for 1st, 2nd, and 3rd person, but if the plural pronoun is used when there are specifically two referents, then the dual suffix will occur on the verb. Most nouns in Ama are not marked for number, but if a rational noun is used when there are specifically two referents, then likewise the dual suffix will occur on the verb, as in (4).

Dual reciprocal

Dual reciprocal transitive clause

- (5) kér bā kwēj-ō kā bāl̥j̥ŋ-ēn-àg
 woman DECL man-ACC VFOC love-DU-NACT
 ‘The woman really loves the man and vice versa.’

Dual reciprocal stative clause

- (6) ē ir ḡū fōrāŋ-ō mòr n-ēn
 and elephant TOPIC rabbit-GEN friend be-DU
 ‘Elephant was a friend of Rabbit and vice versa.’

In transitive reciprocal constructions, the participants of the reciprocal event occur in subject and object positions, and the verb has a nonactive suffix. If there are precisely two participants in the reciprocal event, as in (5), then the dual suffix is also present. In stative clauses, when two participants are stated to be in a reciprocal relationship, the verb is marked as dual, as in (6). The nonactive suffix does not occur in stative clauses.

The Ama dual functions over the whole clause, both in the sense that it can mark dual arguments in a variety of positions in the clause, and in the sense that it can mark two paired participants in different positions of a reciprocal clause. Since it functions over the whole clause it is a category of verbal number. But the Ama dual is an unusual verbal number marker – universally, comparatively, and paradigmatically.

Universally, dual verbal number is *very rare* in known languages, especially as a productive inflection available to all verbs of a language (Corbett 2000).

Comparatively, the verbal dual in Ama is *unique in Nilo-Saharan*, although in Afitti (the language closest to Ama, in the Nyima group within North Eastern Sudanic) it is reported that verbal dual is zero-marked by omission of the verbal plural marker (de Voogt 2011). Pronominal dual markers occur on verbs in some other Nilo-Saharan languages (Otero, this volume), but these are not the same as verbal dual markers that function over the whole clause.

Paradigmatically, the Ama verbal dual is unusual as a *dual without a plural counterpart*. According to Corbett (2000), nouns and pronouns of the world’s languages always comply with Greenberg’s universal 34 that if a language has a dual then it also has a plural, but Norton (2012) argues that universal 34 does not hold for verbal number because Ama has a verbal dual but does not have a true verbal plural. The argument that Ama has no verbal plural is a subtle one, because there is a putative plural of Ama verbs reported in Stevenson (1938, 1957) and Stevenson, Rottland and Jakobi (1992), and it does pattern in the expected way for intransitive verbs where it marks subjects of three or more referents (Norton 2012: 79). On transitive verbs, however, it becomes clear that “plural” marking is more correctly described as a verbal distributive rather than a verbal plural, because it is not obligatory but when added it distributes the event either over several object referents or over a series of sub-events. This is not the same as a verbal plural, which would mark an event that either has several object referents or is repeated several times. The following examples (7) to (9) show distributive marking on sentences and the difference of meaning from verbal plural (Norton 2012: 77–78):

- (7) a. ài bā ḡōlār ḡār mōl tí
 1SG DECL dollar 20x 5 take
 ‘I took 100 dollars.’

- b. àì bā ḍṣlār ṭār mōl bùò
 1SG DECL dollar 20x 5 take:DISTR
 ‘I took 100 dollars in a series of instalments.’ (*‘I took 100 dollars many times over.’)
- (8) a. əŋjí bā dṣrēŋ mōl dē-íḍ-í
 1PL DECL children 5 hit
 ‘We beat 5 children.’
- b. əŋjí bā dṣrēŋ mōl dē-íḍ-í
 1PL DECL children 5 hit-DISTR-TH
 ‘We beat each of 5 children.’, ‘We beat until we had beaten 5 children.’
 (*‘We beat 5 children many times over.’)
- (9) a. àì bā kī=ḍṣ mōdā
 1SG DECL thing=ACC tie
 ‘I tied the thing or things up.’
- b. àì bā kī=ḍṣ mōd-úḍ-ā-r
 1SG DECL thing=ACC tie-DISTR-TH-DISTR
 ‘I tied until I had tied the thing up.’, ‘I tied each thing up.’, ‘I tied until I had tied each thing up.’ (*‘I tied the thing up many times over.’, *‘I tied many things up.’)

There are several other details in Norton (2012) supporting the conclusion that the Ama distributive does not pattern as a successor to the verbal dual in a scale of increasing number. So when we consider the origin of the dual suffix, our reconstruction should account for the development of a verbal dual without the development of a verbal plural counterpart.²

2 Possible origins

We may wonder how there came to be a verbal dual suffix in Ama, given the rarity of dual verbal number in the world’s languages, its absence in other Nilo-Saharan languages, and its presence in Ama without a true verbal plural counterpart. We will consider three possible historical scenarios for the origin of the dual suffix: *incorporation* of a dual word, *borrowing* of a dual suffix, or *semantic shift* in an existing verb suffix.

2.1 Incorporation

Candidates for incorporation would be dual pronouns or quantifiers. Pronoun incorporation on the verb is a widespread feature in language (e.g. Schröder 2012), and the incorporation of a dual pronoun form could offer a natural account of the suffix’s dual agreement and dual reference functions, as the suffix can stand pronominally for a dual noun phrase as in (10) below, as well as agreeing with an overt dual noun phrase as in (1):

- (10) àì bā kōēn-ēn
 1SG DECL light-DU
 ‘I am lighting (two fires).’

In the Nilo-Saharan phylum, the candidate dual pronoun is found in Kunama, whose apparent connection with Ama (Nyimang) was noted by Lionel Bender:

“Kunama is one of the minority of N-S languages having an inclusive-exclusive distinction ... [it] has the most developed such system, having also dual person, found otherwise only in a trace in Nyimang of East Sudanic.” (Bender 1996a: 17)

² Thus, we depart here from the hypothesis in Stevenson, Rottland & Jakobi (1992: 32) that formerly there was singular/dual/plural number inflection on the Ama verb for both subject and object.

The Kunama pronouns are shown in Table 1 in four quadrants, with inclusive pronouns in the first quadrant +speaker +hearer. Each quadrant has lines for singular, dual, and plural.

Table 1: Kunama pronouns (Abraha & Wedekind 1998)

	+Speaker		-Speaker	
+Hearer	-	-	ena	2 SG
	kiime ~ kimme	1 DU.INCL	eeme ~ emme	2 DU
	kime	1 PL.INCL	eme	2 PL
-Hearer	aba	1 SG	unu	3 SG
	aame ~ amme	1 DU.EXCL	iime ~ imme	3 DU
	ame	1 PL.EXCL	ime	3 PL

In all four quadrants, the dual has the form of the plural with the first syllable augmented, either by lengthening the vowel, in the Marda dialect, or geminating the medial consonant, in the Barka dialect. Either way, dual is marked by augmenting the first syllable with a second mora. The problem is that this augmentative dual mora has no segmental content of its own, and therefore no phonological resemblance to the Ama suffix /-ēn/. Without a phonological correspondence there is no prospect of directly reconstructing the Ama suffix /-ēn/ as an incorporated version of the Kunama dual. Moreover, Ama (Nyimang) is Eastern Sudanic and Kunama is not very close to it on the existing Nilo-Saharan trees (Bender 1996b; Ehret 2001), so any dual pronoun form that is the origin of both the Kunama and Ama dual morphemes would have to have disappeared or shifted a great many times in all other branches of Eastern Sudanic and some other branches of Nilo-Saharan. Again, the drastically reduced Kunama dual morpheme offers no independent evidence on the possible form of this hypothetical dual pronominal root.

Perhaps a clearer pathway for dual incorporation might be through cliticisation of a dual quantifier in post-verbal position. This possibility is suggested by the fact that Ama has a post-verbal plural quantifier /gàì/, which optionally highlights a plural argument of the clause as shown in (11).

Ama post-verbal plural quantifier /gàì/ (Norton 2012: 83)

- (11) a. dōrēŋ bā wūḍēŋ ʔèlē gài
 children DECL child see PL
 ‘The children saw a child.’
- b. wūḍēŋ bā dōrēŋ ʔèlē gài
 child DECL children see PL
 ‘The child saw children.’

Perhaps the dual suffix was formerly a post-verbal dual quantifier which cliticised to the verb and then was re-interpreted as a suffix. One advantage of this potential origin is that the post-verbal quantifier in (11) is not specific to subject or to object but can mark either, as the dual suffix also can. In addition, most Ama verb suffixes have a VC shape, and this fact could have supported the re-interpretation of /-ēn/ as a suffix, but left the plural quantifier /gàì/ as a separate word. Hence, this pathway could account for why there is a verbal dual without a verbal plural counterpart in the Ama verb.

Since there currently is no dual post-verbal quantifier in Ama, a place to look for possible quantifier roots is in the noun phrase.³ Ama has a nominal plural particle /gī/ somewhat phonetically similar to the post-verbal plural quantifier /gàì/, and there is also a nominal dual quantifier /kónè/

³ Compare the pre-verbal use of English nominal quantifiers: “All the children saw ...” ~ “The children all saw ...”

‘both’ which shows the adjectival prefix /kw-/ (Stevenson 1981) on a root /nê/,⁴ which is somewhat phonetically similar to the current dual suffix /-ēn/.

A problem is that /gài/ can be added to a dual clause as in (12). So /gài/ means plural in the sense of ‘more than one’, not plural as successor to dual in a scale of increasing number.

- (12) fōrā ḡū mōḡūr-ūḡ bē n-ēn gài
 rabbit TOPIC monkey-GEN enemy be-DU PL
 ‘Rabbit and Monkey they were enemies of each other.’

Perhaps the use of /gài/ with a dual clause developed after the dual quantifier had been incorporated into the verb, when the functions of /gài/ and /-ēn/ began to evolve separately. But that is speculative, and the use of /gài/ with a dual clause hinders an internal reconstruction of a dual quantifier because there is no gap in the system of quantification that a post-verbal dual quantifier would have filled. Also, the VC shape of the dual suffix /-ēn/, which arguably supported its re-interpretation from a word to a suffix, could also simply mean that it always was a suffix.

So although the existence of a post-verbal plural quantifier in Ama opens up the possibility of a cliticisation origin of the dual suffix, the cliticisation theory rests on rather weak circumstantial evidence. As Dimmendaal (2011: 149) states, “today’s morphology is not necessarily yesterday’s syntax”.

2.2 Borrowing

Another scenario to consider for the origin of the Ama dual suffix is borrowing a dual suffix from another language. Ama has borrowed from Nubian languages (Jakobi & Rottland 1991), but Nubian languages do not have a dual suffix. However, Sudanese Arabic has a nominal dual suffix *-een* which matches the form of the Ama dual suffix with a mid front vowel and an alveolar nasal.

A borrowing solution might seem only to shift the problem of the unusualness of the Ama suffix to another language, but in Arabic dual is typologically much less unusual, fitting into a well-attested singular-dual-plural nominal paradigm. What is highly marked in this scenario is the borrowing process itself, where an affix is borrowed from nouns of the source language onto verbs in the recipient language. I am not aware of any clear example where this has happened, but borrowing a noun affix onto verbs might be possible if the verb system in the recipient language is a good fit for the affix typologically (Thomason 2001: 77). This is true of Ama, because Ama nouns are typologically unusual in having no number morphology (Dimmendaal 2000) and number is expressed on the verb through distributive marking (Norton 2012). Hence, in Ama the verb is a better site for a number affix than the noun.

The typological argument alone does not prove a borrowing origin, however, since it would equally favour the development of a dual verb marker internally. The borrowing hypothesis also needs a plausible pathway for the borrowed dual to reach the Ama verb. Language structure is only borrowed when lexical borrowing is also present (Thomason 2001: 70), and a possible pathway for the dual suffix, which is a nominal suffix in Arabic, would be that it occurred on nominalisations of Arabic verbs which were then borrowed into Ama. When finite verbs were created from the borrowed nominalisations, the dual suffix was present and available as a new verb inflection on other verbs. However, this pathway is blocked by the fact that there has been almost no lexical borrowing from Arabic into Ama until very recently as Ama people have settled in towns. There is also a more specific problem that when verbs are borrowed in Ama they are borrowed as co-verbs in a co-verb-light-verb construction, as in the recent loan /rēkb-ēn eḡē/ ‘ride’ < Arabic *rikib*. Confusingly, the co-verb marker is also /-ēn/, but verbal inflection goes on the light verb so that the dual form is /rēkb-ēn eḡ-ēn/. If a dual nominalisation were borrowed from Arabic, then as a finite verb the dual suffix would be on the co-verb rather than the light verb, hence it would not be in a position in which it could be interpreted as an inflection available to other verbs.

The Arabic borrowing hypothesis can also be eliminated by looking at the history of the suffixes, since both of them are known to have had different vowels in the past as shown in (13). The shared

⁴ Compare the Nilo-Saharan root 1297 *à-ré (ADJ-split) ‘two’ in Ehret (2001).

mid front vowel today is actually a coincidence, the result of separate changes. The earlier high front vowel of the Ama dual suffix is found in Stevenson (1938).

Historical Changes to Dual Suffixes

- (13) Sudanese Arabic *-ain(i) > -een
- Ama *-īn > -ēn

Standard Arabic also has dual suffixes *-aa*, *-aan(i)*, *-anni* on verbs in pronominal subject and object paradigms. Although like the Ama dual these are verb suffixes, and are not confined only to subject or only to object, they cannot be the source of the Ama dual because the low vowel does not match the original high Ama vowel, because they are not used in Sudanese Arabic anyway, and because any verb that was borrowed from Arabic would appear as a co-verb so any suffixes would never be in a position where they could be interpreted as inflections available to other verbs.

2.3 Semantic shift

Having considered incorporation or borrowing as the origin of the dual suffix, we turn to a possible origin within the Ama verb itself. The candidate semantic shift that could account for the dual suffix is that it was originally a reciprocal marker. A reciprocal origin is suggested by the observation earlier in example (6) that the dual suffix can reciprocalise a stative clause. This clausal operation would be a surprising innovation if the dual suffix had been incorporated or borrowed from a nominal origin, but would fit well if the suffix had always been verbal. Thus, reciprocalisation of a stative clause looks like a relic of the suffix’s earlier function.

Reciprocal in Ama today is expressed by the voice system, as in Table 2. Passive, middle, and reciprocal have distinct suffixes in perfective aspect, but they all neutralise to the same suffix in imperfective aspect:

Table 2: Ama voice morphology

	perfective	imperfective
passive	āsīdāī-ái ‘was painted’	āsīdāī-āg ‘am being painted’
middle	bālīṅ-éí ‘was loved’	bālīṅ-āg ‘am loved’
reciprocal	bālīṅē-āg ‘loved each other’	bālīṅ-āg ‘love each other’

The suffix used in the imperfective for all non-active voices is identical to the reciprocal suffix in the perfective, where it appears with a theme vowel suffix that distinguishes the perfective from the imperfective. From this we propose the reconstruction that the non-active voice suffix /-āg/ in the imperfective developed a reciprocal reading, then this suffix was also extended to mark reciprocal in the perfective as well.

The proposal that Ama has innovated a reciprocal voice explains the anomaly of stative clauses. Since the voice system is for transitive verbs, the new reciprocal voice can only occur in transitive clauses, not in stative clauses. Therefore, the previous reciprocal marker is still needed for stative clauses. Hence, the dual suffix that reciprocalises stative clauses like (6), repeated here, is the old reciprocal.

Dual reciprocal stative clause

- (6) ē ɪr qū fōrāṅ-ō mòr n-ēn
- and elephant TOPIC rabbit-GEN friend be-DU
- ‘Elephant was a friend of Rabbit and vice versa.’

Without a dual suffix, the meaning is “Elephant was a friend of Rabbit” as expected from the syntax of the clause, but the dual suffix reciprocalises the relationship between the two participants, so the dual suffix functions as a dual-reciprocal operator.

We can see the lexical extent of the dual-reciprocal function of the dual suffix in Table 3. These are all verbs with a symmetric lexical semantics, verbs which conceptually require two participants functioning reciprocally – verbs of greeting, fighting, sex, and mutual relationship states.

Table 3: Verbs where the dual suffix gives a dual-reciprocal reading

<i>gloss</i>	<i>singular</i>	<i>dual reciprocal</i>
be a friend of	mōr n-ē	mōr n-ēn
duel	---	lū-ēn
greet ⁵	wòsò	wòs-ēn
exchange	wōsō	wōs-ēn
make love ⁶	túwárɔ̄	túwárɔ̄-ēn

The first two verbs in Table 3 call for particular comment. The predicate ‘be a friend of’ takes a dual suffix only for the reason already mentioned that the copula verb is not transitive and cannot use voice marking. But it also fits this class of verbs because ‘friend’ is a symmetric relational noun, so that ‘be a friend of’ is a symmetric complex predicate.⁷ Other symmetric relational nouns with the copula verb such as ‘be an enemy of’ would also fit here.

In the verb /lū-ēn/ ‘duel’, the dual suffix is obligatory. The root also occurs in /lū-ā/ ‘die’, so the root /lū/ by itself does not have a symmetrical semantics, rather the symmetrical meaning ‘duel’ arises only from the combination of root and dual suffix. The obligatory role of /-ēn/ for assigning the dual-reciprocal meaning ‘duel’ is further evidence that /-ēn/ is originally reciprocal.

Table 4 shows further examples of dual reciprocals where a voice strategy is used. Here, we see that dual-reciprocal voice marking is distinct from the reciprocal voice given above in Table 2. Dual-reciprocal voice uses a middle suffix which is placed before the dual suffix. Some verbs (greet, make love) have been found to use either dual-reciprocal voice or the dual suffix alone.

Table 4: Ama dual-reciprocal voice

<i>gloss</i>	<i>singular</i>	<i>dual reciprocal</i>
love	bālɔ̄ɔ̄	bālɔ̄ɔ̄-éɪ-ēn (impf. bālɔ̄ɔ̄-ēn-àg)
greet	wòsò	wòsɪ-éɪ-ēn
touch	sōò éí	sōò éí-éɪ-ēn
marry	túóɔ̄ (man), túóɔ̄-éí (woman)	túóɔ̄-éɪ-ēn
make love	túwárɔ̄ (man), túwár-áí (woman)	túwár-éɪ-ēn
copulate	túmīdò (man), tùmīd-éí (woman)	túmīd-éɪ-ēn

This further establishes that in Ama, reciprocal is a voice category, with a dual-reciprocal voice using the middle suffix as a second innovation of the system in addition to a reciprocal voice using the nonactive suffix. The verbs in Table 4 that use dual-reciprocal voice marking include verbs with a conjugal semantics (marry, make love, copulate) which are not fully symmetric because they assume a man participant and a woman participant. Since the middle suffix is used with a woman subject in

⁵ For ‘greet’, symmetry is present in intra-generational greeting but not inter-generational greeting: “People from the same generation greet each other by sliding the palm over the back of the other person’s hand. Youths shake each other’s hands vigorously, which is how they display their strength to one another. When an adult greets a child, they place their hand on the top of their head.” (en.wikipedia.org/wiki/Nyimang_people, accessed 28 November 2013)

⁶ The form shown here includes a punctual suffix /-ɔ̄/.

⁷ See Fiedler (2013) on complex predicates as a general phenomenon in Ama.

these verbs,⁸ I suggest that this is the origin of the use of the middle suffix for dual-reciprocal voice,⁹ first by extension from woman subjects to dual reciprocal subjects in the class of conjugal verbs, and then on other verbs of the language. As a result, middle and dual marking is ambiguous on conjugal verbs between a dual reciprocal reading and a dual reference reading:

- (7) nēŋēn gī kā ɬɔŋ-éɪ-ēn
 sibling PL VFOC marry-MID-DU
 ‘The sisters are both married.’
- (8) kwēī bā kér-ì ɬɔŋ-ìɖ-éɪ-ēn
 man DECL woman-DAT marry-PCT-MID-DU
 ‘The man got married to the woman and vice versa.’

For verbs where dual reciprocal readings are obtained from the dual suffix only, we can make the following internal reconstruction. As Ama adopted a reciprocal voice strategy, the old reciprocal suffix was retained in stative clauses where voice marking cannot occur, and in other predicates that conceptually require two symmetrical participants. Thus, the suffix narrowed from reciprocal to dual-reciprocal meaning. Having become associated with dual meaning, the suffix was then extended to dual reference and dual agreement as well. As a result, dual marking on verbs that conceptually require symmetrical participants is ambiguous between a dual reciprocal reading and a dual reference reading:

- (9) a. àŋi bā wòs-ēn ‘We (two) greeted each other, we (two) greeted someone.’
 b. àì bā wòs-ēn ‘I greeted (two people).’

The shift from reciprocal to dual-reciprocal to verbal dual following the innovation of reciprocal voice also explains the development of a verbal dual in Ama without a verbal plural counterpart.

So we have a new reciprocal voice strategy in transitive clauses through extended uses of the nonactive and middle suffixes, and we have an old reciprocal suffix found in stative clauses and more generally in symmetric predicates, giving it a dual semantics.

3 Reciprocal in Eastern Sudanic

The internal reconstruction of the dual suffix as originally reciprocal is corroborated in Table 5 by its phonetic similarity to reciprocal suffixes in a number of other Eastern Sudanic languages. The vowel of the Ama dual suffix was formerly high (Stevenson, Rottland & Jakobi 1992) and the combination of high front vowel and alveolar nasal lines up well with the other languages.

With reflexes in North, South, and Kuliak branches of Eastern Sudanic, we can suggest #-*ɪn* as the form of the proto-Eastern-Sudanic reciprocal suffix. The South Eastern Sudanic languages show greater diversity of form and function, with erosion of the final *n* and apparent fusion with a preceding consonant for reciprocal/reflexive meaning. The widespread occurrence of #-*ɪn* for reciprocal across Eastern Sudanic also supports the claim that reciprocal voice is an Ama innovation.

⁸ The data is from a male language consultant.

⁹ A disadvantage of using the non-active suffix /-àg/ in perfective dual reciprocals is that there would be neutralisation between perfective and imperfective, as the theme vowel would be lost in hiatus with the dual suffix vowel (Norton 2012: 90). But since this neutralisation occurs in other dual verb forms it cannot be assumed decisive here in selecting the middle suffix over the nonactive suffix. A semantic motivation is also needed for the origin of the middle suffix in dual reciprocals, such as the one proposed in the main text.

Table 5: Similarity to reciprocal in other Eastern Sudanic languages

Suffix	Gloss	Language	Classification	Source
-īn,-ēn	verbal dual	Ama	Nyima, North ES	Norton (2012), Stevenson et al (1992)
-īn,-ēn	reciprocal	Uncu	Kordofan Nubian, North ES	Williams (2009)
-in	reciprocal	Dilling with Kudur	Kordofan Nubian, North ES	Kauczor (1920), Jakobi (<i>pers.comm.</i>)
-in	reciprocal	Kururu (Tagle)	Kordofan Nubian, North ES	Jakobi (<i>pers.comm.</i>)
-Vn	reciprocal	Karko	Kordofan Nubian, North ES	Jakobi (<i>pers.comm.</i>)
-īn,-īm	reciprocal	Ik ¹⁰	Kuliak	Heine (1983)
-din	reciprocal/reflexive	Daju DarDaju	Daju, South ES	Aviles (1997)
-ri	reciprocal/reflexive	Shilluk	W Nilotic, South ES	Gilley (<i>pers.comm.</i>)
-kee	reciprocal/reflexive	Kalenjin	S Nilotic, South ES	Toweett (1979)
-ε	reciprocal/collective	Temein ¹¹	Temein, South ES	Stevenson (1957)
-ən	agentless passive	Gaam ¹²	Eastern Jebel, South ES	Bender & Ayre (1986)

4 Conclusion

The unique Ama verbal dual suffix is internally reconstructed as an old reciprocal suffix. This would mean that the Ama dual suffix is unrelated to Kunama dual pronouns, unrelated to the Ama dual quantifier, unrelated to the Sudanese Arabic dual suffix, and instead a cognate of reciprocal suffixes in other Eastern Sudanic languages.

Abbreviations

ABL	ablative	GEN	genitive
ACC	accusative	INCL	inclusive
APPL	applicative	MID	middle
CFOC	contrastive focus	NACT	non-active
DAT	dative	PCT	punctual
DECL	declarative	POSS	possessive
DEON	deontic	PL	plural
DISTR	distributive	SG	singular
DU	dual	TH	theme
EXCL	exclusive	VFOC	verb focus

¹⁰ Reciprocal verbs in Ik have a long extension /-īn-ós ~ -īm-ós/, of which /-ós/ can be identified as the passive suffix (Heine 1983: 161). Regressive tone spreading from the passive suffix may be responsible for the atypical high tone on the reciprocal suffix (back vowel quality also variably spreads regressively to the reciprocal suffix vowel).

¹¹ Reciprocal verbs in Temein have a long extension /-an-ε/ of which /-an/ can be identified as the passive suffix (Stevenson 1957: 187).

¹²The Gaam reciprocal suffix has a different form /-asa/ (Bender & Ayre 1986: 16).

References

- Aviles, Arthur. 1997. The phonology and morphology of the Dar Daju Daju language. MA Dissertation, University of North Dakota.
- Abraha, John & Klaus Wedekind. 1998. Kunama. Review of Bender (1996). *Notes on Linguistics* 80: 27–30.
- Bender, M. Lionel. 1996a. *Kunama*. Munich: Lincom.
- Bender, M. Lionel. 1996b. *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom.
- Bender, M. Lionel & Malik Agaar Ayre. 1986. *Gaam-English-Gaam Dictionary*. Southern Illinois University.
- Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.
- Dimmendaal, Gerrit. 2000. Number marking and noun categorisation in Nilo-Saharan languages. *Anthropological Linguistics* 42: 214–261.
- Dimmendaal, Gerrit. 2011. *Historical Linguistics and the Comparative Study of African Languages*. Amsterdam/Philadelphia: Benjamins.
- Ehret, Christopher. 2001. *A Historical-Comparative Reconstruction of Nilo-Saharan*. Cologne: Köppe.
- Fiedler, Ines. 2013. Complex predicates in Ama. In Thilo Schadeberg & Roger Blench (eds.), *Nuba Mountain Language Studies*, pp. 433–450. Cologne: Köppe.
- Heine, Bernd. 1983. The Ik language. Unpublished manuscript, University of Cologne.
- Jakobi, Angelika & Franz Rottland. 1991. Loan word evidence from the Nuba Mountains: Kordofan Nubian and the Nyimang Group. In Daniela Mendel & Ulrike Claudi (eds.), *Ägypten im afro-orientalischen Kontext. Aufsätze zur Archäologie, Geschichte und Sprache eines unbegrenzten Raumes. Gedenkschrift Peter Behrens (Afrikanistische Arbeitspapiere)*, pp. 249–269. Institut für Afrikanistik, University of Cologne.
- Kauczor, Daniel. 1920. *Die Bergnubische Sprache*. Vienna: Akademie der Wissenschaften.
- Norton, Russell. 2012. Number in Ama verbs. *Occasional Papers in the study of Sudanese Languages* 10: 75–94.
- Otero, Manuel. 2015. Dual number in Komo verbs. In Angelika Mietzner & Anne Storch (eds.), *Nilo-Saharan: Models and Descriptions*, pp. 127–140. Cologne: Köppe.
- Rilly, Claude. 2010. *Le Méroïtique et sa famille linguistique*. Leuven: Peeters.
- Schröder, Helga. 2012. Incorporated subject pronouns in word order typology. In Matthias Brenzinger & Anne-Marie Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics, Cologne 2009*, pp. 181–192. Cologne: Köppe.
- Stevenson, Roland. 1938. Nyimang grammar. Unpublished manuscript.
- Stevenson, Roland. 1956–57. A survey of the phonetics and grammatical structure of the Nuba Mountain languages. *Afrika und Übersee* 40: 73–84, 93–116, *Afrika und Übersee* 41: 27–65, 117–152, 171–196.
- Stevenson, Roland. 1981. Adjectives in Nyimang, with special reference to k- and t- prefixes. In Thilo Schadeberg & Lionel Bender (eds.), *Nilo-Saharan (Proceedings of the First Nilo-Saharan Linguistics Colloquium, Leiden, September 8–10, 1980)*, pp. 151–165. Dordrecht: Foris.
- Stevenson, Roland, Franz Rottland, & Angelika Jakobi. 1992. The verb in Nyimang and Dinik. *Afrikanistische Arbeitspapiere* 32: 5–64. Institut für Afrikanistik, University of Cologne.
- Thomason, Sarah G. 2001. *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.
- Toweett, Taitta. 1979. *A Study of Kalenjin Linguistics*. Nairobi: Kenya Literature Bureau.
- de Voogt, Alex. 2011. Dual marking and kinship terms in Afitti. *Studies in Language* 35.4: 898–911.
- Williams, Robert. 2009. Reflexivization and reciprocalization in Uncunwee. Presented to the 6th World Congress of African Linguistics, August 2009, University of Cologne.

DUAL NUMBER IN ETHIOPIAN KOMO

Manuel A. Otero

1 Introduction

While dual number is not altogether unheard of in Nilo-Saharan languages, it is nevertheless a rare feature limited to corners of grammars. This paper presents evidence for a tripartite (singular/dual/plural) number contrast in Ethiopian Komo. It argues that this distinction in number is nominal and construction-based, realized through a combination of segmental material, namely bound pronominal inflection on the verb, combined with an alternation of the tonal melody of the verbal complex. To illustrate this contrast, an introduction to Komo verb morphology and a brief overview of tone will be given.¹

Komo [xom] is a Koman language spoken along both sides of the border between Ethiopia and South Sudan. There are populations of Komo speakers in the Benishangul-Gumuz and Gambella regions of Ethiopia. Population estimates from 1971 cite 1,500 Komo speakers in Ethiopia and 10,000 in Sudan (Lewis 2009).² Ethiopian census data for native Komo speakers may be unreliable as speakers of Gwama, a related Koman language spoken in the same area in Ethiopia, identify themselves ethnically as Komo.³ Moreover, Gwama speakers vastly outnumber Komo speakers in Ethiopia and intermarriage is common.

Komo is vastly understudied, as are the remaining Koman languages. Previous work on Komo is limited to a grammar sketch (Burns 1947) and a short phonological sketch (Yehualashet 2008). Bender (1983, 1994) examined Komo within the larger Koman context. Furthermore, the internal structure of the Koman sub-family and its place within Nilo-Saharan has been under considerable debate by scholars (e.g. Bender 1996; Ehret 2001). Fortunately, more data on Koman languages are coming to light as researchers are presently working on the remaining languages of the family. This effort, combined with data presented in this paper, will hopefully contribute to a more definitive classification of the Koman languages within Nilo-Saharan.

This paper is organized as follows, Section 1 gives an overview of the Komo verbal system with specific emphasis on morphology and tone. Section 2 defines and discusses number contrasts and present a schematic overview for the tripartite number contrast in Komo. Section 3 presents Komo data for constructions illustrating the three-way contrast in number.

2 Komo overview

Komo is a relatively isolating language with predominantly SV/AVO word order and scant nominal morphology.⁴ By contrast, the verbal system is rather inflectional. A finite verb in Komo obligatorily inflects for at least one core (S/A) argument employing bound pronominal (BP) suffixes that distinguish person, number, gender and clusivity.⁵ Adjustments in the order of core arguments realized as independent NPs have direct consequences in the obligatory indexing of arguments on the verb. A Komo verb also occurs with a set of suffixes that code aspect/direction (AD). A position class diagram

¹ I would like to wholeheartedly thank all of the patient Komo speakers residing in the village of Yangu who provided the data for this paper. I would also like to thank SIL Ethiopia, the Linguistics department at the University of Oregon and my advisor Dr. Doris Payne for their support.

² I have travelled extensively throughout the Benishangul-Gumuz and Gambella regions of Ethiopia searching for proficient native-speakers of Komo to participate in the Multi-Lingual Education project for the Komo language sponsored by the Benishangul-Gumuz regional Bureau of Education and SIL Ethiopia. I estimate roughly 2,500 Komo speakers of varying proficiency in Ethiopia.

³ A 2006 census cites roughly 5,000 Komo speakers in Ethiopia.

⁴ OVA constructions are also possible.

⁵ Clusivity here refers to whether or not 1PL 'we' includes the addressee.

for the segmental morphemes comprising a Komo verb is seen below (Fig. 1). Note that a finite verb can index up to three core arguments.

Figure 1

STEM	-AD	-BP ₁	(-BP ₂)	(-BP ₃)
------	-----	------------------	---------------------	---------------------

Aspect and direction are inextricably linked in Komo. Currently, there are three major paradigms of contrasting AD morphemes occurring on the same slot on the verb.⁶ The first morpheme /-i/ (AD1) roughly corresponds to Imperfective aspect, where an event can be unspecified for completion, or an Itive directional, where or the literal direction of an action can be construed as moving away from a deictic center.⁷ The second morpheme /-ó/ (AD2) can be categorized as a Ventive directional, where the direction of a particular action occurs towards a deictic center. This Ventive morpheme is also used for events that occurred at a different location to that of the speech act. The third morpheme /-úk/ roughly corresponds to a Perfective aspect, where an event is construed as completed or having been completed in the location of the speech act. While AD morphemes in and of themselves do not have a direct bearing on the present discussion, they do contain vocalic elements which carry an underlying H tone and contribute to the overall tone pattern of an inflected verb.

2.1 Tone

Like many Nilo-Saharan languages, Komo displays contrastive tone. Burns' (1947) grammar sketch of Komo, proposes four level tones. By contrast, the present analysis proposes three level tones (L, M, H) on the surface with rising and falling contour tones occurring on single morphemes or as a result of two morphemes coming together. What is essential for the discussion at hand is the overall tone patterns of inflected verbs and verb roots.

Mono-morphemic verb roots in Komo tend to be monosyllabic (C)V(C) though disyllabic (C)VCV(C) roots also occur.⁸ A verb root in isolation can occur with one of three level tones – L, M and H.⁹ Komo verbs can be categorized into four tone classes according to the behavior of the tone of a given root in isolation compared to the tonal melody of the verb when inflected with a single argument. Tone melody here refers to the tone of the inflected lexical verb stem and its suffixes.

There are four major tone classes grouped according to the tonal melody of inflected verbs indexing a single (S/A) argument with respect to the root tone of the verb (Table 1). Recall that all AD morphemes are underlyingly H tone. Class II and Class III verbs (M and H tone roots in isolation respectively) retain the root tone when inflected. Verbs that have a L tone in isolation split into two classes – the first class retains the L root tone across the three AD paradigms (Class I) and the other and most numerous class of verbs become M.M when inflected in the AD1 paradigm.

Table 1: Inflected verb tone melody (single argument)

	Root tone	AD1 stem-sfx	AD2 stem-sfx	AD3 stem-sfx
Class I	L	L-H	L-H	L-H
Class II	M	M-H	M-H	M-H
Class III	H	H-H	H-H	H-H
Class IV	L	M-M	L-H	L-H

⁶ For simplicity the three aspect/directional morphemes are glossed AD1, AD2 and AD3. Aspect/directional morphemes cannot co-occur on a given verb.

⁷ The Itive and Ventive are also commonly known as the Translocative and Cislocative respectively.

⁸ Monosyllabic roots vastly outnumber disyllabic roots in the lexicon. For the sake of brevity, this paper is limited to monosyllabic roots.

⁹ There is only one mono-morphemic verb root with a contour tone in the lexicon pùl 'jump'.

The following data illustrate a Class I verb *giz* ‘enter’, which has a L tone root in isolation and becomes L.H when inflected with a single argument, regardless of whether the argument is singular or plural (1a-b). Notice how the same verb root in isolation retains its L tone in periphrastic negative auxiliary constructions (2a-b).¹⁰

- | | |
|---|---|
| <p style="text-align: center;">[class I]</p> <p>(1) a. <i>hàr</i> <i>giz-í-r</i> <i>á=gùbí</i>
 3SG.M enter-AD1-3SG.M LOC=hut
 ‘He enters the hut.’</p> | <p style="text-align: center;">[class I]</p> <p>b. <i>amòn</i> <i>giz-án</i> <i>á=gùbí</i>
 1PL.EX enter.AD1-1PL.EX LOC=hut
 ‘We enter the hut.’</p> |
| [class I] | |
| <p>(2) a. <i>hàr</i> <i>baʃ-í-r</i> <i>giz</i> <i>á=gùbí</i>
 3SG.M NEG-AD1-3SG.M enter LOC=hut
 ‘He does not enter the hut.’</p> | |
| [class I] | |
| | <p>b. <i>amòn</i> <i>baʃ-án</i> <i>giz</i> <i>á=gùbí</i>
 1PL.EX NEG-1PL.EX enter LOC=hut
 ‘We do not enter the hut.’</p> |

The following data illustrate a Class IV verb *bàb* ‘bury, plant’, which has a L tone in isolation (3b, 4b) and becomes M.M when inflected in the AD1 paradigm (3a, 4a).

- | | |
|--|---|
| <p style="text-align: center;">[class IV]</p> <p>(3) a. <i>aka</i> <i>bab-a</i> <i>bàmbé</i>
 1SG bury.AD1-1SG potato
 ‘I plant the potato.’</p> | <p style="text-align: center;">[class IV]</p> <p>b. <i>aka</i> <i>baʃ-á</i> <i>bàb</i> <i>bàmbé</i>
 1SG NEG-1SG bury potato
 ‘I do not plant the potato.’</p> |
| <p style="text-align: center;">[class IV]</p> <p>(4) a. <i>amòn</i> <i>bab-an</i> <i>bàmbé</i>
 1PL.EX bury.AD1-1PL.EX potato
 ‘We plant the potato.’</p> | <p style="text-align: center;">[class IV]</p> <p>b. <i>amòn</i> <i>baʃ-án</i> <i>bàb</i> <i>bàmbé</i>
 1PL.EX NEG-1PL.EX bury potato
 ‘We do not plant the potato.’</p> |

The tonal behavior of inflected verbs in Komo is crucial to the tripartite distinction in number being discussed. The following section will discuss number more in depth as well as propose a hypothesis as to how dual number is expressed in Komo.

3 Number

In an extensive treatment of number, Corbett (2000) challenges the preconceived notion that number is solely a nominal category. He makes a distinction between nominal number and verbal number, where the former refers to quantifying real-world entities and the latter to the multiplicity of events. How these types of number distinctions are encoded within a particular grammar varies across languages. What is important for the current discussion is that nominal number need not solely be expressed on a nominal. The following English data show a morphological singular/plural distinction made on the noun ‘dog(s)’ (5a-b), contrasting with a nominal distinction in number marked on the verb (5c-d). These data show that nominal number in English can be marked on both verbs and nouns. What is clear about these examples is that the distinction in number clearly refers to the number of participants, not the number of events.

¹⁰ See 3.2 and 3.5 for further discussion of periphrastic constructions.

pronominal morphology and the tone melody on the verb. One hypothesis as to how the dual is expressed in Komo is through the combination of an element that codes ‘singular’ with another element that codes ‘plural’. This is not altogether uncommon typologically as Goddard (1911) cites the combination of [\pm singular] elements code the dual in Hupa (Athabaskan). More specifically in Komo, the dual is expressed by combining a tonal melody that indicates one entity (‘singular’) coupled with bound pronominal inflection that marks more than one entity (‘plural’).

What is meant by a ‘singular’ tone melody? The vast majority of Komo verbs retain the tonal melody regardless of whether or not the S/A argument is singular. This is seen above in (1–4) as the lexical verbs do not change their tone class as a function of the number of the S/A argument. In this regard, a given verb with one tone melody (or tone class) coupled with [+singular] bound pronominal morphology codes ‘singular’ number while the same verb with the same tone melody with ‘plural’/[–singular] bound pronominal morphology codes ‘plural’ number (Fig. 2). This patterning is extremely common in Komo verbs.

Figure 2: Komo verbs (majority)

melody	bound pronominal	
one melody	[+singular]	singular
	[–singular]	plural

Nevertheless, there is a subset of Komo verbs that are inherently specified for number. These verbs have a distinct tone melody for [+singular] S/A arguments and another melody for [–singular] arguments. Thus a given lexical verb employs one tone melody used with [+singular] morphology to code ‘singular’ number and a distinct melody occurs with [–singular] bound pronominal morphology to code ‘plural’ number (Fig. 3).¹²

Figure 3: Komo verbs (rare)

melody	bound pronominal	
melody x	[+singular]	singular
melody y	[–singular]	plural

An example of a verb inherently specified for number, which alters in tone melody as a function of the number of the S/A argument is seen in (8a–b). In (8a), the verb *sókir* ‘(he) sits’ has a M.M (class IV) tone melody and the S argument is singular. In (8b) *sókin* ‘(they) sit’, the S argument is plural and the verb has a H.H (class III) melody.

	[class IV]		[class III]	
(8) a.	<i>hàr sòk-i-r</i>	<i>á=láú</i>	b.	<i>hòn sók-i-n á=láú</i>
	3SG.M sit-AD1-3SG.M	LOC=home		3PL sit-AD1-3SG.M LOC=home
	‘He is at home.’			‘They are at home.’

The hypothesis presented in this paper for dual constructions is the following: the dual is expressed by combining a [+singular] tone melody with [–singular] bound pronominal morphology. This is schematized below in (Fig. 4) where the shaded area indicates the dual distinction.

¹² There are also verbs that have suppletive roots for singular/plural number (c.f. *fá* ‘eat.SG’ / *ùf* ‘eat.PL’)

Figure 4: SG/DU/PL constructions

melody	bound pronominal	
melody x	[+singular]	Singular
	[-singular]	Dual
melody y		Plural

4 Komo tripartite number contrast constructions

The following sections provide evidence to illustrate tripartite number contrasts in Komo. In all cases, the same mechanism is employed: singular and dual constructions have the same tonal melody with different bound pronominal morphology, while dual and plural constructions have the same bound pronominal morphology (i.e. are segmentally identical) and differ in the overall tone melody. The constructions covered and the argument on which the number distinction is made is seen below (Fig. 5).

Figure 5:

Construction	Number contrast
<i>Declarative Transitive</i>	O argument
<i>Auxiliary Declarative Transitive</i>	O argument
<i>Intransitive Relative clause</i>	S argument
<i>Transitive Relative clause</i>	O argument
<i>Imperative</i>	A argument
<i>Auxiliary Negative Imperative</i>	A argument

4.1 Declarative transitive construction (double argument)

The double argument declarative transitive constructions has a [V-A-O] structure, indexing both A and O arguments on the verb. The following data show a three-way contrast in number for O arguments in this construction (9a-c, 10a-c). Notice that the singular and dual constructions retain the same tonal melody (9a-b, 10a-b) yet differ morphologically, while the dual and plural constructions are segmentally identical (9b-c, 10b-c), distinguished solely by an alternation of the overall tone melody.¹³

	[class IV]	[class IV]	[class III]
(9)	a. pàm-g-àp' touch-1SG-3SG.F 'I touch her.'	b. pàm-g-òn touch-1SG-3PL 'I touch them (two).'	c. pám-g-òn touch-1SG-3PL 'I touch them (3+).'
	[class I]	[class I]	[class III]
(10)	a. dùm-g-àr hit-1SG-3SG.M 'I hit him.'	b. dùm-g-òn hit-1SG-3SG 'I hit them (two).'	c. dúm-g-òn hit-1SG-3SG 'I hit them (3+).'

In Komo, a left-dislocated O argument realized as an independent NP is obligatorily indexed on the verb with the corresponding BP suffix. Moreover, independent NP arguments can occur with a series

¹³ When the A argument is 1SG, the O argument is always realized with a L tone in the double argument transitive construction.

of proclitics which make a singular/plural distinction.¹⁴ The resulting verbal complex indexes both A and O arguments on the verb and allows for a tripartite number contrast (11a-c).¹⁵

- [class IV]
 (11) a. à=bib tuʃ-i-p'-ar=ké
 SG=cow tie-AD1-3SG.F-3SG.M=ké
 'She tied the cow up.'

- [class IV]
 b. gò=bib tuʃ-u-p'-on=ké
 PL=cow tie-AD2-3SG.F-3PL=ké
 'She tied the (two) cows up.'

- [class III]
 c. gò=bib túʃ-ó-p'-on=ké
 PL=cow tie-AD2-3SG.F-3PL=ké
 'She tied the (3+) cows up.'

The following data show a similar tripartite distinction in number for O arguments when the A argument is plural. Example (12a) has singular verb morphology on a class IV verb and the nominal O argument also has a singular proclitic. Dual and plural constructions only differ in the tone melody on the verb (12b-c).

- [class IV]
 (12) a. à=bib tuʃ-am-ar=ké
 SG=cow tie-1PL.EX-3SG.M=ké
 'We tied the cow up.'

- [class IV]
 b. gò=bib tuʃ-am-on=ké
 PL=cow tie-1PL.EX-3PL=ké
 'We tied the (two) cows up.'

- [class III]
 c. gò=bib túʃ-ám-on=ké
 PL=cow tie-1PL.EX-3PL=ké
 'We tied the (3+) cows up.'

In certain constructions, there also appears to be a restriction on the class of verbs which can express the number contrast. Whether or not the restrictions are due to the semantics or the tone class of a given verb is still unknown. In the double argument declarative transitive constructions, dual distinction is not possible with the following verbs: *yé* 'slice' (class III), *k'óʃ* 'kill' (class III), *tùt* 'aim at' (class IV), *nàb* 'turn over (while roasting)' (class IV), *jil* 'see' (class I), *gàm* 'find, meet' (class IV).

¹⁴ Perhaps it is more accurate to describe the proclitics as distinguishing [+singular] from [-singular] number. These proclitics may also code definiteness, gender and/or be tied into information structure, more research is needed. The number distinction is relevant to the intents of this paper.

¹⁵ The morpheme *ké* is not yet fully understood. It appears to mark completion or recent completion of an action or event and does not have a bearing on the current discussion.

4.2 Auxiliary declarative transitive construction (double argument)

In Komo, future events or future intentions are expressed through periphrastic auxiliary verb constructions. Periphrastic auxiliary verb constructions contain an inflected auxiliary verb preceding a bare lexical verb root. In the future/intentative construction, the auxiliary verb *ó* takes bound pronominal morphology to index core arguments followed by the bare lexical verb root. To achieve the dual/plural contrast in number, the segmental morphology and tone of the inflected auxiliary verb is identical but the tone on the bare lexical stem changes (13a-c). This gives further evidence that it is the elements of the overall construction which allows for the tripartite contrast in number.

- [class IV]
- (13) a. *ó-nàn-ar* *t'òr*
 FUT-1PL.IN-3SG.M slaughter
 'We will slaughter him.'
- [class IV]
- b. *ó-nàn-òn* *t'òr*
 FUT-1PL.IN-3PL slaughter
 'We will slaughter them (two).'
- [class III]
- c. *ó-nàn-òn* *t'ór*
 FUT-1PL.IN-3PL slaughter
 'We will slaughter them (3+).'

The following data illustrate a three-way number contrast in periphrastic constructions with a left-dislocated NP_O argument. The lexical verbs in (14a-b) are of the same tone class and the morphology on the auxiliary verb distinguishes number. By contrast, (14b-c) are segmentally identical except for the tone class of the lexical verb *t'ór* 'slaughter' which facilitates the dual/plural distinction respectively.

- [class IV]
- (14) a. *à=mé* *ó-g-àr* *t'òr*
 SG=goat FUT-1SG-3SG.M slaughter
 'I will slaughter the goat.'
- [class IV]
- b. *gò=méí* *ó-g-òn* *t'òr*
 PL=goats FUT-1SG-3PL slaughter
 'I will slaughter the (two) goats.'
- [class III]
- c. *gò=méí* *ó-g-òn* *t'ór*
 PL=goats FUT-1SG-3PL slaughter
 'I will slaughter the (3+) goats.'

4.3 Intransitive relative clause construction

Relative clause constructions in Komo have the following structure [NP REL V], where the relativizer (REL) is specified for [±singular] number. Tripartite number distinctions can be made on S arguments of particular verbs in an intransitive relative clause constructions. Thus far only the verbs of posture *dɔf* 'stand' and *sɔk* 'sit', which also serve as existential copulas, allow for the three-way distinction in

- | | | | | | | |
|---------|----------------------------|----------|----------------------------|----------|---------------------------|----------|
| | [class I] | | [class I] | | [class II] | |
| (19) a. | dùm-í | t'at'òmò | b. dùm-í-m | t'at'òmò | c. dum-í-m | t'at'òmò |
| | hit.AD1-2SG | door | hit-AD1-2PL | door | hit-AD1-2PL | door |
| | '(You one) Hit the door!' | | '(You two) Hit the door!' | | '(You 3+) Hit the door!' | |
| | [class IV] | | [class IV] | | [class III] | |
| (20) a. | k'et-i | k'òlò | b. k'et-i-m | k'òlò | c. k'èt-im | k'òlò |
| | grab-AD1.2SG | hand | grab-AD1.2PL | hand | grab-AD1.2PL | hand |
| | '(You one) Grab the hand!' | | '(You two) Grab the hand!' | | '(You 3+) Grab the hand!' | |

Negative imperative constructions are periphrastic, employing the negative imperative auxiliary verb *lāk*, which takes the bound pronominal person marking, followed by a bare lexical verb root. A tripartite number contrast is also possible for certain verbs in these constructions. In (21a), the negative auxiliary verb indexes two singular arguments and the lexical verb root is class IV. The dual and plural constructions in (21b-c) are identical except for the shift in tone class of the bare lexical root.

- | | | |
|---------|---------------------------------------|-------------|
| | | [class IV] |
| (21) a. | lāk-i-in | p'òt'-ó |
| | NEG.IMP-AD1.2SG-3N | pluck-AD2 |
| | '(You one) Don't pluck it (towards)!' | |
| | | [class IV] |
| b. | lāk-i-m-in | p'òt'-ó |
| | NEG.IMP-AD1-2PL-3N | pluck-AD2 |
| | '(You two) Don't pluck it (towards)!' | |
| | | [class III] |
| c. | lāk-i-m-in | p'ót'-ó |
| | NEG.IMP-AD1-2PL-3N | pluck-AD2 |
| | '(You 3+) Don't pluck it (towards)!' | |

5 Conclusion

This paper has discussed dual number constructions in Komo, a little-known phenomenon in a lesser-known language of the Koman subfamily. Dual number distinctions in Komo form part of a three-way singular/dual/plural contrast in nominal number realized on the verbal complex. More specifically, dual number is construction-based, only expressed through a combination of a [–singular] tone melody of a given verb coupled with [+singular] bound pronominal morphology that indexes core arguments. It is argued that this tripartite contrast distinguishes nominal number and is not to be confused with the marking of verbal number in Komo pluractional constructions, which employed to express the multiplicity of events.

Abbreviations

1	first person	M	masculine
2	second person	N	neuter
3	third person	NEG	negative
A	transitive subject	O	object
AD	aspect/direction	PFV	Perfective
AUX	auxiliary	PL	plural

EX	exclusive	PLU	pluractional
IMP	imperative	REL	relative
IN	inclusive	S	intransitive subject
intrs	intransitive	SG	singular
LOC	locative	trans	tansitive
		V	verb

Tone: \acute{v} = High \grave{v} = low
 v =mid (unmarked)

References

- Bender, M. Lionel. 1983. Proto-Koman phonology and lexicon. *Afrika und Übersee* 66: 259–297.
- Bender, M. Lionel. 1994. Comparative Komuz grammar. *Afrika und Übersee* 77: 31–54.
- Bender, M. Lionel. 1996. *Kunama: Languages of the World/Materials* 59. Munich: Lincom.
- Bender, M. Lionel. 1996. *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom.
- Burns, Samuel J. 1947. *Notes Toward a Grammar of the Koma Language*. London: SIM, MS.
- Connell, Bruce A., Hayward, Richard J. & John A. Ashkaba. 2000. Observations on Kunama Tone. *Studies in African Linguistics* 29: 1–42.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- de Voogt, Alex. 2011. Dual Marking and Kinship Terms in Afitti. *Studies in Language* 35: 898–911.
- Ehret, Christopher. 2001. *A Historical-Comparative Reconstruction of Nilo-Saharan*. Cologne: Köppe.
- Goddard, Pliny E. 1911. Athapascan (Hupa). In Franz Boas (ed.), *Handbook of American Indian Languages* (Part 1), pp. 85–158. Washington: Smithsonian Institution, Bureau of American Ethnology.
- Hieda, Osamu. 2011. *Kumam Vocabulary with Grammatical Notes* (Studies in Nilotic Linguistics 4). Tokyo: Research Institute for Languages of Asia and Africa.
- Lewis, Paul M. (ed.). 2009. *Ethnologue: Languages of the World, sixteenth edition*. Dallas, TX: SIL International.
- Newman, Paul. 1980. *The Classification of Chadic within Afroasiatic*. Leiden: Universitaire Pers Leiden.
- Newman, Paul. 1990. *Nominal and Verbal Plurality in Chadic* (Publications in African languages and Linguistics 12). Dordrecht: Foris.
- Noonan, Michael. 1992. *A Grammar of Lango*. Berlin: Mouton de Gruyter.
- Norton, Russell. 2013. The Ama dual suffix: an internal reconstruction. Paper presented at the 2013 Nilo-Saharan Linguistics Colloquium, Cologne.
- Plank, Frans. 1983. On Humboldt on the dual. In Roberta Corrigan, Fred Eckman & Michael Noonan (eds.), *Linguistic Categorization*, pp. 293–333. Amsterdam/Philadelphia: Benjamins.
- Stevenson, Roland. 1938. Nyimang grammar. MS.
- Thompson, E. D. 1989. Kunama: phonology and noun phrase. In Lionel Bender (ed.), *Nilo-Saharan Language Studies*, pp. 280–322. East Lansing: African Studies Center, Michigan State University.
- Tucker, Archibald N. & Margaret A. Bryan. 1966. *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press.
- Yehualashet, Teshome. 2008. A brief phonological description and nomenclature of Kwama/Komo (A Nilo-Saharan language). Addis Ababa: Addis Ababa University, MS.

THE PRONOMINAL SYSTEM OF BAKA, A CENTRAL SUDANIC LANGUAGE OF SOUTH SUDAN

Christine Waag & Martin Phodunze

1 Introduction

BAKA is one of the Bongo-Bagirmi languages which belong to the Central Sudanic group of the Nilo-Saharan phylum. It is spoken in the area of Maridi and Yei River county in South Sudan. Some speakers are also found in the Democratic Republic of Congo (DRC). In total, they number about 25,000 to 30,000 (Gurtong, Jan. 2013, Internet source).

Most of the documentation on Bongo-Bagirmi languages discuss only the phonology and the listings of which languages belong to this group differ. The Ethnologue (Lewis et al. 2014) lists 41 Bongo-Bagirmi languages:

- **Bongo-Baka** (8)
 - Baka (A language of South Sudan)
 - Bongo (A language of South Sudan)
 - Morokodo-Beli (6)
 - Beli (A language of South Sudan)
 - Jur Modo (A language of South Sudan)
 - Mittu (A language of South Sudan)
 - Morokodo-Mo'da (3 languages of South Sudan)
- **Kara** (3)
 - Furu (A language of Congo (Kinshasa))
 - Gula (A language of Central African Republic)
 - Yulu (A language of Central African Republic)
- **Sara-Bagirmi** (29)
 - Birri (A language of Central African Republic)
 - Fongoro (A language of Chad)
 - Bagirmi (8 languages of Chad)
 - Sara (19 languages of Chad and Central African Republic)
 - **Sinyar** (A language of Chad)

Persson (1997) collected wordlists and some grammar to group the Bongo-Baka languages, a sub-group of the Bongo-Bagirmi languages. He did not have grammar data from Bongo and Baka, but “from both dialects of Beli and from Sopi”, which he groups together as the “BELI dialect cluster”. (Persson 1997: 24). In his grammar comparison he also includes languages from the other sub-groups of his listing of the Bongo-Baka languages, which are the Morokodo-Mo'da sub-group (Morokodo and Mo'da languages) and the Modo-Nyamusa sub-group (Modo and Nyamusa dialect clusters), see Persson (1997: 24ff.). He includes a section about pronouns in which he writes: “In all the dialects there are three sets of personal pronouns, an emphatic set which is used to emphasise subject or object of verbs, a second set which may be used for the object of verbs, and a possessive set. In dialects which distinguish intimate and non-intimate possession the possessive pronouns indicate non-intimate possession, and the object set is used for intimate possession. Third person pronominal subject is usually taken from the emphatic set, but first person is a prefix on the verb and second person is usually a tone change on the verb root.” (Persson 1997: 25). In third person singular Baka distinguishes between alienable and inalienable possession, which relates to Persson's ‘non-intimate’ and ‘intimate’ possession respectively. As in the languages Persson has investigated, in Baka the inalienable third person singular has the same form as the object pronoun.

Persson (1997: 26) lists the following pronouns in the emphatic set (his data of Wira is not complete):

	singular			plural		
	1	2	3	1	2	3
Mo'da	ma	wi'i	uně	oze	wohe	oga
Morokodo	ma	yī	mo	ze	ye	zě
Nyamusa	ma	yī	ne	je	ene	nega
Molo	ma	yī	ne	je	'je	pee
Wira	ma	nī	neko	ze		
Modo	ma	nī	bo	ze	kpe	lijë
Loro	ma	nī	vo	ze	kpe	ijě
Sopi	ma	yī	mini	je	se	ge
'Beli B.G.	ma	yī	ně	je	ye	mene
'Beli Wulu	ma	yī	ne	je	ye	mene

If the tone is neglected, the emphatic set of independent pronouns in Baka is, for comparison,

Baka	<i>maa</i>	<i>yīī</i>	<i>wo</i>	<i>zee</i>	<i>see</i>	<i>yee</i>
------	------------	------------	-----------	------------	------------	------------

The basic root forms of the Baka pronominal morphemes are comprised of the same three sets Persson mentioned for the languages above. There are variations only by lengthening or copying the vowel and alternating the tone (see Table 3 in the section about independent pronouns below). These variations differentiate meaning. Still, the pronominal system of Baka is not as complex as some other Central Sudanic languages, like Moru-Madi (Kilpatrick 2006) or Aringa (D. Andrew Angupale, personal communication). In Aringa and the Moru-Madi languages, pronouns and word order differ with different aspect, and thus they require more pronoun sets than the Bongo-Baka languages.

The Bongo-Bagirmi languages are fairly isolating. Baka differentiates subject and object by word order in the clause (Subject – Verb – Object). If subject and object are not expressed by explicit nouns, pronominal markers and the choice of connectors clarify the relations (“pronominal marker” as defined by Creissels et al. 2008: 91). The language has lexical as well as grammatical tone. Tone is so important that it is marked in the orthography.

The first section of this article introduces the pronominal affixes to verbs, their forms and use. The next section discusses the independent pronouns. It deals with subject, object, and possessive pronouns including the forms that are attached to words other than verbs, and it contains a discussion of logophoric pronouns and their use. The last section explains negation in Baka and a special use of possessive pronouns to indicate negation.

2 Pronominal affixes on verbs

Pronominal paradigms in Baka differentiate 6 forms: first, second, and third person each with singular or plural. This applies to verb paradigms as well as to pronoun sets.

In regard to pronominal markers, the verb forms differentiate three types: an uninflected form, an inflected form that involves tone changes as well as affixes, and a form with optional subject reference suffixes. Uninflected verbs may be combined with one of the latter two types in one clause.

The uninflected form of the verb always has a high tone on the first syllable of the stem, as used in the heading of Table 1. The tone on other syllables of the verb root is lexical. Table (1) shows the paradigms of the subject reference affixes of the inflected form for intransitive and transitive verbs. The two intransitive verbs *ndéré* ‘go’ and *ógu* ‘come’ are chosen as examples, as one is with a high and one with a low tone on the last syllable of the root.

Table 1: Subject reference paradigms of intransitive verbs and transitive

	intransitive		transitive	tone on first syllable of stem
	<i>ndéré</i> 'go'	<i>ógu</i> 'come'	<i>ĩmbjító</i> 'lift and put'	
1s	má-ndéré	má-ógu	má-ĩmbjító	High
2s	ndéré	ógu	ĩmbjító	High
3s	nderé	ogu	ĩmbjító	Low
1p	nderé-zé	ogu-zé	ĩmbjító-zé	Low
2p	ndéré-sé	ógu-sé	ĩmbjító-sé	High
3p	nderé-ní	ogu-ní	ĩmbjító-ní	Low

In this set of verb paradigms, transitive and intransitive verbs have the same subject reference affixes and also tone is used to differentiate the forms. In the singular the subject reference is a prefix, in the plural it forms a suffix. The tone on the affix is high. The tone on the first syllable of the verb stem is low in any third person and in first plural of this paradigm. In all other verb forms, including infinitive and derivations, the tone on the first syllable of the verb stem is high. (1) to (3) show sentences in which the forms of Table 1 are used, in (1) with a transitive and in (2) with an intransitive verb. The verbs with these forms in (1) to (3) are printed in bold.

- (1) kágá e **'dütü-ní** misi'di go bi ndéré=ye ndaá.
 tree PL **cover-3p** road eventually place go=3p.POSS it.is.not
 'Trees covered the road, so that they could not go through the place.' (lit.: a place of their going through was not there)
- (2) simi bi **ogu-ní** íri ní, zí=ye ínyi-yé úfu kobji.
 when **come-3p** there ECL DM=3p start.off-3p kill buffalo
 'When they came there, they started killing the buffalo.'

The verb forms of Table 1 do not need independent subject pronouns. The verb is sufficient to identify the subject, as *lurú* (third singular of *lúru*) and *ndjiszé* (first plural of *ndjisi*) in example (3).

- (3) **lurú** so'do máa kí lafúma ga bi **ndjiszé** gámá kéye
3s.look.to 3s.search 1s with companion PL who **stay.1p** walk with.3p
 ní máá
 ECL long
 'He looked for me and (my) companions, whom I used to walk with, for a long time.'

The verb form with optional subject reference suffixes has to be used in clauses with the connective *zí*. The forms and functions of *zí* are discussed in Manderson (1948). The most frequent use in narrative texts is as the marker of the event line ("development marker" by Dooley et al. 2000), as shown in (4). In this function it is introducing a main clause. The connective *zí* can also introduce a subordinate clause which is usually the result or purpose of the preceding clause. In all clauses, the connective *zí* is immediately followed by the subject. The subject precedes the verb and is additionally referred to by a suffix on the intransitive verb, as in *ndéréne* and *óguné* in (4).

- (4) Zí kára ba nda go ndéré-ne ógu-né 'be.
 DM woman this then eventually go-3s come-3s home
 'Then this woman went home.'

Transitive verbs do not usually carry this person suffix, since they are followed by an object or object pronoun instead, as shown in (5a) and (6). Example (5a) is a transitive clause with explicit nouns for the subject and object. In example (6), the object is the pronoun *a*.

If the subject is the same as in the previous clause, the connector *do* can be used instead of *zí*. Then the subject pronoun is not used. This can be seen in the second clause of example (5b). Here *do* is used instead of *zíye* ‘DM.they’. Again, the object pronoun *a* is used, referring to the explicit object of the first clause.

(5a) *zí nduwu e gírágá mongú landa*
DM anteater PL dig big mountain
‘Then anteaters dug the big mountain.’

(5b) *zí nduwu e gírágá mongú landa do gírágá phùtrụ a*
DM anteater PL dig big mountain CON dig break.through 3s
‘Then anteaters dug the big mountain and dug (a way) through it.’

(6) further shows that a subject pronoun is attached to the connective *zí*, if the clause is without an explicit noun for the subject. In this case the attached pronoun is obligatory, since the connective *zí* must be followed by the subject.

(6) *zí=ye ódó a mbá*
DM=3p tie 3s all
‘(and) they tied it all.’

Table 2 shows the paradigms of the subject reference suffixes and clitics, as used with the connective *zí*. The first column shows the subject pronoun clitics with the connective *zí*, the other columns show the verb paradigms with their suffixes. The tone of the subject reference suffix is contrastive to the last tone of the verb to which it is attached (polar tone), which is a high tone after *ógu* and low after *ndéré*. The transitive verb ‘lift and put’ is also listed in Table (2) for convenience, though transitive verbs do not usually have subject reference suffixes in this paradigm, but occur in the basic form only.

Table 2: Subject reference paradigms of verbs with *zí*-

		intransitive	transitive	
		<i>ndéré</i> ‘go’	<i>ógu</i> ‘come’	<i>ímbjító</i> ‘lift and put’
1s	<i>zí=ma</i>	<i>ndéré-ma</i>	<i>ógu-má</i>	<i>ímbjító</i>
2s	<i>zí=yí</i>	<i>ndéré-yí</i>	<i>ógu-yí</i>	<i>ímbjító</i>
3s	<i>zí=a</i>	<i>ndéré-ne</i>	<i>ógu-né</i>	<i>ímbjító</i>
1p	<i>zí=ze</i>	<i>ndéré-ze</i>	<i>ógu-zé</i>	<i>ímbjító</i>
2p	<i>zí=se</i>	<i>ndéré-se</i>	<i>ógu-sé</i>	<i>ímbjító</i>
3p	<i>zí=ye</i>	<i>ndéré-ye</i>	<i>ógu-yé</i>	<i>ímbjító</i>

In a clause that begins with the connector *zí*, the verb has to be followed by a nominal expression, either the suffix or a noun or pronoun, as in the clauses above. The suffix refers to the subject, while a noun or independent pronoun represents the object. The object can also be reflexive or reciprocal as shown in (7) and (8), respectively.

(7) *Zí=ye phíréphe kobí bí ufu-ní ba zí=ye ndá'balúgu*
DM=3p skin buffalo which kill-3p this DM=3p return.something
ro=yé 'be.
self=3p house
‘They skinned the buffalo which they had killed and returned home.’

(8) *Zí=ye tónó ócó ro=yé.*
DM=3p start beat self=3p
‘Then they started fighting each other.’

If the subject reference suffix is used with a transitive verb, it has a middle voice meaning, as shown in (9).

- (9) Zi=yi ótoómo a zí=a 'dǐ'dǐ-né cúku'dée
 DM=2s leave 3s for=3s cook-3s for.while
 'Then you leave it boil for a while,' (lit.: '...that it boils for a while')

3 Independent pronouns

As could be seen in the examples of the previous sections, independent pronouns are frequently used in Baka to represent the object after the verb. Other pronouns are used for subjects and possession. They differentiate the same 6 forms as the affixes to the verb: first, second, and third person each with singular or plural.

Table 3 shows the paradigms of the independent pronouns. Except for third person singular they resemble the pronominal suffixes to the verb with *zí* (see Table 2). They only have an additional vowel, which is lacking in the attached forms. The attached pronoun forms have been introduced with *zí* in Table 2. They are listed in Table 3 as well. They carry polar tone to the word they are attached to, except for the third singular =a, which is with low tone. The clitic =ne/=né is the attached form of either the alienable possessive pronoun *ené* or of the logophoric pronoun *née*, as opposed to the verbal suffix of Table 2, which is -ne/-né for any third singular subject.

Table 3: Independent pronouns

	SBJ		SBJ/OBJ	OBJ	POSS	POSS	enclitic
	SBJ	significant OBJ	logophoric	OBJ	POSS inalienable	POSS alienable	
1s	máa			máa	amá	amá	=ma/=má
2s	yíi			yíi	eyí	eyí	=yi/=yí
3s	wo	wo	née	a	a	ené	=ne/=né/=a
1p	zée			zée	ezé	ezé	=ze/=zé
2p	sée			sée	esé	esé	=se/=sé
3p	yée		yée	yée	eyé	eyé	=ye/=yé

Table (3) shows that subject and object pronouns are not distinguished by their form in first and second person and in the plural. The difference between subject and object is the place before or after the verb, respectively. Only third person singular discriminates between several forms. Here the subject pronoun is different from the normal object pronoun, and the possessive set differentiates alienable from inalienable in third person singular. In addition to the usual pronouns, the speaker of semi-direct reported speech has his/her own logophoric pronoun *née*. This will be demonstrated later.

3.1 Subject / emphatic pronouns

Since the subject is usually referred to by affixes on the verb, independent subject pronouns are not frequently used. They are employed for emphasis or clarification, as is shown in the following examples. So the subject pronoun set of Table (3) corresponds to the emphatic pronoun set of Persson (1997: 25) found in the other Bongo-Baka languages (see introduction).

In (10) and (11) the subject pronoun *wo* is used for an emphasised topic. In (10) it is used as 'the one who ...' in combination with a relative clause, and in (11) as 'that one' in combination with a demonstrative.

- (10) **wo bi** kí owú simi=né ba ni nda tí ndólo yée
3s who with child in=3s this 3s.be although call them
 idíní ..., **yée ga bi** gbre ya
 shall.3p **3p those** two that
 ‘The one with the child in her, though calling them to ..., those two (answered) that,’

In the context of (10) three women went to get firewood. One of them is pregnant and needs help, which is refused by the other two. Thus the one (*wo* ‘she’) is contrasted to the other two (*yée* ‘they’)

- (11) **wo née** ndĩsì 'dĩ'bi tú'dú 'yì e
3s that 3s.stay take many person PL
 ‘That one (usually/always) takes many people.’

In the context of (11) different ways of cultivation are explained. (11) is about one of these ways.

In (12) clarification is needed, since the participants before (12) are a woman’s brother and her brother-in-law. Now, in (12) a second brother is added. The pronoun *yée* ‘they’ refers to the two brothers, and the sentence (12b) explains why there are now three participants instead of two, and who these participants are.

- (12) a. Zí=ye ìnyì-yé 'yì e ota ndéré-ye simi súwú.
 DM=3p start.off=3p person PL three go-3p in forest
 ‘Three of them started off to go to the forest.’
- b. **Yée** gbre kí lúndu=né nda go kí yoko=yé.
3p two with brother=3s.LOG then now with brother.in.law=3p
 ‘They two, (he) with his brother, and their brother in law.’

The third person singular pronoun *wo* can also be used in object position. It then signifies the topic or an important participant. In (13) it is employed for the object within a relative clause, which is the subject of the main clause, the main topic of the sentence, and the main participant of the story.

- (13) ìnyìní kí 'yì bándága ga bi ndĩsìní bándá **wo** ní ndéré-ye.
 get.up.3p with person his.guards PL REL stay.3p look.after **3s** ECL go-3p
 ‘..., he got up with his guards, who used to look after him, and they went.’

(14) shows the independent pronoun *zée* ‘we’ as a complement.

- (14) Mizefì 'yì ga bi ndĩsì-ní ógúútù simi sùkùlù nda
 first person those stay-3p come.first in school then
 'duo zée.
 straight.away 1p
 ‘We became the people who arrived at school first.’ (lit.: ‘The first people, who (continuously) arrived at school first, were we straight away.’)

3.2 Possessive pronouns

Possessive pronouns are used in possessive constructions as *okó eyé* ‘their fight’ in (15), *ledre ezé* ‘our matter’, *pho'du eyì* ‘your (SG) fire’, *phala amá* ‘my spear’, or *éyì amá ga gére née* ‘those things of mine’. In Baka, possessive pronouns are also used for marking negation. This use in negation will be explained later.

The *yée* in (15) is the object of the first verb which is different from the subject; else the reflexive *yé* would have been used, as was shown in (7) and (8) above. The subject of the second verb refers to

the same participant as this object. The object of the second verb is a possessive construction with the possessive pronoun *eyé*.

- (15) *Zí=ye ótoómo yée méngi okó eyé.*
 DM=3p leave 3p do fight 3p.POSS
 ‘They left them fighting.’ (lit.: ‘doing their fight’)

The possessive pronoun can also be used as the agent of an uninflected verb, as in (16), where *ezé* ‘our’ is referring to the agent of *ndísjí ndéndeke* ‘playing’. *ndísjí* in *ndísjí ndéndeke* has the grammatical function of continuous aspect. The whole phrase *ndísjí ndéndeke ezé* ‘our playing’ is a secondary complement to the verb *azé* ‘we were’, with the primary complement *sími baba mánga* ‘in the mango forest’.

- (16) *Ní ógu ní, azé yí ezé sími baba mánga ndísjí ndéndeke ezé.*
 3s.be come ECL 1p.be self 1p.POSS in forest mango stay play.a.play 1p.POSS
 ‘He came while we were playing in the mango forest.’

If the agent is less prominent, the attached form can be used, as =*ye* ‘their’ on the last verb in (17). This pronominal element is not the usual subject reference suffix, since the same construction can be used without it as in (18), a sentence from the same story.

- (17) ...; *kágá e 'dutu-ní misi'di go bi ndéré=ye ndaá.*
 tree PL cover-3p road now place go=3p it.is.not
 ‘Trees covered the road, there was no way for them to go through’ (lit.: ‘... a place for/of their going through was not existent.’)

- (18) *Zí=ye ógu ndiki mongú mingi'di kí séngú 'dutu-ní*
 DM=3p arrive find big thorny.vine.sp. with sharp.grass.sp. cover-3p
misi'di go bi ndéré ndaá.
 road now place go it.is.not
 ‘They arrived finding big thorny vines and specific sharp grass covering the road, there was no way to go through.’ (lit.: ‘... a place for/of going through was not existent.’)

The possessive pronoun in third person singular differentiates if the possessed item is alienable or not. Thus, *a* is inalienable, e.g. used for relatives and body parts as *síli a* ‘his hand’ or *lúndu a* ‘her brother’, as shown in (19), while *ené* is used for alienable items as *ká'dá ené* ‘his rope’ or *phala ené* ‘his spear’, as is shown in (20) and (21).

- (19) *Zí lúndu a ndéré-ne méngi a tí kenée*
 DM brother 3s.POSS.inalienable go-3s do 3s exactly like that
 ‘And her brother did (lit.: went doing) it exactly like that.’
- (20) *Zí 'bua ókúómo sama zí=a; zí=a yémeómo phala ené.*
 DM father prepare.leave bow for=3s DM=3s prepare spear 3s.POSS
 ‘Then his father prepared for him a bow; and/then he prepared (for him) his spear.’
- (21) *zí=a ólúógu-né geré ndíki kéré ené, sama ené*
 DM=3s come.out-3s immediately find arrow 3s.POSS bow 3s.POSS
kí phala ené.
 with spear 3s.POSS
 ‘..., he comes out and finds his arrow, his bow, and his spear.’

The pronouns can differentiate if something is owned (*kífú ené* ‘his/her seed’) or a part of something (*kífú a* ‘its seed’, the seed of the plant). (22) and (23) demonstrate this contrast in elicited clauses, while (24) and (25) are drawn from natural text. The context of (24) is a recipe how to make a sauce with the seeds of a special plant, while in (25) a child needs to get his food by himself.

- (22) Zí owú ánu kífú=a
DM child eat seed=3s.POSS
‘The child ate its seed (seeds of the sorghum plant).’
- (23) Zí owú ánu kífú=ne
DM child eat seed=3s.POSS
‘The child ate his seeds (seeds of the child).’
- (24) zí=yì lúgu kífú a née drá.
DM=2s wash seed 3s.POSS this clean
‘(When you have finished shelling it all,) then you wash the shelled seeds clean.’ (lit.: ‘... wash these its seeds clean’)
- (25) Zí=a ndéré-ne 'dəcì ngbúrádá ené ógu ánu a.
DM=3s go-3s break.off fruit, sp. 3s.POSS come eat 3s
‘(Then) he went and cut his fruit and ate it.’

Also aptitudes that cannot be passed on, such as *cóngóro a* ‘her strength’, are inalienable, as shown in (26).

- (26) Cóngóro a ndaá gi zía ni kí owú simi=né.
strength 3s.POSS it.is.not because 3s.be with child in=3s
‘She had no strength because she had a child in her.’

3.3 Logophoric pronouns

According to Creissels et al. (2008: 144) logophoricity is very “common in African languages, in an area ranging from Senegal in the west to Ethiopia in the east, and cutting across genetic boundaries.” Clements (1975: 141f.) explains that “several languages of the eastern branches of Niger-Congo are known to employ logophoric pronouns which are morphologically distinct from personal and reflexive pronouns. For instance, in Ewe, the pronoun *yé* [...]”. This pronoun contrasts with personal pronouns:

- (1) Kofi be yè-dzo ‘Kofi said that he (Kofi) left.’
say LOG-leave
- (2) Kofi be mè-dzo ‘Kofi said that I left.’
- (3) Kofi be é-dzo ‘Kofi said that he/she (≠Kofi) left.’

Here he states (Clements 1975: 141) that “logophoric pronouns (are employed) to distinguish reference to the individual whose speech, thoughts, or feelings are reported or reflected in a given linguistic context, from reference to other individuals.” More generally, Creissels et al. (2008: 144) define that “logophoricity involves special third-person pronouns occurring in dependent clauses only and expressing coreferentiality with the subject of the main clause.”

Baka has a logophoric pronoun *née* for third person singular, and *yée* for plural, which is mainly used in reported speech. Reported speech is traditionally distinguished in two types: direct as ‘he said, “I will go”’ and indirect speech, as ‘he said that he will go’. “With direct speech reports ... the perspective of the original speaker is maintained; with indirect speech reports ... the perspective shifts towards that of the reporting speaker and his speech situation” (De Roeck 1994: 332). However, Baka can be analysed to distinguish two types of reported speech: direct speech and speech with a mixed

perspective. In a strict sense, Baka has no indirect speech according to the traditional definition. Thus, in the following, the speech with a mixed perspective will be called semi-direct speech, as defined by Aikhenvald (2011: 348), “While in indirect speech reports the person reference must ‘shift’ to the perspective of the narrator, there is no such shift in direct speech reports and in quotations. And in semi-direct speech, the reference for some participants is shifted, while for others it is not”. In semi-direct reported speech in Baka, the reported speaker is referred to in third person, as is usually the case for indirect reported speech in English.

Before further discussing the reporting of speech, (27) gives an example for comparison, if the speech is not reported. As can be seen in this example, the pronouns are used the same in a conversation as in a story. Independent subject pronouns are not needed, since the verb form shows the subject.

- (27) a. ndéré karana 'da
[2s].go yesterday where?
'Where did you go yesterday?'
- má-ndéré simi gara.
1s-go in town
'I went to town.'
- b. Zi=ma ndíkí kí Alékisi zí=ze odrozé do ledre
DM=1s meet with Alex for=1p speak.1p on matter
'bi lúndu a
of brother 3s.POSS
'I met with Alex to discuss (with him) the issue of his brother.'

In reporting of speech, the logophoric pronoun is employed in semi-direct speech, not in other forms of reported speech. The verbs in semi-direct speech take the same third person forms as are employed in other contexts. The logophoric pronoun *née* is used for the speaker within his or her speech, as in example (28).

- (28) Togú **née** mengi omo éyí miánu go **Née** utúasá
if **3s.LOG** 3s.make 3s.leave thing NLMZ.eat eventually **3s.LOG** 3s.can
í'bí a zí owú née wá.
give 3s to child that NEG
'(She₁ said that) if she₁ makes food and leaves it, she₁ will not give it to that child.'

If the speaker is the subject or object in semi-direct reported speech, this pronoun is obligatory. In the plural, *yée* has to be used, as in (29).

- (29) ... ya **yée** idi.ní ndéré gámá só'do ngírí.
that **3p** let.3p go walk search fire.wood
'... that they should go to look for fire wood.'

A third person subject in semi-direct reported speech without this pronoun is not the speaker, but possibly the addressee, as can be seen in (30). In the translation direct speech is used, since the English would be ambiguous in this case.

- (30) *bi romo do ledre e go za mbá, misi'di*
 as 3s.defeat on matter PL eventually Completely all road
bi née 'dutu wo nderé phutu go mbá ní,
 REL 3s.LOG 3s.close 3s 3s.go 3s.break.through eventually all ECL
ní miútúásáne zí=a ékí-ne komo Wála ba.
 3s.be appropriate for=3s climb-3s eye mahogany.tree this
 ‘(The big chief ... spoke the matter that,) “As you have defeated all the matters on the road that I closed it (with), (as) you have gone through all of them, it is appropriate for you to climb (the eye of) this mahogany tree.” (‘the eye’ here means ‘up’).’

In (30) the pronouns differentiate the participants though they are all third person singular. In this semi-direct reported speech of a chief to a wise child, the first subject is the addressee, the child. Here the third person singular form of the verb (with a low tone on the first syllable) is used without any independent pronoun. The second clause, *bi née 'dutu wo* ‘that he closed it with’ is a relative clause describing the road. The subject of this relative clause is the speaker (the chief). Thus, in addition to the third person singular form of the verb, the pronoun *née* has to be used. The road is referred to with the pronoun *wo*, since it is the topic of the relative clause (in object position). In the rest of the sentence, the subject is the child again and no independent pronoun is used. Third singular -a is used with *zí* for the child in the last clause.

The alternative way of reporting speech in narrative is direct reported speech. Then first and second person references are used for the speech participants within the speech, as in example (31).

- (31) *Zí kára ba úku a yaá, “Má-áyí íbì késí zí=yí.”*
 DM woman this say 3s that “1s-will give Money to=2s
 ‘This woman spoke, “I will give you money.”’

The logophoric pronouns *née* and *yée* are also used in object position and for second person, as can be seen in examples (32) to (39).

- (32) *zí=a úku a zí=ma kídí ísì naná née*
 DM=3s say 3s to=1s that dog 3s.bite 3s.LOG
 ‘He₁ told me that the dog bit him₁ (the speaker).’
- (33) *zí=a úku a zí=ma kídí ísì naná wo*
 DM=3s say 3s to=1s that dog 3s.bite 3s
 ‘He₁ told me that the dog bit him₂ (not the speaker).’

In (32) the object in the reported speech is the speaker (the speaker has been bitten), while in (33) it is someone else. Thus, in (32) the pronoun *née* has to be used instead of *wo*. The object pronoun *wo* has to be used in (33), since it refers to a salient participant. An insignificant participant would be referred to by an explicit noun, since it would not have been active from the previous clause.

Alternatively to (32), direct speech could be used, as in (34).

- (34) *zí=a úku a zí=ma kídí ísì naná máa*
 DM=3s say 3s to=1s that dog 3s.bite 1s
 ‘He told me, “The dog bit me.”’

(34) is ambiguous as to who is bitten. It can be the speaker or the addressee, since the addressee is first person. There is no way to express unambiguously that the addressee is the one bitten in a sentence as (34).

(35) is itself direct speech, but includes semi-direct reported speech within the direct speech. In it the logophoric pronoun is used for the speaker of the semi-direct reported speech, who is the addressee

of the direct speech. In Baka third person is used within the semi-direct reported speech, though it has to be translated with second person in English.

- (35) úku zǐ=a kídí **née** ní ndéré gara
 [2s].say to=3s that **3s.LOG** 3s.be go town
 ‘You told him/her that you (the speaker) go to town.’

Since the person going to town is the speaker of the semi-direct reported speech (‘you’), the logophoric pronoun *née* has to be used in addition to the third person marking on the verb (*ní*), as in (35). Without *née*, as in (36), the person going to town is not the speaker, but may be either the addressee of the speech (*zǐa*) or someone else.

- (36) úku zǐ=a kídí ní ndéré gara
 [2s].say to=3s that 3s.be go town
 ‘You told him/her that he/she (not the speaker) go to town.’

In example (35), as in any reported speech with the use of the logophoric pronoun, the third person form of the verb is used. (37) is itself direct speech, and includes indirect reported speech within the direct speech. For this reported speech within the direct reported speech, the second person form of the verb is instead used, as shown in (37), where *áyí* is used instead of *ní*. This way of reporting the speech is without the logophoric pronoun.

- (37) úku zǐ=a kídí **áyí** ndéré gara
 [2s].say to=3s that **2s.is** go town
 ‘You told him/her, that you go to town.’

The same applies to plural participants. (38) indicates reported speech (semi-direct speech) with a logophoric pronoun and (39) indicates reported speech without a logophoric pronoun. There is no difference in meaning, if semi-direct speech (with logophoric pronoun) or indirect speech (with second person subject) is used.

- (38) uku-sé zǐ=a kídí **yée niyí** ndéré gara
 say-2p to=3s that **3p 3p.are** go town
 ‘You (pl) told him/her that you (the speakers) go to town.’

- (39) uku-sé zǐ=a kídí **ásé** ndéré gara
 say-2p to=3s that **2p.are** go town
 ‘You (pl) told him/her that you go to town.’

In Baka, the logophoric pronoun is not used, if the speaker is first person. If a speaker is reporting his/her own speech, first person pronouns and verb forms have to be used, as in (40) for singular and (41) for plural.

- (40) zǐ=ma úku zǐ=a kídí **má-áyí** ndéré gara
 CM-1s say to=3s that **1s-be** go town
 ‘I told him/her that I go to town.’

- (41) ukuzé zǐ=a kídí **azé** ndéré gara
 say.1p to=3s that **be.1p** go town
 ‘We told him/her that we go to town.’

In Baka, usually a speech follows a speech introducer, as in the previous examples. It is possible that the speech precedes the speech act verb for emphasis on the speech, as in example (42). In this case, the logophoric pronoun is used as a cataphoric referential. Since indirect speech cannot precede the speech act verb in English, direct speech is used in the translation.

- (42) “**née** ili lolu o'do née wá,”
 3s.LOG 3s.want any.more man that NEG
 zí Miyáká úku ledre zí lúndu=né kenée
 CM Miyaka say matter to brother=3s.LOG like.that
 ‘“I do not want that man any more” is what Miyaka said to her brother.’

In combination with the connector *zí*, the usual third person pronominal is =*a*, as in (43). The logophoric form with this connector is =*ne*, as can be seen in (44).

- (43) *zí=a* ógu-né go 'be.
 DM=3s come-3s eventually home
 ‘She reached home.’
- (44) *ídí* ékí za do odu wála née 'dága íri
 3s.let climb completely on top mahagony tree that up there
 gí ro *zí=ne* ówo 'dí'bi ma'dii ené.
 for to=3s.LOG see well truth his
 ‘Let you climb all the way to the top of this mahagony tree up there for me to know your truth.’

The logophoric pronoun can also be found outside of speech, though this is very rare. The following gives an example.

- (45) Misindíkéří sembe owú ba gí ro *zí=ne* ánu a máá,
 one.legged.man 3s.wait child this for to=3s.LOG eat 3s for long
 misi'di bi gí ro *zí=a* ánu owú ba ndaá 'bi ené wá.
 way that for to=3s eat child this is not of his NEG
 ‘The one-legged man₁ waited a long time to eat this child (lit.: for this child₂ that he₁ eat him₂), but there was no way for him₁ to eat that child₂.’

The expression *gí ro zine ánu a* ‘that he₁ eat him₂’ in example (45) uses the logophoric referential =*ne* with the connector *zí*. It clarifies that the subject of the main clause, the man waiting, is also the subject of the dependent action of eating. The logophoric referential cannot be used in the second part of this sentence, since here the expression *gí ro zia ánu owú* is referring to the way instead of the waiting of the one-legged man. Thus, the object *owú* is made explicit to avoid ambiguity, so that the subject =*a* ‘he’ in *zia* is referring to the man and not to the child.

The independent logophoric pronoun *née* is not used in subject position in this kind of clauses, as is shown in examples (46) and (47). The logophoric subject form only shows in connection with the marker *zí*, as in (47).

- (46) bi owú inyi gí ro ndéré lúru ngére ní
 which child 3s.get.up for go see chief ECL
 ‘..., in which the child got up to go and see the chief, ...’
- (47) bi owú inyi gí ro *zí=ne* ndéré lúru ngére ní
 which child 3s.get.up for to=3s.LOG go see chief ECL
 ‘..., in which the child got up to go and see the chief, ...’ (emphasis on who the child is)

Nevertheless, the logophoric pronoun *née* is found in object position in such sentences, as in (48), adapted from (45).

- (48) *Misindjékérí sembe owú ba gí ro zí=a ócò née máá,*
 one.legged.man 3s.wait child this for to=3s beat 3s.LOG for long
 ‘The one-legged man₁ waited a long time for this child₂ that he₂ beat him₁.’

Here the subject pronominal in the expression *gí ro zía ócò née* is the usual =*a* again, referring to the child, while the subject of the main clause, the man waiting, is the object of the beating in this sentence.

As discussed in the previous section, the third singular possessive pronoun differentiates between alienable and inalienable possession. In the attached form it is =*ne* and =*a*, respectively. The attached inalienable possession can also have the logophoric form, as shown in the following examples. In the alienable possession a logophoric reference cannot be differentiated, since the logophoric pronoun has the same form as the alienable possessive pronoun (see Table 3 above).

In (49) the speaker (the chief) is handing over the chieftom from his hand to the child's (the addressee's) hand. Since the hand is inalienable, the normal possessive pronoun is *a* as in *sílí a* ‘his hand’. But when the possessive pronoun is referring to the speaker, it becomes =*ne* as in *sílíne* ‘his (the speaker's) hand’.

- (49) *bí ba ní ngére bí do sílí=ne ba ní née nda go*
 as this ECL chief which on hand=3s.LOG this ECL 3s.LOG then now
ógu í'bí a do sílí=a, ...
 come give 3s on hand=3s
 ‘(The chief said,) “... , as it is like this, the chieftom that I have (lit.: ‘which is in my hand’) I am going to hand it over to you (lit.: ‘give into your hand’), ...”’

The attached logophoric possessive pronoun is even used for objects that are inalienably possessed by the subject of the clause, as can be seen in (51). In (50) *lúndu=a* ‘her brother’ is the subject of the clause and the inalienable possessive pronoun *a* is used. In (51) *lúndu=né* ‘her brother’ within the prepositional phrase refers to the brother of the subject. Thus, the logophoric form =*ne*/*né* is used. If *lúndu=a* would be used instead, he would be the brother of someone else. The same applies for the object *komo=né* ‘her eye’ in (51).

- (50) *zí lúndu=a útú-ne kí kúkugú*
 DM brother=3s.POSS fall-3s with laughter
 ‘Her brother laughed.’ (lit.: ‘... fell with laughter’)
- (51) *zí=a ónzó komo=né ro lúndu=né*
 DM=3s put eye=3s.LOG on brother=3s.LOG
 ‘She saw her brother.’ (lit.: ‘She put her eyes on her brother’)

4 Politeness and pronouns in negative constructions

Possessive pronouns are also needed in polite negative constructions. Negative clauses in Baka can be formed with *ndaá* ‘be not’, as in (52).

- (52) *togú sáka ndaá zí=yí rára do kú'bí.*
 if mat it.is.not for=2s [2s].spread on winnowing.plate
 ‘If you have no mat you can spread it on a winnowing tool.’

Usually the negative clause ends with the particle *wá*, as in (53).

- (53) *née utúasá í'bí a zí owú née wá.*
 3s.LOG 3s.can give it to child that NEG
 '(This woman said that) ... she will not give it to that child.'

In the Baka language, the possessive pronouns *amá* 'my', *ené* 'his/her/its', *eyí* 'your', *ezé* 'our', *esé* 'your, PL', and *eyé* 'their' play a prominent role in making a statement negative. Using the possessive pronoun is a polite way of making a negative statement, since it signals the negation to the audience early in the clause, as in examples (54) to (56).

- (54) *má-ndéré amá gara wá*
 1s-go 1s.POSS town NEG
 'I am not going to town.'

The possessive pronoun used for marking a negative statement has the person and number of the subject, which is first person singular in (54), and it immediately follows the verb. In (55) the subject is third person plural, and in (56) is first person plural.

- (55) *ya yéé ìmbjotoní eyé ngírí wá.*
 that 3p put.3p 3p.POSS fire.wood NEG
 '(The two ... said that) they will not put the fire wood (on her head).'

If the possessive pronoun in (55) referred to the owner of the firewood, it would have been singular, since it belonged to a third woman. The only plural participants in this clause are the two women who are the subject.

- (56) *nderézé ezé simi sùkùlù akpa káa zí 'bí lafú=ze wá.*
 go.1p 1p.POSS to school early like of mate=1p.POSS NEG
 'We did not go to school early like the others.' (our classmates)

When making the negation without a possessive pronoun, it is an impolite response as in example (57).

- (57) *má-ili wá.*
 1s-want NEG
 'I do not want.'

Politely you need to say,

- (58) *má-ili amá a wá.*
 1s-want 1s.POSS 3s NEG
 'I do not want (it).'

With the negation words *ndá* 'do not' and *ndaá* '(it) is not', the preposition *'bí* 'of' will precede the possessive pronouns making a polite statement, as in example (59).

- (59) *ndá 'bí eyí ánu ndùku née wá*
 do.not of 2s.POSS [2s].eat food that NEG
 'You do not eat that food.'

When using the negation *ndá* without the preposition *'bí* and a possessive pronoun *amá*, *ené eyí*, *ezé*, or *eyé*, it is a direct command. Then the *ndá* is always found at the beginning of the clause and the clause is concluded by *wá* as in example (60).

- (60) ndá ánu ndụkụ née wá
do.not [2s].eat food that NEG
'Do not eat that food!'

Ndaá 'is not' is the negative form of the verb 'to be' and is another negation word. It can be found anywhere in the clause or sentence. It is again followed by the preposition 'bi'. *Wá* is usually found at the end of the negative clause in this kind of sentences, as in example (61).

- (61) Lomo ndaá 'bi ené kí rokinyi kí Káyina wá
God be.not of 3s.POSS with happiness with Cain NEG
'God is not happy with Cain.'

As can be seen in examples like (13), here repeated for convenience, relative clauses are formed with the relativiser *bi* and the end of clause marker *ní*. In negative relative clauses, the negation word *ngára* is added, as in example (62). Because of *ngára*, the possessive pronoun is not needed to indicate negation early in the clause.

- (13) ìnyìní kí 'yì bándága ga **bi** ndịsini bándá wo **ní** ndéré-ye.
get.up.3p with person his.guards PL REL stay.3p look.after 3s ECL go-3p
'..., he got up with his guards, who used to look after him, and they went.'

- (62) ídí lígíçí 'yì ga **bi ngára** iliní lịkpí cika **wá ní**
[2s].let forget person PL REL **who.be.not** want.3p speak story NEG ECL
'Forget those people who do not want to tell stories.'

In most cases, two or more of the Baka negation words (*ndaá*, *wá*, *ndaá*, *ngára*, *lolụ* - 'any more') can be found in a clause or sentence, such as (62) above.

5 Conclusion

The pronominal elements in Baka are comprised of subject reference markers on the verb and independent pronouns which may also occur in attached forms (enclitics). A pronoun paradigm is comprised of 6 forms: first, second, and third person each with singular or plural. The independent pronouns differentiate only two cases – subject/object and possessive, except for third person singular. Some subject reference markers on verbs employ tone in addition to affixes.

Baka has a variety of pronouns for third person singular. In third singular, this language not only differentiates subject and object, but also alienable and inalienable possession with possessive pronouns, and has a separate third singular pronoun to mark the participant as the topic in a sentence or as important for the story. Further, Baka employs logophoric pronouns to differentiate the speaker from other participants in semi-direct reported speech, some of which may also be employed in other special cases. The third singular logophoric pronoun is again different from other third singular pronouns, while the plural logophoric pronoun resembles the subject/object pronoun.

Possessive pronouns also play an important role in negative clauses. Namely, negation may be marked by a possessive pronoun after the verb in addition to the negation particle at the end of the clause. The person of this possessive pronoun is always the subject.

Abbreviations

1P/1S	first person plural/singular
2P/2S	second person plural/singular
3P/3S	third person plural/singular
CON	connector
DM	developmental marker (Dooley and Levinsohn, 2000), thematic continuant (Edward B. Manderson, 1984)
ECL	end of clause marker
LOG	logophoric
NEG	negation
NMLZ	nominalizer/nominalization
OBJ	object
PL	plural
POSS	possessive
REL	relativiser
SG	singular
SBJ	subject

References

- Aikhenvald, Alexandra Y. 2011. Semi-Direct Speech in Typological Perspective. In Alexandra Y. Aikhenvald & Robert M.W. Dixon, (eds.), *Language at Large: Essays on Syntax and Semantics*, pp. 327–367. Leiden: Koninklijke Brill NV.
- Clements, George N. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 10.2: 141–177.
- Creissels, Denis, Gerrit J. Dimmendaal, Zygmunt Frajzyngier & Christa König. 2008. Africa as a Morphosyntactic Area. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*, pp. 86–150. Cambridge: Cambridge University Press.
- De Roeck, Marijke. 1994. A Functional Typology of Speech Reports. In Elisabeth Engberg-Pedersen, Lisbeth F. Jakobsen & Lone S. Rasmussen (eds.), *Function and Expression in Functional Grammar*, pp. 331–351. Berlin: de Gruyter.
- Dooley, Robert & Stephen Levinsohn. 2000. *Analysing Discourse: A Manual of Basic Concepts*. Dallas: SIL.
- Gurtong, 2013 – Internet source Jan. 2013. (<http://www.gurtong.net/Peoples/PeoplesProfiles/Baka/tabid/181/Default.aspx>).
- Kilpatrick, Eileen. 2006. Pronoun Sets in the Moru-Madi Languages. In Al Amin Abu-Manga, Leoma Gilley & Anne Storch, (eds.), *Insights into Nilo-Saharan Language, History and Culture. Proceedings of the 9th Nilo-Saharan Linguistics Colloquium Institute of African and Asian Studies, University of Khartoum, 16–19 February 2004*, pp. 263–277. Cologne: Köppe.
- Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2014. *Ethnologue: Languages of the World*, 17. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- Manderson, Edward B. 1984. ZI and its related forms in Baka discourse. *Occasional Papers in the Study of Sudanese Languages* 3: 24–30.
- Persson, Andrew B. 1997. Grouping of the Bongo-Baka languages. *Occasional Papers in the Study of Sudanese Languages* 7: 19–39.

STAMP MORPHS IN CENTRAL SUDANIC LANGUAGES*

Gregory D.S. Anderson

1 Overview

STAMP morphs (Anderson 2012) are portmanteau morphs that simultaneously encode the person/number(/gender) properties of a verbal argument – usually the syntactic subject – together with TAM and polarity categories.¹ In most previous work on African languages, such elements have also been called ‘tensed’ or TAM-encoding pronouns or fused subject/TAM auxiliaries (Anderson 2006, 2011) or the ‘tense-person complex’ (Creissels 2005: 50–51; 55–59). In this paper I offer some comments on the nature and history of STAMP morphs in the languages of the Central Sudanic family (Boyeldieu & Nougayrol 2004, 2008; Boyeldieu, Nougayrol & Palayer 2006) and the varied role these elements have played in the history of the verbal inflectional system of the languages of this family.

In the following sections, an overview of what STAMP morphs are is provided and then a brief presentation is given of the two conjugational types of personal inflection found in Central Sudanic languages – a bound prefix series and an unbound series. In section 3, I offer some evidence that the unbound series is a STAMP morph construction in various individual languages and families within Central Sudanic, and that this likely developed from the fusing of a pronominal marker with an original auxiliary verb. Similar developments are commonly seen in languages from throughout the Macro-Sudan Belt (Güldemann 2008; Anderson 2011). Further, it is clear that within language families found in the Macro-Sudan Belt, bound inflectional prefix series from other language groups generally derive from the fusing of a STAMP morph and following lexical verb. Thus, in section 5 it is argued that the bound inflectional prefix series in Central Sudanic languages also likely derives historically from the fusing of an original STAMP morph and a following lexical verb.

2 Introduction to STAMP morphs in equatorial African languages

STAMP morphs are found in numerous languages in equatorial Africa across a wide number of genetic lineages, Central Sudanic included. In its most basic form, a STAMP morph construction consists of the STAMP morph followed by the uninflected/unmarked form of the verb. All the relevant obligatory morphosyntactic indices that make the construction syntactically and morphologically well-formed and finite are encoded within the STAMP morph. Such a formation can be found in the Benue-Congo language Tarok of Nigeria (1). In Tarok, two segmentally distinct forms of the first singular STAMP morph are found, the perfective/realis in *n* and the irrealis in *mi*; each combine with the basic form of the verb to create different TAM forms:

[TAROKOID PLATEAU]

	<u>Tarok</u>			<u>Tarok</u> (Sibomana 1981: 238)					
(1)	a. <i>n</i>	wá	ù-diŋ	b. <i>mi</i>	wá	a-tí	ipín		
		1.PFV	drink	CLSFR-water		1.IRR	drink	CLSFR-tea	tomorrow
			‘I have drunk the water’				‘I will drink tea tomorrow’		

Another common pattern is represented by Kulango of Côte d’Ivoire (2). Here the perfect and subjunctive forms of the first singular STAMP morph are segmentally identical but tonally distinct.

* Thanks to the members of the audience at the Nilo-Saharan Linguistics Colloquium-11 in Cologne who offered valuable comments during the discussion following the oral presentation of the paper, especially Colleen Ahland, Torben Andersen and Pascal Boyeldieu. I have maintained the original transcription of the source insofar as was possible. Thus *y*- refers to the palatal glide not the IPA symbol for the front rounded high vowel. All errors however remain the author’s responsibility.

¹ STAMP stands for Subject-Tense-Aspect-Mood-Polarity. Evidentiality is here subsumed under Mood.

Ma'di (Blackings & Fabb 2003: 14)

- (6) ɔ-ŋa gbándà rá
 3.DEF-eat cassava AFFRM
 'he ate cassava'

Should Ma'di *ka* be analyzed as a pronominal (cf. Bender 1989, 1992) or as an auxiliary? The answer is that Ma'di *ka* and similar elements are STAMP morphs, which means that they share syntactic properties and historical features of both types of elements, but function as synchronic entities, albeit ones that are embedded within larger constructional frameworks or projections syntagmatically and which stand in paradigmatic opposition to other functionally similar forms. In Ma'di's sister language Moru, according to the analysis of Andersen (1986), the cognate to Ma'di *ka* thus retains an internally definable structure that consists of a pronominal marker in *k-* and an auxiliary in *-a*. This element thus remains embedded within a synchronic auxiliary verb construction, in a familiar and typical AUX-headed configuration (Anderson 2006, 2011), with subject marking on the auxiliary and the lexical verb appearing in a construction-specific dependent form *ù* (with *-mú* at least):

- | | |
|---|--|
| <p><u>Moru</u></p> <p>(7) a. kòcé mū tē
 dog 3.run COMPL
 'the dog ran'</p> | <p><u>Moru</u> (Andersen 1986: 31)</p> <p>b. kòcé k-á ù-mú
 dog 3-AUX DEP-run:PRTCPL
 'the dog is running'</p> |
|---|--|

Indeed, various researchers in their analyses of the structure or history of individual African languages or language families have been troubled by how exactly to analyze such STAMP morphs which arise from various phonological processes of fusion and erosion affecting pronominals and auxiliary elements. Thus, when discussing the development of what I here call STAMP morphs in Benue-Congo, Babaev (2010b: 35) states that "... various phonological processes of merging person markers with predicative markers of tense, aspect, modality and polarity have made the situation in many languages obscure".⁴ Similarly, when speaking of the system in the Leko-Nimbari language Zing Mumuye, Shimizu (1983: 101) cautions us that "(t)he surface differences in subject pronouns are in fact due to the TAM markers, which are realized on them or contracted with them".

While STAMP morphs occur in a few restricted and definable areal clusters apart from the Macro-Sudan belt (see section 4 below), e.g., northern South America or eastern and insular Papua New Guinea (Anderson 2006), they are particularly common in equatorial Africa (Anderson 2011). Moreover, some stage in the development of STAMP morphs occur in the form of the so-called pronominal contractions in Colloquial Spoken American English.

Colloquial Spoken American English

- (8) a. I gotta think about it [I AUX Verb Prep ProN]
 b. I'll think about it [1.FUT Verb Prep ProN]
 c. I'm thinking about it [1.PROG.PRS Verb<PROG> Prep ProN]
 d. I've thought about it [1.PRF Verb<PRF> Prep ProN]

3 STAMP morphs and the two inflectional series of Central Sudanic

There is strong evidence that one must reconstruct two separate sets of inflectional paradigms for Proto-Central Sudanic. One such inflectional series, the perfective or 'definite' series, is often phonologically bound to the verb, appearing in most languages as a series of inflectional prefixes.⁵ The other inflectional series, the imperfective series, is, on the other hand, often marked with a free-standing or at least a phonologically less integrated (or quasi-free-standing) pronominal marker. I

⁴ Cf. Babaev (2010a) and Vydrin (2006) for similar statements on Mande.

⁵ In the Bongo first, second and third plural, and in some paradigms in Baka, both of the Sara-Bongo-Bagirmi family, suffixed agreement forms may be found. Mamvu (Vorbichler 1971; cf. Vorbichler 1965) of the Mangbutu family might also be developing a suffixal series, see note 5 below.

argue below that both likely have their origins in STAMP morphs (or STAMP morph constructions), and that while the imperfective series is often synchronically free-standing, and thus may transparently remain encoded by a STAMP morph construction, the perfective series has rather been fused with the following verb into a prefixal inflectional series. Such an opposition between a bound ‘definite’ series and an unbound ‘indefinite’ series can be found across languages representing various families or branches of Central Sudanic, e.g., in Bongo of the Sara-Bongo-Bagirmi family (9), Lugbara (Crazzolaro 1960; Barr 1965; Weber 1994) of the Moru-Ma’di family (10), Lendu of the Lendu family (11), and Ok’ebu (12), which has been classified – perhaps inaccurately – as a member of the Mangbutu family.⁶ This suggests strongly that we should reconstruct this particular configuration for Proto-Central Sudanic, although only traces remain of the original opposition in individual branches and languages.

[SARA-BONGO-BAGIRMI]

- | | | | |
|-----|--|----|--|
| | <u>Bongo</u> * <i>mi-</i> < <i>m-i</i> | | <u>Bongo</u> (Tucker & Bryan 1966: 75) |
| (9) | a. mi-bi | b. | ma bi |
| | 1.DEF-give | | 1.INDEF give |
| | ‘I give, gave’ | | ‘I am giving’ |

[MORU-MA'DI]

- | | | | |
|------|--------------------------------|----|--|
| | <u>Lugbara</u> | | <u>Lugbara</u> (Tucker & Bryan 1966: 46–7) |
| (10) | a. á-tso mvá | b. | ma mvá tso |
| | 1.DEF ⁷ -beat child | | 1.INDEF child beat |
| | ‘I beat the child’ | | ‘I am beating the child’ <i>má-</i> form? |

[LENDU]

- | | | | |
|------|-------------------|----|--|
| | <u>Lendu</u> | | <u>Lendu</u> (Tucker & Bryan 1966: 46) |
| (11) | a. má-drr mbi | b. | má mbi drf |
| | 1.DEF-pull rope | | 1.INDEF rope pull |
| | ‘I pull the rope’ | | ‘I am pulling the rope’ |

[MANGBUTU]

- | | | | |
|------|----------------------|----|---|
| | <u>Oke’bu</u> | | <u>Oke’bu</u> |
| (12) | a. l-ómá ùnzu | b. | láá unzú òma |
| | 2.DEF-beat child | | 2.INDEF child beat |
| | ‘you beat the child’ | | ‘you are beating the child’ (Tucker & Bryan 1966: 48) |

In addition to the opposition of subject encoding in verbs being divided into a bound/definite/perfective vs. unbound/indefinite/imperfective series, the reader may already be aware, or have noticed in the examples above, that in many branches of Central Sudanic these two series are distinguished by the clausal syntax as well: the definite series takes VO order while the indefinite series correspondingly appears with OV order (13). Primarily this syntactic opposition is true of the

⁶ Constance Kutsch-Lojenga (2013) suggested in her presentation on Mvuba of the Mangbutu group at NSLC-11 that Oke’bu did not fit within Mangbutu but it was also not clear where else it might fit. It does diverge from the usual profile of Mangbutu with respect to the inflectional phenomena discussed here, but given that it is simply archaic in this regard, this fact alone would not necessarily exclude it from Mangbutu *a priori* (there are other reasons for this uncertainty regarding its classification as well of course). Taxonomically, Central Sudanic consists of six family-level branches: Sara-Bongo-Bagirmi [SBB], Kresh, Moru-Ma’di, Mangbetu, Mangbutu, and Lendu. More controversial is whether intermediate nodes exist, such as Western Central Sudanic, consisting of the first two taxa mentioned above, vs. Eastern Central Sudanic, consisting of the remaining four. Regardless of whether these controversial intermediate taxa gain widespread acceptance or not, STAMP morph constructions and the opposition between the two inflectional series project back to the Proto-Central Sudanic level.

⁷ NB: *á- ~ má- ~ m-’* as 1.DEF projects back to Proto-Moru-Ma’di.

languages of the Moru-Ma'di family (Andersen 1984, 1998), Lendu and Mangbutu as well at least, and may thus project back as far as the putative Proto-Eastern Central Sudanic level.⁸

(13)	Series	'Definite'/Perfective	'Indefinite'/Imperfective
	Degree of Integration	Bound (usually prefix)	Free-standing
	Syntax	VO	OV

Such a distribution suggests that an unbound 'indefinite' series and a bound 'definite' series of inflectional markers or conjugational types may reflect inheritances from an original system dating back all the way to the Proto-Central Sudanic level, each with their own corresponding phrasal syntax. The bound or 'definite' series is examined in slightly more detail in section 5. First, however, in section 3 I explore the origin of the unbound 'indefinite' series. This series is indeed more subject to restructuring than the relatively more stable 'definite' series within and across the different branches of Central Sudanic, and it also more clearly and obviously shows that the oldest forms of the 'indefinite' series originated in, or indeed synchronically may remain, STAMP morph formations.

4 Auxiliary verb construction origins of STAMP morphs in C. Sudanic languages

Based on observable cross-linguistic tendencies (Anderson 2006, 2011), known historical origins of STAMP morphs in other equatorial African languages families (Anderson 2012), and on internal evidence within Central Sudanic languages themselves, it seems likely that the unbound 'indefinite' series of pronominal markers arose through the fusing of a pronominal element and a following auxiliary verb, synchronically melded into the unitary element here called a STAMP morph.

Certain languages of the Sara-Bongo-Bagirmi group, in particular those of the Western sub-family (Boyeldieu & Nougayrol 2008; Boyeldieu, Nougayrol & Palayer 2006; Boyeldieu 2006a, 2006b, 2000) possess two distinct conjugational types, a bound 'definite' series and an unbound 'indefinite' series. As mentioned previously, most treatments of this phenomenon, the unbound series is said to be encoded by subject pronouns, as they appear to be in such languages as Bongo (or Lugbara or Lendu above). However in languages like Sara (14), the indefinite series requires an infinitive form of the verb:

[SARA-BONGO-BAGIRMI]

	<u>Sara</u>		<u>Sara</u> (Tucker & Bryan 1966: 75)
(14) a.	m-usa	b.	ma k-usa
	1-eat		1:DEF INF-eat
	'I ate/eat'		'I am eating', 'I eat (HAB)'

In other words, (14b) appears to lack a finite verb altogether, and to consist of a subject pronoun and an infinitive verb form. However, if one analyzes the subject 'pronoun' as a STAMP morph instead, and thus containing an erstwhile auxiliary that encoded TAM categories, then the use of an infinitive complement to an auxiliary head in this type of constructional configuration may thus be expected, or at least is in no sense aberrant. Other Western SBB languages show similar phenomena, e.g., Kenga, where the first singular future STAMP morph *m̄* likewise requires a following verb to be in the cognate *k*-infinitive form (15c):

	<u>Kenga</u>		<u>Kenga</u>		<u>Kenga</u>
(15) a.	m-ḡsḡ	b.	m-ḡs-gà	c.	m̄ k-ḡsḡ
	1-eat:PRS		1-eat-PRF		1.FUT INF-eat
	'I eat'		'I have eaten'		'I will eat'
	(Neukom 2010: 104; 116–117)				

⁸ This pattern might therefore be evidence for considering the controversial Eastern Central Sudanic to be a valid intermediate taxonomic node. This is a complex subject and must remain the topic of a different study.

Sara's close sister language Ngambay reflects a similar situation. Here, one STAMP morph construction is formally cognate with Sara and similarly triggers the infinitive verb form, but like Kenga, the construction marks future, not progressive action.

- | | | | |
|---------|-----------------------|----|---------------------------------|
| | <u>Ngambay</u> | | <u>Ngambay</u> |
| (16) a. | mā k-ào àl ngà | b. | á k-ùsā né ngà uà |
| | 1.FUT INF-go NEG REP | | 2.FUT INF-eat thing REP Q |
| | 'I will not go again' | | 'you are already going to eat?' |
| | (Vandame 1963: 118) | | |

More distantly related but still Western SBB, Kabba (17) shows a similar construction to that in Ngambay:

- Kabba (Moser 2004: 220)
- (17) má k-àw lò t́ ááng
 1:FUT INF-go place LOC NEG
 'I shall go nowhere'

Other SBB languages show (semi-)cognate structures with infinitive marked verbs in the 'indefinite' series which contrast with bound forms of the 'definite' series; these, however, use a different form of the infinitive, such as *y-* in the Baka variety (18b) described by Tucker & Bryan (1966) or Santandrea (1976),⁹ or *l-* in the Yulu (19) of Santandrea (1970; cf. Boyeldieu 1987)

- | | | | |
|---------|---------------------|----|---------------------------------------|
| | <u>Baka</u> | | <u>Baka</u> (Tucker & Bryan 1966: 75) |
| (18) a. | m-áne yí | b. | má y-ane yí |
| | 1.DEF-eat thing | | 1.INDEF INF-eat thing |
| | 'I ate (something)' | | 'I am eating (something)' |

- Yulu (Santandrea 1970: 25)
- (19) ma le'ε
 1:FUT INF:go
 'I shall go'

Kresh (Santandrea 1976; Brown 1991a, 1991b, 1994) shows a similar pattern to the one found in Sara-Bongo-Bagirmi languages, with bound definite series but the indefinite series consisting of an unbound STAMP morph followed by a verb in the infinitive form (20).¹⁰

⁹ But not the interesting Baka variety described by Waag and Phodunze (2013), which has a restructured agreement system that now uses suffixes or enclitics in all persons and numbers in the bound conjugation attached to a clause-initial discourse narrative particle *zi-* 'so, then, and...' (ii), and not just optionally in the conjugation of 1PL/2PL forms in the 'definit' series, as in Bongo (Santandrea 1963; Boyeldieu & Nougayrol 2004), see (iii).

- | | | | | | |
|----------|----------------------------------|----|---------------------------------|----|------------------|
| | <u>Baka</u> | | <u>Baka</u> | | <u>Baka</u> |
| (ii) a. | nderé-ni | b. | zi-ye ndéré(-ye) | c. | ani/niyí ndéré |
| | go.3.PFV-3PL | | NARR-3PL go-(3PL) | | ASP.3PL go |
| | 'they went' | | 'they went' (in narrative text) | | 'they are going' |
| | (Waag & Phodunze 2013-handout) | | | | |
| | <u>Bongo</u> | | <u>Bongo</u> | | |
| (iii) a. | j-óŋ ~ òŋ-jé | b. | h-óŋ ~ òŋ-hé | | |
| | 1PL-eat eat-1PL | | 2PL-eat eat-2PL | | |
| | 'we eat' | | 'you all eat' | | |
| | (Boyeldieu & Nougayrol 2004: 25) | | | | |

¹⁰ Again, this may be an innovation that could lend support to the putative intermediate Western-Central Sudanic taxon.

[KRESH]

- | | | | | | |
|---------|-------------------|-------|---|-----------|-------|
| | <u>Kresh</u> | | <u>Kresh</u> | | |
| (20) a. | m-omò | nòmò | b. ma | y-òmò | nòmò |
| | 1-drink | drink | 1.INDEF | INF-drink | drink |
| | 'I drink a drink' | | 'I am drinking' (Tucker & Bryan 1966: 75) | | |

There are traces of this same proto-system found in the Eastern Central Sudanic languages as well. Some forms of the Mangbutu language Oke'bu verb appear in a dependent-form marked by low-toned *ù-* in the indefinite series. This too suggests that the 'indefinite' series 'pronominal' should be analyzed as a STAMP morph containing an original auxiliary element.

[MANGBUTU]

- | | | | | | |
|---------|--------------------|-------|--|-------|----------|
| | <u>Oke'bu</u> | | <u>Oke'bu</u> | | |
| (21) a. | m-ómá | ònzò | b. máá | ònzó | ù-má |
| | 1-VWL:beat | child | 1:INDEF | child | DEP-beat |
| | 'I beat the child' | | 'I am beating the child' (Tucker & Bryan 1966: 46) | | |

Note that it is certainly plausible and indeed even likely that the floating low tone projected in the Ma'di 'indefinite' series is a relic of a low-tone marked dependent verb form originally marked by a low-tone dependent (or quasi-infinitive) prefix; see (22), which repeats (5) above.

- | | | | | | |
|---------|-----------------------------|---------|--------------|------------------------|--------------|
| | <u>Ma'di</u> | | <u>Ma'di</u> | | |
| (22) a. | ká | gbándà | `na | b. kó | `na-ʔa |
| | 3SG | cassava | NPST:eat | 3SG | NPST:eat-OBJ |
| | 'he eats/is eating cassava' | | | 'he eats/is eating it' | |
| | (Blackings & Fabb 2003: 13) | | | | |

STAMP morphs generally stand in formal opposition with overt contrastive person deictics (pronouns), with the latter typically being longer in form and contrastively intonated and used primarily in specific types of informational focus discourse contexts (like English *as for me, I...* or French *moi, je...* etc.), while the former are typically shorter and occur widely in all contexts.

5 STAMP morphs as a feature of the Macro-Sudan Belt

As demonstrated eloquently by Güldemann (2008), a number of features that cross-cut genetic linguistic boundaries can be found in a large band that stretches west to east across equatorial Africa that he called the Macro-Sudan Belt, with eastern Central Sudanic languages representing the easternmost extension of this macro-area. In Güldemann's scheme, Central Sudanic languages are listed within two not necessarily related genera: his Bongo-Bagirmi and Moru-Mangbetu basically correspond to the proposed Western and Eastern intermediate nodes of Central Sudanic.

STAMP morph constructions are widely used across the languages of the Macro-Sudan Belt (Anderson 2011), found for example in as disparate a set of languages as Angas of the West Chadic family (Burquest 1986) and Wobé of the Kru family:

[WEST CHADIC]

- Angas (Burquest 1973/1980: 38/ANG 4)
- | | | |
|------|---------------|------|
| (23) | ɲâ: | ji |
| | 1.COMPL | come |
| | 'I have come' | |

[KRU]

- | | | | |
|---------|----------------------------------|----|-----------------------------------|
| | <u>Wobé</u> | | <u>Wobé</u> |
| (24) a. | ɛ ² gyi ³² | b. | ma ² gyi ³² |
| | I.PST come | | I.NPST come |
| | ‘I have come’ | | ‘I am coming’ |
- (Hofer & Link 1973/1980: WOB 3)

The forms in (23–24) encode all TAM categories in a portmanteau manner with the person/number features of the subject used together with an unmarked verb. As with auxiliary verb constructions more widely, the verb can appear in an unmarked form, such as with English modals or the future (25a), but more typically the functional semantics of auxiliary verb constructions are encoded by the combination of a particular auxiliary verb together with a semi-inflected or construction-specific (often dependent, semi-finite or non-finite) form of the verb, as with the English progressive, perfect and passive formations, seen in (25b) – (25d).

- | | | | |
|---------|--------------------|----|---------------------------|
| | <u>English</u> | | <u>English</u> |
| (25) a. | I will go | b. | I am going |
| | I FUT go | | I I:AUX go:PRTCPL.I |
| | ‘I will go’ | | ‘I am going’ |
| | <u>English</u> | | <u>English</u> |
| c. | I have gone | d. | she was seen |
| | I AUX go:PRTCPL.II | | she AUX:PST see:PRTCPL.II |
| | ‘I have gone’ | | ‘she was seen’ |

Thus, because STAMP morphs derive from the fusing of pronominal markers with former auxiliaries, it should come as little surprise that many languages in the Macro-Sudan Belt require specific dependent (often described as ‘aspectual’) forms of the verb to be used in combination with STAMP morphs, i.e., in a constructional configuration. This latter type of formation can be found in numerous families within the Niger-Congo phylum for example, e.g. the Senufic language Nafanra, the Gbaya language ‘Bozom, or Non of the Cangin (North Atlantic) family.

[SENUFIC]

- | | | | |
|---------|----------------|----|-----------------|
| | <u>Nafanra</u> | | <u>Nafanra</u> |
| (26) a. | me pan | b. | ni paan |
| | I.IRR come | | I.RLS come:IPFV |
| | ‘I will come’ | | ‘I am coming’ |
- (Jordan & Jordan 1975/1980 NAF: 3)

[GBAYA]

- | | | | |
|---------|-------------------|----|------------------|
| | <u>‘Bozom</u> | | <u>‘Bozom</u> |
| (27) a. | ʔā ré | b. | má rè |
| | he.RLS enter.IPFV | | he.IRR enter.PFV |
| | ‘he enters’ | | ‘he will enter’ |
- (Moñino 1995: 159)

[CANGIN ATLANTIC]

- | | | | |
|---------|----------------|----|-----------------|
| | <u>Non</u> | | <u>Non</u> |
| (28) a. | ma hayɛn | b. | mɛ haya |
| | I.PRF come:PRF | | I.PRG come:IPFV |
| | ‘I have come’ | | ‘I am coming’ |
- (Pichl 1973/1980 NON: 3)

6 Bound prefix pronominal series from fusing of STAMP constructions

As alluded to before, not only does the unbound ‘indefinite’ series in Central Sudanic languages likely represent STAMP morph constructions (historically at least), so too does the bound ‘definite’ prefix series likewise similarly derive historically from the univerbation of a STAMP morph together with the following verb. Indeed, such a trend towards fusion and synthesis in the creation of inflectional prefix series is widespread among the languages of the Macro-Sudan Belt, and Central Sudanic languages are no exception in this regard.

To exemplify this process of univerbation and creation of an inflectional series of prefixes, first consider the present and past tense forms of the first singular inflections in the Biu-Mandara Chadic language Mbuko. An imperfective STAMP morph construction realizes the progressive present (29b), while a fused form of the original perfective STAMP morph – now realized as an inflectional prefix series – forms the anterior (29a); note that the verb stem appears in different construction-specific tonal forms in each of these Mbuko formations.

[BIU-MANDARA CHADIC]

- | | | | | |
|---------|--------------------|----|-----------------|--|
| | <u>Mbuko</u> | | <u>Mbuko</u> | |
| (29) a. | nə-zlāmbál | b. | nī zlāmbāl | |
| | 1PRF/ANT-throw:ANT | | 1.IMPF throw | |
| | ‘I threw’ | | ‘I am throwing’ | |
| | (Gravina 2001: 7) | | | |

A similar distribution is found in the unrelated Burak language (30), which belongs to the Bikwin-Jen family within the Gur-Adamawa complex of Atlantic-Congo:

[BIKWIN-JEN]

- | | | | | |
|---------|----------------------------|----|---------------------|--|
| | <u>Burak</u> | | <u>Burak</u> | |
| (30) a. | n-táŋ-lénú | b. | mī ka táŋ-é | |
| | 1-eat-PRF | | 1.NPST PROG eat-DEP | |
| | ‘I have eaten’ | | ‘I am eating’ | |
| | (Jungrathmayr 1968/9: 203) | | | |

Thus, not only do STAMP morphs underlie both inflectional series found in these Afroasiatic and Niger-Congo languages, but like Central Sudanic languages, they show a bound series in the ‘definite’/perfective and an unbound series in the ‘indefinite’/imperfective. Further, as the reader may have noticed, the two series exemplified above show cognate elements in the person-marking systems among Burak in (30), Nafaanra in (20), and Tarok in (1), but while the definite series is found as prefixes in Burak, they remain free-standing STAMP morphs in Tarok and Nafaanra. Indeed, the phenomenon of the fusing of pronominal markers and auxiliaries into verbal prefixes has been suggested for more than a century in West African linguistics, as Seidel already in (1898: 211) alerted us to the fact that in Kwa languages of Togo “die verbalen Präfixe, wahrscheinlich Reste ehemaliger Hilfsverben, verschmelzen nicht selten mit den Pronominalpräfixen zu einer Silbe” (the verbal prefixes, most likely remnants of former auxiliary verbs, meld together not infrequently with pronominal prefixes into a single syllable).

To be sure, various dialects or sister-lects of the Kwa language Nchumuru of Ghana vary along this precise parameter of univerbation of the STAMP morph with the following verb into inflectional prefix series. Thus, in Bejamso-Grubi Nchumuru (31a–b), both constructions appear to use free-standing STAMP morphs, while in Banda Nchumuru (32a–b), both series appear as bound inflectional prefix elements, reflecting a perfective/realis vs. imperfective/irrealis opposition (Anderson 2012).

[KWA]

- (31) a. Bejamso-Grubi Nchumuru
 màá bà
 I.PRF come
 ‘I have come’
 (Price 1975/1980: NCH 5)
- b. Bejamso-Grubi Nchumuru
 mĕ bà
 INPST come
 ‘I am coming’. ‘I will come’
- (32) a. Banda Nchumuru
 mà-ba
 IPST:PRF-come
 ‘I have come’
 (Cleal 1973b/1980: NCH 4)
- b. Banda Nchumuru
 mi-ba
 IPRS/FUT-come
 ‘I am coming’; ‘I will come’

Two closely related varieties that show different diachronic developmental stages of this sort can be seen in Central Sudanic as well, even with what is historically the ‘indefinite’ series. In other words, the trend towards synthesis appears to be ongoing in the region. Compare in this regard the following forms for ‘I don’t know’ in the closely related lects Furu and Bagiro as described by Boyeldieu (1990), where Bagiro appears to have a fused form and Furu a free-standing one:

- (33) Furu
 mí gáli gɔ
 I know NEG
 ‘I don’t know’
 (Boyeldieu 1990: 91)
- (34) Bagiro
 mú-gá⁺li gɔ
 I-know NEG
 ‘I don’t know’

Based on these observations, I suggest that the ‘definite’ series of inflectional prefixes in Central Sudanic (35 – 37) also historically reflect unverbated STAMP morph constructions, these latter thus belonging to a pre-Proto-Central Sudanic stage.

- Bongo (Tucker & Bryan 1966: 75) *mi- < m-i
- (35) mi-bi
 1.DEF-give
 ‘I give, gave’

- Lendu (Tucker & Bryan 1966: 46)
- (36) má-drr mbi
 1.DEF-pull rope
 ‘I pull the rope’

- Oke’bu (Tucker & Bryan 1966: 48)
- (37) l-ómá ùnzu
 2.DEF-beat child
 ‘you beat the child’

The ‘definite’ series of bound inflection can also be used with auxiliaries of course in the Mangbetu language Meje, such as in the following typologically and historically interesting form:

- Meje (McKee 1991: 168)
- (38) m-á ú m-óólí-a
 1-AUX LOC 1-walk-NPST.I
 ‘I am walking’

This form consists of a person-marked auxiliary followed by a locative-headed lexical complement phrase, the verb of which itself bears finite morphology.¹¹ *Meje* is typical of Mangbetu languages in reflecting only the definite series, or in other words, Mangbetu languages have only the historical bound inflectional series of pronominal markers, even if they are realized within a structure that looks very much like a potential STAMP morph, e.g., attached to an auxiliary *-a*, as in 1ST.SG *m-á* – one nearly reconstituting the indefinite series marker in this instance.¹² These forms in turn can be drawn into the verb as prefix conjugational markers and thus the process may begin anew. In other words, although historically belonging to the bound definite series, new forms can be created combining these markers with phonologically light auxiliary forms (like *-a*), that yield something very much like the historical unbound indefinite series. These prosodically weak functional elements can in turn be drawn into the verb stem and create new, historically secondary formations that in turn appear very similar in form to the archaic, historically bound definite series. In this manner, the cyclic process can be created and continued indefinitely.

In Mangbetu languages there appears to be evidence for two different inflectional series historically, but the system has been reinterpreted as having a partly phonological basis. Thus the

¹¹ Central Sudanic languages have auxiliary structures of various formal sub-types. Among closely related languages, these may even reflect variation with the same auxiliary in the same function in one language (Anderson 2011). For example, compare the following forms of the progressive in various Gula lects of the SBB family. In the Gula Sara form in (iv), the auxiliary requires a subject-marked complement, while in Gula Zura, the progressive auxiliary requires an infinitive-marked form (v). Gula Méré on the other hand (vi) shows both patterns optionally, either the doubly-subject marked form (vi-a) or the infinitive marked pattern (vi-b).

Gula Sara (Nougayrol 1999: 137)

- (iv) zō-ndō z-ūsā ī ngá
 1PL-AUX 1PL-eat EXCL thing
 ‘we are eating’

Gula Zura (Nougayrol 1999: 129)

- (v) mō-nā kās cá
 1-AUX INF:eat thing
 ‘I am going to eat’

- | | | | |
|---------|--|----|--|
| | <u>Gula Méré</u> | | <u>Gula Méré</u> |
| (vi) a. | mō-ndō m-ūsā jō
1-AUX 1-eat thing
‘I am eating’
(Nougayrol 1999: 137) | b. | mō-ndō kūsá jō
1-AUX INF:eat thing
‘I am eating’ |

Mbay of the Sara branch shows yet another variation pattern. In Mbay with a first person subject, the cognate construction involves either a doubly marked form with subject marking on both the lexical verb and the auxiliary verb (vii-b), or one in which only the lexical verb appears with subject inflection (vii-a). In the first plural on the other hand, the subject prefix is doubled or absent on the auxiliary verb, but the number suffix appears only on the lexical verb component of the construction (viii) in either instance.

- | | | | |
|-----------|--|----|---|
| | <u>Mbay</u> | | <u>Mbay</u> |
| (vii) a. | ndi m̄-sá yáá
AUX 1-eat food
‘I am/was eating’
(Keegan 1997: 69) | or | b. m̄-ndi m̄-sá yáá
1-AUX 1-eat food |
| (viii) a. | ndi kò-sà-n̄ yáá
AUX 1PL-eat-PL food
‘we are/were eating’
(Keegan 1997: 69) | or | b. kò-ndi kò-sà-n̄ yáá
1PL-AUX 1PL-eat-PL food |

¹² Indeed, this subject prefix plus auxiliary *m-á* reconstitutes and virtually mimics the indefinite series STAMP morph in SBB languages, but may be a pseudo-archaism – something that recreates or only appears to be an archaic form but is in fact regularly but secondarily derived and not a direct inheritance of an older form.

second singular future marker in Mangbetu and Meje may be realized differently depending on whether they occur after vowel-initial or consonant-initial verb stems:¹³ This represents a restructuring of the system where the trigger had grammatical/functional conditioning, to one with phonological conditioning; such developments are most likely due to conventional specialization and re-interpretation with specific original auxiliary elements, now eroded to Ø.¹⁴

	<u>before vowel-initial stem</u>	<u>before consonant-initial stem</u>
(39) Mangbetu:	<i>ni-'</i>	<i>mú-</i>
(40) Meje:	<i>ni-'</i>	<i>i/ú</i>
<u>Mangbetu</u>		<u>Meje</u>
(41) ni-ési-(a)	(42) ni-ési-a	
2.PRS/FUT-do-TAM	2.PRS/FUT-do-TAM	
'you (will) do'	'you (will) do'	
(Larochette 1958: 106)		
<u>Mangbetu</u>		<u>Meje</u>
(43) mú-ta	(44) ú-ta	
2.PRS/FUT-carry	2.PRS/FUT-carry	
'you (will) carry'	'you (will) carry'	
(Larochette 1958: 106–7)		

Like Mangbetu, most Mangbutu languages also have only the bound/'definite' series (as suggested by Tucker & Bryan 1966), but nevertheless reflect traces of two different series. In Mangbutu languages this is realized mainly through variation in the following TAM vowel appearing between the subject prefix and the verb stem. The vowels in these forms likely reflect now fused but originally distinct STAMP morphs, with the different vowels encoding distinct TAM categories (in turn of course reflecting what were originally distinct TAM auxiliaries historically). Indeed, one can posit some of these formations back to the Proto-Mangbutu level, e.g. **m-a* 1.PRS vs. **m-ú* 1.FUT.

(45) Lese	<i>m-á</i> 1.PRS.PROG	vs.	<i>m-ó</i> 1.FUT
(46) Mamvu	<i>m-a</i> 1.PRS	vs.	<i>m-ú</i> 1.FUT

LeseLese

¹³ Meje complicates this a bit since if one examines the full set of paradigms offered by Larochette (1958) one finds some places where this distribution appears to break down, so one finds both *i-ŋ-* and *mú-ŋ-* in the PRS.PRF, but *i-b-* in the PST.PRF.

¹⁴ Another trace of what were likely two distinct series originally (or historically) in Mangbetu inflection is realized synchronically as two different tonal projections onto the following verb stem of the same segmentally identical inflectional prefix, such as the projection of a high tone in the subjunctive vs. the projection of a middle tone with the present/future with first singular subjects in Meje.

(ix) a.	<u>Meje</u> [má-+H]	b.	<u>Meje</u> [má-+M]
	má-tá		má-ta
	1.SBJNCTV-carry		1.PRS/FUT-carry
	(Larochette 1958: 107; 111)		

- (47) a. m-á-dzò kòdí à-nó b. m-ó-dzá kòdí à-nò
 1-TAM-AUX meat DEP-eat 1-TAM-AUX meat DEP-eat
 ‘I am eating meat’ ‘I shall be eating meat’
 (Tucker & Bryan 1966: 51)

As a last word of caution, note that it is possible if not likely that more Central Sudanic languages may have bound inflectional prefixes than has to date been identified. As Creissels et al. (2008: 93) have cautioned “(m)any descriptions of African languages do not identify pronominal markers appropriately, treating them as independent words”. For example, in the Boko/Busa cluster of Eastern Mande, the pronominal markers are orthographically written separately as individual words, but phonologically are bound (Jones 1998). Resolving this issue in Central Sudanic languages alas must remain the subject of a different study however.¹⁵

7 Summary

Central Sudanic languages often have two inflectional series, one a bound/‘definite’/perfective series and one an unbound/‘indefinite’/imperfective series, traditionally viewed as being encoded by a subject ‘pronoun’. Data from a range of Central Sudanic languages suggest that this subject ‘pronoun’ is something more than just a pronoun, but often exhibits functional semantics and syntactic projections more typical of auxiliary verbs. It is suggested that these elements in the ‘indefinite’ series be re-interpreted such that they are neither pronominals nor auxiliaries *per se*, but rather a kind of in-between category here called STAMP morphs – portmanteau Subject/Tense-Aspect-Mood/Polarity elements that result typically from the fusing of a pronominal marker and an auxiliary diachronically. Such STAMP morphs are in fact commonly attested in genetic units across the Macro-Sudan Belt of equatorial Africa. In these languages, one frequently finds in addition to the free-standing STAMP morph series, a phonologically bound set of inflectional prefixes (or clitics, depending on the element and language in question) that derives from the fusing of a STAMP morph with the following verb. Indeed, when one series is bound and the other free-standing, it is always the ‘definite’/perfective/realis series that is bound, and the ‘indefinite’/imperfective/irrealis series that is free-standing in multiple unrelated lineages. It is therefore proposed here that the bound/‘definite’/perfective series in Central Sudanic languages likewise derived from the fusing of an original STAMP morph construction in the history of these languages. Some branches of Central Sudanic only preserve the bound series (e.g. Mangbetu, Mangbutu) but with traces of an original contrast, often restructured or collapsed. Other individual Central Sudanic languages (e.g. Bagiro) have started the process anew such that the historical (unbound) indefinite series is also becoming bound as well, and new prefixal agreement systems appearing. The process is thus always potentially recycling itself. This study is another step towards unraveling the many complex issues in the history of the inflectional systems of the Central Sudanic languages.

¹⁵ As shown by a map presented by Anderson (2012), the more eastern the location of a language or branch within the different genetic units, the more likely it is to have bound prefixes rather than free-standing elements.

Abbreviations

1	1 st person	LOC	locative
2	2 nd person	NARR	narrative
3	3 rd person	NEG	negative
AFFRM	affirmative	NPST	non-past
ANT	anterior	OBJ	object
AOR	aorist	PFV	perfective
ASP	aspect	PL	plural
AUX	auxiliary	Prep	preposition
CLSFR	classifier	PRF	perfect
COMPL	completive	PROG	progressive
CONT	continuous	ProN	pronoun
DEF	definite	PRS	present
DEP	dependent	PRTCPL	participle
DIR	directive	PST	past
EXCL	exclusive	Q	question
FUT	future	REP	reported
IMPF	imperfect	RLS	realis
INDEF	indefinite	SBJNCTV	subjunctive
INF	infinitive	SG	singular
IPFV	imperfective	TAM	tense-aspect-mood
IRR	irrealis	VWL	vowel

References

- Andersen, Torben. 1984. Aspect and word order in Moru. *Journal of African Languages and Linguistics* 6: 19–34.
- Andersen, Torben. 1986. Verbal inflexion in Moru. *Afrika und Übersee* 69: 19–43.
- Anderson, Torben. 1998. Verb-coding in Madi relative clauses. *Sprachtypologie und Universalienforschung* 51.4: 295–326.
- Anderson, Gregory D. S. 2012. STAMP morphs in the history of Benue-Congo and Niger-Congo conjugation. Presented at Niger-Congress, Paris, September 2012.
- Anderson, Gregory D. S. 2011. Auxiliary verb constructions in the languages of Africa. *Studies in African Linguistics* 40.1–2: 1–409.
- Anderson, Gregory D. S. 2006. *Auxiliary verb constructions*. Oxford: Oxford University Press.
- Babaev, Kirill. 2010a. Person marking in South-West Mande languages: a tentative reconstruction. *Mandenkan* 46: 3–48.
- Babaev, Kirill. 2010b. Reconstructing Benue-Congo person marking II. *Journal of Language Relationship/Voprosy Jazykovogo Rodstva* 4: 1–45.
- Barr, L. I. 1965. *A Course in Lugbara*. Nairobi: East African Literature Bureau.
- Bender, M. Lionel. 1989. Nilo-Saharan pronouns/demonstratives. In M. Lionel Bender (ed.), *Topics in Nilo-Saharan Linguistics*, pp. 1–34. Hamburg: Buske.
- Bender, M. Lionel. 1992. Central Sudanic segmental and lexical reconstructions. *Afrikanistische Arbeitspapiere* 29: 5–61.
- Blackings, Mairi & Nigel Fabb. 2003. *A Grammar of Ma'di*. Berlin: Mouton de Gruyter.

- Boyeldieu, Pascal. 1987. *Les langues Fer (Kara) et Yulu du Nord Centrafricain: esquisses descriptives et lexiques*. Paris: Librairie Orientaliste Paul Geuthner.
- Boyeldieu, Pascal. 1990. L'extension oubanguienue des langues "sara". *Journal des Africanistes* 60.1: 77–106.
- Boyeldieu, Pascal. 1991. The tonal identity of Gbaya ("Bagiro") verbs. In Franz Rottland & Lucia A. Omondi (eds.), *Proceedings of the 3rd Nilo-Saharan Linguistics Colloquium*, pp. 37–52. Hamburg: Buske.
- Boyeldieu, Pascal. 2000. *La langue bagiro (République Centrafricaine). Systématique, textes et textique*. Frankfurt: Lang.
- Boyeldieu, Pascal. 2006a. Présentation des langues Sara-Bongo-Baguirmiennes. Paris: CNRS-LLACAN. <http://sumale.vjf.cnrs.fr/SBB/> (accessed 11 April 2013).
- Boyeldieu, Pascal. 2006b. Reflexes of a labiovelar series in Central Sudanic. In Al-Amin Abu-Manga, Leoma Gilley & Anne Storch (eds.), *Insights into Nilo-Saharan Language, History and Culture: Proceedings of the 9th Nilo-Saharan Linguistics Colloquium, Khartoum, February 16–19, 2004*, pp. 129–151. Cologne: Köppe.
- Boyeldieu, Pascal. 2007. Compound verbs and modalities of process in Yulu (Central Sudanic). In Doris Payne & Mechthild Reh (eds.), *Advances in Nilo-Saharan Linguistics: Proceedings of the 8th Nilo-Saharan Linguistics Colloquium, Hamburg, August 22–25, 2001*, pp. 25–39. Cologne: Köppe.
- Boyeldieu, Pascal. (forthcoming). Entre lac Albert et lac Tchad: l'expansion sara-bongo-baguirmiennne. Migrations et mobilité spatiale dans le bassin du Lac Tchad. *Actes du 13e Colloque International du Réseau Mégatchad, Maroua, October 27 – November 4, 2005*.
- Boyeldieu, Pascal & Pierre Nougayrol. 2004. Les marques personnelles des langues SBB: traits systématiques et perspectives historiques. In Dymitr Ibriszimow & Guillaume Segerer (eds.), *Systèmes de Marques Personnelles en Afrique*, pp. 23–42. Louvain/Paris: Peeters.
- Boyeldieu, Pascal, Pierre Nougayrol & Pierre Palayer. 2006. *Lexique comparatif historique des langues Sara-Bongo-Baguirmiennes*. Paris: CNRS-LLACAN. <http://sumale.vjf.cnrs.fr/SBB/> (accessed 15 May 2013).
- Boyeldieu, Pascal & Pierre Nougayrol. 2008. Les langues soudaniques centrales: essai d'évaluation. In Marianne Bechhaus-Gerst, Dymitr Ibriszimow & Rainer Vossen (eds.), *Problems of Linguistic-Historical Reconstruction in Africa* (SUGIA 19), pp. 9–30. Cologne: Köppe.
- Brown, D. Richard. 1991a. Information focus in Kresh. In M. Lionel Bender (ed.), *Proceedings of the 4th Nilo-Saharan Linguistics Colloquium, Bayreuth, August 2–4, 1989*, pp. 325–345. Hamburg: Buske.
- Brown, D. Richard. 1991b. Noteworthy features of Kresh phonology and orthography. In Franz Rottland & Lucia N. Omondi (eds.), *Proceedings of the 3rd Nilo-Saharan Linguistics Colloquium, Kisumu, August 4–9, 1986*, pp. 53–79. Cologne: Köppe.
- Brown, D. Richard. 1994. Kresh. In Peter Kahrel & René van den Berg (eds.), *Typological Studies in Negation*, pp. 163–189. Amsterdam/Philadelphia: Benjamins.
- Burquest, Donald. 1973. Angas. In Mary E. Kropp Dakubu (ed.), *West African Language Data Sheets 1*, pp. 35–43. Accra: West African Linguistic Society.
- Burquest, Donald. 1986. The pronoun system of some Chadic languages. In Ursula Wiesemann (ed.), *Pronominal Systems*, pp. 71–101. Tübingen: Narr.
- Cleal, Alison M. 1973a/1980a. Krachi. In Mary E. Kropp Dakubu (ed.), *West African Language Data Sheets 1*, pp. 366–378. Accra: West African Linguistic Society.
- Cleal, Alison M. 1973b/1980b. Nchumuru. In Mary E. Kropp Dakubu (ed.), *West African Language Data Sheets 1*. Accra: West African Linguistic Society.
- Crazzolaria, Joseph P. 1960. *A Study of the Logbara (Ma'di) Language*. London: Oxford University Press.

- Creissels, Denis. 2005. A typology of subject and object markers in African languages. In F. K. Erhard Voeltz (ed.), *Studies in African Linguistic Typology*, pp. 43–70. Amsterdam/Philadelphia: Benjamins.
- Creissels, Denis, Gerrit J. Dimmendaal, Zygmunt Frajzyngier & Christa König. 2008. Africa as a morphosyntactic area. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*, pp. 86–150. Cambridge: Cambridge University Press.
- Elders, S. 2007. Complex verb morphology in Kulango (Gur): similarities and dissimilarities with Bantu. In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special Issue on Bantu linguistics (SOAS Working Papers in Linguistics 15)*, pp. 187–200. London: University of London.
- Gravina, Richard. 2001. *The Verb Phrase in Mbuko*. Yaoundé: SIL.
- Güldemann, Tom. 2008. The Macro-Sudan belt: towards identifying a linguistic area in northern sub-Saharan Africa. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*, pp. 151–185. Cambridge: Cambridge University Press.
- Hofer, Verena & Christa Link. 1973/1980. Wobé. In Mary E. Kropp Dakubu (ed.), *West African Language Data Sheets 2*. Accra: West African Linguistic Society.
- Jones, Ross McCallum. 1998. *The Boko/Busa Language Cluster* (Lincom Studies in African Linguistics, 30). Munich: Lincom.
- Jordan, Dean & Carol Jordan. 1975/1980. Nafaara. In Mary E. Kropp Dakubu (ed.), *West African Language Data Sheets 2*. Accra: West African Linguistic Society.
- Jungraithmayr, Herrmann. 1968/1969. The class languages of the Tangale-Waja district. *Afrika und Übersee* 52: 161–204.
- Kropp Dakubu, Mary E. (ed.). 1975–1980. *West African Language Data Sheets 1–2*. Accra: West African Linguistic Society.
- Kutsch Lojenga, Constance. 2013. Mvuba. Paper presented at Nilo-Saharan Linguistics Colloquium 11, May 22–24, Cologne.
- Larochette, Joe. 1958. *Grammaire des dialectes mangbetu et medje, suivie d'un manuel de conversation et d'un lexique* (Sciences de L'homme, 18.) Tervuren: Annales du Musée Royal du Congo Belge.
- McKee, Robert G. 1991. “Here”, “There” “Yonder”, and beyond with Meje aspect. In Franz Rottland & Lucia A. Omondi (eds.), *Proceedings of the 3rd Nilo-Saharan Linguistics Colloquium, Kisumu, August 4–9, 1986*, pp. 165–80. Hamburg: Buske.
- Moñino, Yves. 1995. *Le Proto-Gbaya: Essai de linguistique lompative sur vingt-et-une langues d'Afrique Centrale*. Paris: Peeters.
- Neukom, Lukas. 2010. *Description grammaticale du Kenga (Langue Nilo-Saharienne du Tchad; Nilo-Saharan, 25)*. Cologne: Köppe.
- Nougayrol, Pierre. 1999. *Les parlers Gula (République Centrafricaine, Soudan, Tchad): éléments de grammaire comparative et lexicque*. Paris: CNRS.
- Pichl, Walter J. 1973b/1980b. Non. In Mary E. Kropp Dakubu (ed.), *West African Language Data Sheets 2*. Accra: West African Linguistic Society.
- Price, Norman. 1975/1980. Nchumuru. In Mary E. Kropp Dakubu (ed.), *West African Language Data Sheets (Volume 2)*. Accra: West African Linguistic Society.
- Rottland, Franz & Lucia A. Omondi (eds.). 1991. *Proceedings of the 3rd Nilo-Saharan Linguistics Colloquium. Kisumu, August 4–9, 1986*, pp. 165–80. Hamburg: Buske.
- Santandrea, Stefano. 1956. Brief notes on Biŋa. *Afrika und Übersee* 40: 25–38.
- Santandrea, Stefano. 1963. *A Concise Grammar Outline of the Bongo Language*. Rome: Sodality of St. Peter Claver.
- Santandrea, Stefano. 1970. *Brief Grammar Outlines of the Yulu and Kara Languages*. Rome: Sodality of St. Peter Claver.

- Santandrea, Stefano. 1976. *The Kresh Group, Aja and Baka Languages (Sudan)*. Bologna: Editrice Missionaria Italiana.
- Seidel, August. 1898. Beiträge zur Kenntnis der Sprachen im Togo. *Zeitschrift für afrikanische und oceanische Sprachen* 4: 201–286.
- Shimizu, Kiyoshi. 1983. *The Zing dialect of Mumuye. A descriptive grammar. With a Mumuye-English dictionary and an English-Mumuye index*. Hamburg: Buske.
- Sibomana, Leo. 1981/82. Tarok III: Das Verbalsystem und der Satz. *Afrika und Übersee* 64.2: 237–247.
- Tucker, Archibald N. & Margaret A. Bryan. 1966. *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press.
- Vandame, Charles. 1963. *Le Ngambay-Moundou: Phonologie, grammaire et textes*. Dakar: IFAN.
- Vorbichler, Anton. 1965. *Die Phonologie und Morphologie des Balese (Ituri-Urwald, Kongo)*. Glückstadt: Augustin.
- Vorbichler, Anton. 1971. *Die Sprache der Mamvu*. (Afrikanistische Forschungen, 5). Glückstadt: Augustin.
- Vydrin, Valentin F. 2006. Lichnye mestoimenija v juzhnykh jazykakh mande [Personal pronouns in South Mande languages]. *Trudy instituta lingvističeskikh issledovanij. Tom 2, chast' 2*, 327–413.
- Waag, Christine & Martin Phodunze. 2013. The pronominal elements in Baka, a Central Sudanic language of S. Sudan. Paper presented at Nilo-Saharan Linguistics Colloquium 11, May 22–24, Cologne.
- Weber, Heather. 1994. A syntactic and morphological sketch of Logbara. M.A. thesis, State University of New York at Buffalo.

SYNTACTIC SENSITIVITY AND PREFERRED CLAUSE STRUCTURE IN MAJANG

Andreas Joswig¹

1 Introduction

This paper will present two syntactic phenomena of the Majang language, which can be seen as a manifestation of a high sensitivity of the language to deviations from a preferred clause structure.

The Majang language is a Surmic language spoken by about 30,000 speakers in various scattered pockets in western Ethiopia. In their writings on Majang, both Bender (1983: 132) and Unseth (1989: 111) note the presence of the marker *-ŋ*, the use of which defies an easy description. Whereas Bender does not even try to provide an analysis, Unseth proposes that the marker should best be seen as an intransitivity marker. He concedes, however, that this analysis does not account for all the data.

Examining the range of data in which this marker appears (and just as importantly, where it does not appear), this paper reaches the following conclusions:

- Although Bender and Unseth call the marker a suffix, there are reasons to treat it as an enclitic.
- The marker has a purely syntactic function, without any reference to semantic or discourse-pragmatic information. It is therefore not optional, but obligatory if the right syntactic conditions apply.
- The marker is triggered by the absence of any noun phrase or adverb in the clause.

The Majang language, therefore, displays a high degree of sensitivity to the syntactic composition of a clause, forcing the speakers to react to the absence of particular clause components to make the clause grammatical. This also becomes apparent from another syntactic device, the absolutive-disjoint marking on the verb, which seems to be a clear instance of syntactic ergativity in an African language. It appears that the language has a preferred clause structure of Verb – NP_{abs}.

The language has a basic word order of VSO. Majang has a well developed morphological ergative system, leaving transitive objects and intransitive subjects unmarked, and using the ergative case for transitive subjects. The language is tonal, with two phonemic levels and downstep.

The Majang language features a marker *-ŋ*, whose presence or absence has long puzzled the linguists describing the language. Students of the language will encounter this morpheme early when they collect their first verb paradigms, as such verbs invariably are accompanied by this marker:

(1)	kò:w-á:† -ŋ	‘I dig’	kò:w-î:-ŋ†	‘we dig’
	kò:w-ún† ²	‘you _{sg} dig’	kòw-âri-ŋ	‘you _{pl} dig’
	kòw-ŋ	‘he digs’	kòw-âr-ŋ	‘they dig’

Bender (1983: 132) notes the marker’s “*optional*” presence and includes it in the description of the verbal paradigms. Unseth (1989: 111) also acknowledges the marker, noting that its function is to mark a clause as intransitive. But he also notes that “*there are still some sentences which do not have this intransitivity marker, which I cannot explain: there appears to be a degree of optionality in the use of -ŋ*”. In this assessment of optionality he agrees with Bender.

¹ I am deeply indebted to Ato Joseph Kalakun, who patiently provided the language data used in this study, and to Colleen and Michael Ahland, who gave valuable input into early versions of this paper. I am also grateful to those who offered comments after the presentation at the Nilo-Saharan Linguistics Colloquium in Cologne. Particularly I’d like to acknowledge Gerrit Dimmendaal who pointed me to the conjoint/disjoint phenomena of Bantu languages.

² For phonological reasons, the marker *-ŋ* is never observed following a 2s suffix.

In spite of this, Unseth provides some good data on the use of the marker, which may help towards a more definite understanding. He observes the following:

- *-ŋ* does not appear if the clause has any kind of complement.
- *-ŋ* does not appear in negative clauses (Unseth suggests that the negative construction in Majang has the lexical verb as a syntactic complement of the auxiliary negative verb).
- *-ŋ* does not appear if the clause has a direct object.
- *-ŋ* does not appear if the clause has an instrument NP.
- *-ŋ* does not appear if the clause has a locative NP.

According to these facts, the analysis of *-ŋ* as an intransitivity marker can only be maintained if the term *intransitive* is defined in wider terms than just the absence of a direct object in a clause. It appears that Unseth had such a more generous definition in mind.

2 The limits and the morphological status of *-ŋ*

Both Bender and Unseth describe the marker *-ŋ* as a suffix. Bender only seems to encounter it in the environs of a verb (including the TAM markers), so he may truly have assumed that it is a verbal suffix, whatever its function. Unseth (1989: 111), however, notes that *-ŋ* “*is suffixed to the verb if there is no subject NP, but suffixed to the subject NP if it is specified*”. Throughout the paper he gives a number of examples where a nominal suffix *-ŋ* is glossed as IV for ‘*intransitive*’. If this analysis were correct, it would then appear reasonable to call this morpheme not a suffix, but an enclitic. We will therefore use it with an enclitic connector =*ŋ* in all following interlinearisations. It seems necessary, however, to investigate the cases where =*ŋ* appears on a noun. In spite of Unseth’s claim that =*ŋ* is placed on the subject NP in each intransitive clause, many of his examples with intransitive clauses featuring a specified subject NP do not use this morpheme, as in the following clauses taken from his paper (Unseth 1989: 102, 112):

(2) (transcription and glosses as in the original)

- | | | | | |
|----|-------------|---------|--------|--------------------|
| a. | mɛl+ki+ko | maŋa | eet+ak | ‘My sibling came.’ |
| | come+DR+PST | sibling | me+GEN | |
| b. | den+ɛ+ko | bokaye | | ‘Bokaye saw.’ |
| | see+3S+PST | Bokaye | | |

This is in line with the data analysed for this study: No intransitive clause with a specified subject NP displays the morpheme =*ŋ*, neither on this NP, nor on the verb.

- | | | | |
|--------|-----------|-----------|-------------------|
| (3) a. | mèlki | jò:sép | ‘Joseph arrived.’ |
| | arrive.3S | Joseph | |
| b. | *mèlki | jò:sép=íŋ | ‘Joseph arrived.’ |
| | arrive.3S | Joseph=ŋ | |

When eliciting the clause ‘*Joseph arrived*’, the informant returned example (3a). When pressed whether (3b) would also be possible, he stated that it sounded very odd. There may be a context where this is possible, but normally this would not be used. He could not specify what such a context may be. When producing clause (3b), the informant insisted on the tonal differences on the verb and noun, which suggests that the syntactic relations in (3b) are not along the lines of verb-absolutive (see section below).

A possible explanation may be provided by Unseth (1989: 116f) himself, who encounters a homophonous morpheme *-ŋ* as a marker for conditional clauses. When describing complement clauses, once more he interprets the same nominal morpheme as the intransitive marker (Unseth 1989: 118f). Without being currently able to provide a clear analysis for the nominal morpheme *-ŋ* myself, I submit that this is a marker employed for complex sentences and clause conjoining.

Therefore we will discard all data where *-ŋ* shows up on nouns, assuming that this is another morpheme used for different purposes. But this leaves still another part of speech on which *=ŋ* may appear: the tense particles of Majang. Although these are part of the verb phrase, there are phonological reasons which indicate clearly that these need to be seen as separate words. As seen in (2), Unseth seems to sometimes interpret them as verbal suffixes; in other examples he rather interprets them as particles (Unseth 1989: 110). Bender (1983: 132ff) proposes an analysis as particles, but also is not consistent in his treatment (see Bender 1983: 134). Because of the Majang downstep rules³, an analysis as particles is preferable. Now, if such a tense particle is present in a clause, the morpheme *=ŋ* will always attach itself to this:

- (4) a. pá:kk⁺-i: kô:=ŋ. 'We will soon be hot.'
 be.hot-1P NFUT=**ŋ**
- b. á:dór kó:=ŋ. 'It recently became ripe.'
 become.ripe.3S RECPST=**ŋ**

Therefore, because *=ŋ* is not entirely found on verbs, it is advisable to analyse *=ŋ* as an enclitic which attaches itself to the end of a verb phrase. The tense particle, if present, is the last element of a verb phrase to which *=ŋ* may attach itself.

3 The syntactical status and function of *=ŋ*

The fact that *=ŋ* is an enclitic added to the verb phrase, and not to the verb as such, suggests that it serves not as a marker for a particular verbal category⁴, but in a function relating to the clause or predication as a whole. Whatever this function is, it must be found among the grammatical or pragmatic parameters which modify a clause. Bender's and Unseth's notion of optionality would suggest that this needs to be rather a pragmatic parameter – such pragmatic functions would at least require considerable context beyond what is found in the predication itself to be predictable with any amount of accuracy. A syntactically governed marker, instead, would need to appear in all circumstances where the syntactic conditions are met, and, just as importantly, it would be absent in only those environments where the conditions are not met.

The claim of this paper is that the marker *=ŋ*, in spite of Bender's and Unseth's assumption of optionality, has a clearly defined syntactic function.

3.1 *Verbless clauses*

Verbless clauses in Majang usually have two noun phrases, the first of which serves as the subject, and the second as the predicate.

- (5) cè:g ágálti:r. 'They are thieves.'
 they thieves

In such clauses the morpheme *=ŋ* is never used, in spite of their intransitivity.

3.2 *Intransitive constructions without complements and overt subjects*

This is the situation when verbs are elicited in isolation – they will always be produced with the clitic *=ŋ* in attendance.

³ If the tense particles were suffixes, for instance the tone on *kó:=ŋ* in example (4b) would be downstepped, but it is not.

⁴ Particularly, it cannot be analysed as a derivational marker, turning transitive verbs into intransitive verbs. This function is covered in the Majang language by the derivational suffix *-di*, which is placed between the verb root and the person suffixes. The position between verb root and inflectional affixes is the universally preferred place for derivational morphology, which is another reason why the interpretation of *=ŋ* as an intransitivity marker is unlikely.

- (6) a. båndúr=ŋ. ‘He groans.’
groan.3S=ŋ
- b. ŋájór-àrí=ŋ. ‘You (PL) smell.’
smell-2P=ŋ
- c. tòp-ùr=ŋ. ‘They continue.’
continue-3P=ŋ

In this situation, there is no optionality whatsoever – the use of =ŋ is required by the grammar of Majang.

3.3 Intransitive constructions without complements, but with overt subjects

This kind of construction assumes an intransitive verb in a situation where the subject is overtly expressed, either because it is the focus of the construction, or it needs to be (re-)introduced as the topic. In this case, the marker =ŋ is always absent. Most examples here are gleaned from elicited data.

- (7) a. kúc-ú jègúj. ‘The ox comes.’
come-3S ox
- b. dë:gár-ár ágáltí:r. ‘The thieves sleep.’
sleep-3P thieves
- c. cě:n tóç-è. ‘This is right.’
it correct-3S

Example (7c) illustrates that the overt subject does not have to be a full noun-phrase in order for =ŋ not to show up. As long as the subject is more than just referenced anaphorically by verbal agreement, the conditions for placing the marker =ŋ are not given. This is also the case when the verb in question is a derived intransitive verb:

- (8) káw-dí wár[†]. ‘The dog bites.’
bite-INTR dog

In this example, the transitive verb *kaw* ‘bite’ is used in an antipassive construction, involving the intransitiveriser *-dí*, and the subject is expressed in a full noun phrase. =ŋ cannot be used in such a context.

These examples make it clear that =ŋ cannot be analysed as an intransitive marker of Majang. There are countless intransitive constructions in the language which make no use of this marker.

3.4 Intransitive constructions with complements

As Unseth (1989: 111) correctly pointed out, the marker =ŋ never appears when there is any kind of complement or adverbial element beyond the subject in a clause.

- (9) a. nè kè: dũk-à cê ‘and she went to the forest’
CONJ go.3S forest-DAT SPEC
- b. nè ròg-èr Jét ‘and they laughed hard’
CONJ laugh-3P very

In these examples, the subject is only marked on the verb, so it can be concluded that the absence of a subject NP in itself does not trigger the marker =ŋ.

3.5 Transitive constructions

As would be expected from Unseth (1989: 111), most transitive constructions, that is those constructions involving an overt direct object, will not display the morpheme =*ŋ*.

- (10) a. nè b̂á⁺ ɪĵá:g òlà ká:rín-òŋk ‘and he made weapons’
 CONJ REM PST make.3S things fighting-GEN
- b. nè b̂òŋk-ù órkán cìnè b̂à kòbè-òŋk ‘and he took the bark of the
 CONJ take-3S bark which of Kobetree.GEN Kobetree’

It is probably this behaviour that prompted Unseth to assume a detransitivising function of =*ŋ*. But there are indeed transitive constructions which do feature the marker =*ŋ*:

- (11) p̂ò:c-á-ni=*ŋ*. ‘I thank you.’
 praise-1S-2S=*ŋ*

In this example, both subject and object of the transitive construction are expressed by anaphoric agreement only. Like in the intransitive clauses of this type in (6), the marker =*ŋ* needs to be present.

3.6 Conclusion on the function of =*ŋ*

Since the available data seems to rule out any analysis of =*ŋ* as a detransitiviser, another explanation for its use must be found. The only place where =*ŋ* consistently shows up is in clauses which do not have any noun phrase alongside the verb. Conversely, any clause with any kind of noun phrase (including pronouns or adverbs) never displays the morpheme =*ŋ*. Therefore, the function of =*ŋ* must be described along the lines provided by these facts.

It appears that the Majang language is characterised by a high degree of sensitivity towards the syntactic configuration of a clause. There seems to be a constraint in the language that does not allow for clauses without at least one fully expressed argument or adjunct. At the same time, Majang is an *anaphoric-agreement* or *pro-drop language* (see Payne 1997: 223), allowing for clauses where subjects and even objects are expressed by verbal agreement only. Therefore, the marker =*ŋ* may be viewed as an empty placeholder for the missing syntactic element, providing a dummy noun phrase to clauses which otherwise don't have one. In this sense we may revise our earlier statement that the enclitic attaches to the verb phrase of a clause; instead, it attaches to the clause as a whole.

Another perspective to view this phenomenon may be that =*ŋ* serves as a predictive syntactic device, warning the hearer about the things still to come (or better, not to come) in the clause. Since the verb is usually at the beginning of the clause (although there are some pragmatically conditioned exceptions to this), the grammar has developed ways to prepare the hearer for what else is to be expected in the rest of the clause. One of these ways is the marker =*ŋ*, which, added to the verb phrase, warns the hearer that there will be no information besides that which is already marked on the verb.

Other Surmic languages display phenomena which look quite similar to this marker. The closest may be in Me'en, a Southeast Surmic VSO language, about which Will (1989: 141) writes

“If the verb in the unmarked tense is the final element of a sentence unit, the suffix <Do> has to be added. [...] it is obligatory not only when the object is deleted after the verb, but also when it is preceding. In addition it is not only an object following a verb that prevents <Do> from appearing [...].”

It may be argued that a clause with a preceding object does not qualify as an anaphoric-agreement-only clause, but it may well be that this preceding object is in fact outside the clause. The phonetic shape (alveolar stop with vowel /o/), however, does not look close to the Majang velar nasal clitic. In line with Me'en is Suri (Tirmaga), on which Bryant (1999: 95f.) states that

“the Tirmaga verbs make use of a suffix on the end of a verb to indicate that no direct object will follow. I have called this a phrase final suffix. [...] In Tirmaga, the verb does not have to be the last element of a sentence, though it is commonly the case. It is possible for a locative, adverb, or adverbial phrase to follow a verb with a phrase final

suffix. The object is the only item restricted from occurring post-verbally when the phrase final suffix is used. When a verb has a phrase final suffix, the sentence may have an overt direct object, but the direct object may only occur in the pre-verbal position.”

This Tirmaga suffix is very similar in shape to the Me'en suffix, but this time it really seems to mark the verb as intransitive, as other noun phrases do not block it from appearing. Both the Me'en and the Tirmaga suffixes have in common with Majang that they appear at the very end of the verb. This is very different from Southwest Surmic Murle, where there is an intransitivity marker for a number of verbs. This is a stem extension (before the person suffixes), only used “*when there is no object*” (Arensen 1982).

4 Absolutive-disjoint marking

There is another way in which the Majang language prepares the hearer for the syntax of the clause elements to follow the verb. A verb could face ambiguities with regard to the noun-phrase following the verb: this could either be the subject of the clause, or a noun phrase with another syntactic status. Therefore it would help the hearer to decode a clause if clues are given in advance whether the NP following the verb is the subject or not. But this is not what the Majang language does. Most 3*S*-verbs come with two different tonal patterns: they either have their lexical tone melody (usually *H*, *L*, or *LH*), or they come with a tone replacement of *L* all across the verb, with a *H* on the last syllable.

- | | | | | | |
|------|----|----------------------------------|-------------------------------|-------------------------------|---------------------------|
| (12) | a. | mèlk-i
arrive-3 <i>S</i> | tô:n
boy.ABS | | ‘The boy arrived’ |
| | b. | ɲù:lé
break-3 <i>S</i> | tôn
boy.ERG | béá ⁺
spear.ABS | ‘The boy broke the spear’ |
| | c. | ɲù:l-è
break-3 <i>S</i> | béá ⁺
spear.ABS | | ‘He broke the spear’ |
| | d. | mèlk-i
arrive-3 <i>S</i> | kâtámé-á
town-DAT | | ‘He arrived in town’ |
| | e. | mèlk-i=ɲ
arrive-3 <i>S</i> =ɲ | | | ‘He arrived’ |

The two verb-stems *ɲù:l-* ‘break’ and *mèlk-* ‘arrive’ both have a lexical low tone, which also becomes apparent when these verbs are used in 1st or 2nd person. This lexical low tone shows in (12a), where the verb is followed by the subject noun phrase, and also in (12c). In examples (12b), (12d) and (12e), however, the alternative tone pattern is used. In (12d) the following noun phrase is marked by the dative case; in (12e) there is no NP at all (which is also indicated by the anaphoric-agreement-only-marker =*ɲ*).

The difficulty is with examples (12b) and (12c): The NP following the verb in (12b) is the subject, and it agrees with the subject-marking on the verb, but still the alternative tone pattern is employed. The noun phrase following the verb in (12c) is not the subject, but the object; still, the lexical tone pattern is used. So clearly the use of the alternative tone pattern is not governed by the presence or absence of a subject NP following the verb. But since Majang has a morphological ergative-absolutive case marking system (Joswig, forthcoming), the subject noun phrase in (12b) is marked by the ergative case, in this instance expressed through a different vowel length. And it appears to be this different case marking on the NP following the verb that triggers the alternative tone pattern on the verb. This also explains why in (12c) the original tone pattern is used. The following NP is not the subject, but it is in the absolutive case. We therefore conclude that this alternative tone pattern is employed to prepare the hearer that the verb is not followed by an absolutive case NP, either because the absolutive NP is dropped, or because it follows later in the clause, after a noun-phrase in a different case (as in example (12b)).

In cases when the subject of the verb is not in the third person singular, this absolutive disjoint marking does not tonally affect the stem. Only the tone of the person-suffix changes. For example, the first person singular suffix *-a*⁺ always has a fixed low tone (as in (13a)). In an absolutive disjoint context, however, it comes with a fixed high tone and a short vowel (13b):

- (13) a. *lāk-à.* ‘I have’ (*when followed by an absolutive NP*)
 have-1S
- b. *lāk-á⁺* ‘I have (abs. disjoint)’
 have-1S

Similar phenomena have also been observed in Bantu languages, from which I have borrowed the terminology *conjunct/disjunct marking* (see Creissels 2012 and van der Wal 2011 for recent accounts). Majang’s absolutive disjoint marking could conceivably be described along the lines of the Bantu disjoint marking. It is interesting to note, however, that in Bantu languages the origin of the conjunct/disjunct distinction is usually traced back not to syntactic, but to pragmatic needs – in the case of Makhuwa the conjunct marking links the verb closer to an exhaustive-focused NP in the immediate after-verb position (van der Wal 2011: 1749). In Tswana (Creissels 2012: 18), although now of a superficially syntactic nature, the conjoined-disjunct distinction can diachronically be tracked to the presence or absence of non-topical NPs following the verb. Buell (2006) instead sees the conjunct-disjunct distinction in Zulu as a syntactic feature of the language, more in line with the observations in Majang. Also interesting is the notion that in Makhuwa the conjunct forms seem to be the morphologically marked ones, whereas the disjunct appears to be unmarked. In Tswana, both forms are marked tonally. In Majang instead, the disjunct form is clearly the morphologically marked one, and is opposed to no marking, which could be conceived as conjunct. So it would probably not be right to speak of a conjunct-disjunct distinction in Majang, but of disjunct marking only. This, again, seems to be more along the lines of Zulu as described by Buell (2006).

This absolutive-disjunct marking on the verb throws an interesting light on the nature of ergativity in Majang. Most languages with ergative systems are not expected to display any syntactic ergative structures (Dik 1989: 243–244; Givón 1984: 165–166; Anderson 1976: 11; Andrews 1985: 130; Van Valin & LaPolla 1997: 580) – although the morphological marking on the noun phrases groups intransitive subject and transitive object together, the syntactic relationships in the clause usually follow a nominative-absolutive pattern. This is also the case in Majang, where the subject agreement on the verb always points to the actor of a transitive sentence (Joswig, forthcoming). But in the case of the absolutive-disjunct marking, there is a clear case where the syntactic relations of subject and object are ignored, and the morphological absolutive case is the basis for a syntactic process. So, in a very restricted context, the Majang language not only displays morphological, but also syntactic ergativity.

Abbreviations

1, 2, 3 S/P	1 st , 2 nd , 3 rd person singular/plural	NP	noun phrase
ABS	absolutive	O	object
CONJ	conjunction	PL	plural
DAT	dative	PST	past tense
DR	derivational affix	RECPST	recent past tense
GEN	genitive	REMPST	remote past tense
ERG	ergative	S	subject
H	high tone	SG	singular
INTR	intransitive marker	SPEC	specific reference
L	low tone	V	verb
NFUT	near future tense		

References

- Anderson, Stephen R. 1976. On the notion of subject in ergative languages. In Charles N. Li (ed.), *Subject and Topic*, pp. 1–23. New York: Academic Press.
- Andrews, Avery. 1985. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Vol. I. Clause Structure*, pp. 62–154. Cambridge: Cambridge University Press.
- Arensen, Jonathan. 1982. Murle grammar. *Occasional Papers in the Studies of Sudanese Languages*, 2.
- Bender, Lionel M. 1983. Majang phonology and morphology. In Lionel M. Bender (ed.), *Nilo-Saharan Language Studies*, pp. 114–147. East Lansing, MI: Michigan State University, African Studies Center.
- Bryant, Michael Grayson. 1999. *Aspects of Tirmaga Grammar*. Arlington: University of Texas at Arlington.
- Buell, Leston. 2006. The Zulu conjoint/disjoint verb alternation: focus or constituency? *ZAS Papers in Linguistics* 43: 9–30.
- Creissels, Denis. 2012. *Conjoint and Disjoint Verb Forms in Tswana and other Bantu Languages*. Paris: Fédération Typologie et Universaux Linguistiques.
- Dik, Simon C. 1989. *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Dordrecht: Foris Publications.
- Givón, Talmy. 1984. *Syntax: Volume I. A Functional-Typological Introduction*. Amsterdam/Philadelphia: Benjamins.
- Joswig, Andreas. (forthcoming). Ergativity and antipassive in Majang: another look. In *Proceedings of the 7th World Congress of African Linguistics, Buea, August 2012*. Cologne: Köppe.
- Payne, Thomas E. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press.
- Unseth, Peter. 1989. Sketch of Majang syntax. In Lionel M. Bender (ed.), *Topics in Nilo-Saharan Linguistics*, pp. 97–127. Hamburg: Buske.
- van der Wal, Jenneke. 2011. Focus excluding alternatives: conjoint/disjoint marking in Makhuwa. *Lingua* 121.11: 1734–1750.
- Van Valin, Robert D. & Randy J. LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.
- Will, Hans-Georg. 1989. Sketch of Me'en grammar. In Lionel M. Bender (ed.), *Topics in Nilo-Saharan Linguistics*, pp. 129–150. Hamburg: Buske.

COMPLEX VERBAL PREDICATES IN UDUK

Don Killian

1 Introduction to complex predicate constructions

Complex predicates are multi-headed predicates in which two or more grammatical elements (such as morphemes or words) act as a single semantic predicate (Alsina et al. 1997). That is, the semantic meaning of a head (in this case, the focus is on verb heads) is not contained solely in the verb itself, but rather distributed across the predicate. Different researchers take different approaches as to which types of constructions fall under the description complex predicate, but common correlations have to do with empty semantics of the head verb, as well as retention of valency of the co-occurring noun (when it is a noun).

Germanic phrasal verbs are a well known of complex predicates, as seen below:

- (1) Sie sag-t-en das Konzert ab
3PL say-PST-3PL the concert off
'They cancelled the concert.' (lit: they said the concert off)

There are a variety of possible constructions of this type, in fact, ranging from the templatic verbal morphology of Athabaskan languages, in which the basic meaning of a verb is distributed across different slots in the template, to coverb constructions in northern Australia and Papua New Guinea, in which there is a small amount of inflecting verbs which combine with an open class of 'coverbs', forming a syntactically integrated construction which only together give the meaning of the phrase. Similar constructions called light verbs are found across Ethiopia and Sudan (as well as other parts of the world).

Note the following examples in Amharic, from Amberber (1995), in which the verb 'say' is shown to combine with a variety of coverbs, to form a complex predicate.

- (2) Ləmma yi-hedal alə
Lemma 3MS-go.IMP.3MS say.PFV.3MS
'Lemma said "He will go." '
- (3) t'ərmusu sibbir alə
glass.DEF break say.PFV.3MS
'The bottle broke.'
- (4) Ləmma rot' alə
Lemma run say.PFV.3MS
'Lemma ran.'
- (5) Ləmma zim alə
Lemma quiet say.PFV.3MS
'Lemma kept/became quiet.'

A new type of complex predicate construction is introduced here, namely that of the *Partargument* construction, a term coined to refer to a construction in which body parts are incorporated into the verbal predicate to change the basic meaning of the verb, or to classify the arguments of the verb. These resemble (and are related to) both coverb/light verb constructions as well as noun-verb idioms, but are distinct from both.

In the following example in Uduk, for instance, the verb *'ci'th*, 'to cut/cross', combines with the partargument *'twā?*, 'mouth', in a discontinuous manner, creating a new meaning for the verbal predicate, 'to hinder'.

- (6) à shō'k 'ci'th únī 'twā?
 CL2 rain cross:PFV 3PL mouth
 'Rain hindered them.' (lit: the rain crossed them mouth)

This construction can also be found in Berta, a Nilo-Saharan (although not closely related) language spoken in the same area as Uduk. Note the following examples from Neudorf (2013):

- (7) Gali gag-í mia
 dog chase-PRES goat
 'The dog is chasing a goat.'
- (8) fāmiili gag-í gali alú
 people chase-PRES dog head
 'The people are chasing the dogs away.'

Many languages in West Africa appear to have a similar construction, variously called a (grammaticalized) Double Object construction in Likpe (Ameka 2009) or Inherent Complement Verb (Essegbey 1999). The following Ewe examples come from Essegbey (1999).

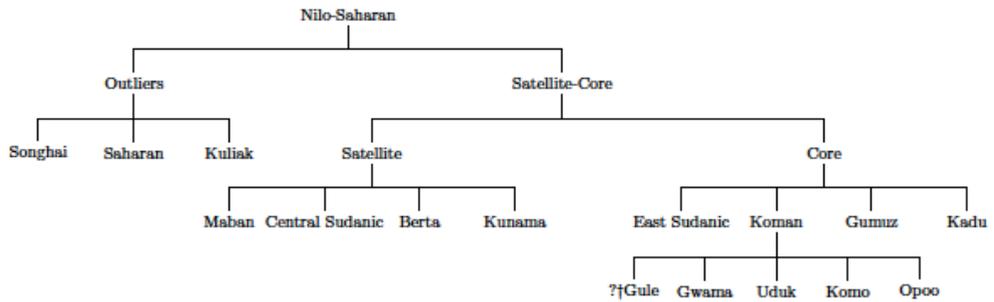
- (9) Kofi fú du
 Kofi ICV IC
 'Kofi ran.'
- (10) Kofi fú tsi
 Kofi ICV water
 'Kofi swam.'
- (11) Kofi fú kó Ami
 Kofi ICV fist Ami
 'Kofi knocked Ami.'

In this article, we give an overview of necessary basic syntactic information in Uduk, followed by a discussion of how partargument constructions work. We end with a short discussion of the semantics, and suggest that the structure is similar to a range of typologically related phenomena.

2 Introduction to Uduk

Uduk [udu], also known as Tw'āmpā (Orthographic: 'Twampa), is a little-studied Koman language of the Sudan-Ethiopian borderlands. There are approximately 20-25,000 speakers of Uduk, according to James (2007); however, due to the political turmoil in Sudan beginning from the 1950's, the Uduk have been forced to flee as refugees to a number of different countries around the world, and it is difficult to estimate the actual number of speakers.

In terms of classification, Uduk forms a node on the Nilo-Saharan language tree along with Komo, Opuuo, Gwama in the Koman family, and somewhat more distantly, Gumuz (Bender 1997, Ahland 2012). The extinct language Gule was also placed into Koman by Greenberg (1963) with relatively little data available, and may not belong to the Koman languages.

Figure 1: Koman and Nilo-Saharan languages (Bender 1997)

Linguistic research began in the 1940's, when the Forsberg missionaries working for the organization SIM opened a school in Chali, and began bible translation and some linguistic study. While living in Chali, the Forsbergs kept in contact with Roland Stevenson, a linguist living at the time in Khartoum, who then helped the Forsbergs to understand the language. Around 1944, Stevenson completed a grammar sketch of Uduk based on information provided by the Forsbergs via correspondence, although it remained unpublished, and numerous areas such as tone were not addressed. Although the grammar sketch was not published, it was used by Tucker and Bryan (1956, 1966) in their covering of the Koman language family.

The Forsbergs meanwhile were working on the bible and orthography, with portions of the bible translated between 1947 and 1963. 1963 saw the first edition of the entire new testament of the bible (Beam et al. 1963). Also made during this time were some literacy primers, teaching Uduk children how to read; these remain as of yet unpublished. Later periods also saw some anthropological research by Wendy James (e.g. James 1979, 1988), as well as an Uduk-English dictionary (Beam and Cridland 1970).

The alphabet used in the orthography is as follows, with the IPA representation below:

a	b	ʔb	c	ḥ	ʔc	d	ʔd	dh	e	g	h	i	j	k
a, ʌ	b	ḃ	c	ḥ	cʔ	d	dʔ	ḏ	e, ε	g	h	i, ɪ	ʝ	k
ḵ	ʔk	l	m	n	ŋ	ny	o	ḡ	p	ʔp	r	s	sh	t
ḵʰ	kʔ	l	m	n	ŋ	ɲ	o, ɔ	p	pʰ	pʔ	r	s	ʃ	t
ṭ	ʔt	th	ṭh	ʔth	u	w	y	/						
tʰ	tʔ	ṭ	ṭʰ	ṭʔ	u, ʊ	w	j	ʔ						

As the orthography does not mark tones, it is not entirely suitable for a linguistic analysis. However, the characters actually used in the orthography are generally more visually distinctive than the standard characters of the IPA, particularly when it comes to the dental versus alveolar distinction. This work will henceforth use a modified version of the orthography; all consonants and vowels will be written as in the orthography, with the exception of the glottal stop /, written here as ʔ. Furthermore, tone is marked in all instances.

Uduk has a 5-vowel system underlyingly, but there are phonetic variations of the vowels on the surface due to interactions with tone, consonants, and syllable structure; vowels are written here with the basic Roman letters a, e, i, o, and u, despite their phonetic values. Labialized consonants will be written as a sequence of Cw rather than Cʷ.

An overview of the phonology can be found in Appendix 1.

3 Alignment and clause structure

There are three main factors to take into consideration in order to understand basic alignment and clause structure. These considerations are necessary in order to understand complex verbal predicate constructions, and are as follows: the *Gender* of the arguments involved in the construction, which is lexically determined; the *Case* of the arguments, which has to do with the number of arguments (and

partarguments) as well as their syntactic roles in the clause; and the *Voice* of the construction, a crucial factor dependent on information structure, which can change the whole construction.

Uduk generally is a V2, or Verb Second language, with a constituent order in bivalent independent clauses as follows:

Top V A IO O Partarg Particle Adv

A prototypical example could be seen as follows¹⁰⁵:

- (12) wàhí? 'cí'th cwá mò
 man cut:PFV tree COMPL
 'The man cut down a tree.'

At first glance, this could appear to be an instance of an AVO language; however, clauses of the following type are also fairly common in Uduk.

- (13) tāshá wò'c mā 'ká
 snake bite:IPFV ERG.CL2 dog
 'The dog bit the snake.'

Thus, as can be seen in the previous example, Uduk exhibits case marking, but case marking is partially dependent on the noun's position in the clause, and not just its role in the syntax. The case form specific to postverbal A is called *Ergative*, whereas preverbal (and most other) nouns are marked with the *Absolutive*. In monovalent clauses, S occurs in the preverbal topic position and does not generally occur postverbally (in the rare instances in which it does, it does not take case marking); adverbials do not affect this placement.

For any bivalent construction, there are two voices, depending on whether A is foregrounded (Agent Voice, or *A-Voice*), or whether A is backgrounded (Object Voice, or *O-Voice*). Note that although O-Voice is used here as a convenient label, it is not necessary that O itself occupies the preverbal topic position; any other argument may also fill this slot.

A-Voice constructions are those in which A is the topic of the construction, in the preverbal position. They correlate more closely with the typological notion of a prototypical bivalent predicate. The A argument shares more semantic features with Agent, such as higher degrees of animacy, volitionality, control, and initiation. The O argument on the other hand shares more semantic features with the Patient, lacking volitionality, control, and initiation, and high in affectedness.

- (14) á'dī dhàl-kí wà'c í mè'd
 3SG let:PFV-ITV fish CL1.LOC hand
 'He dropped the fish from his hand.'

O-Voice constructions are those in which A is not the topic, and the A argument occurs in a postverbal position. O generally occurs in the preverbal position. If there is another argument, however, there are two places in which the additional argument may occur. First, it may occur after A, if O is the topic; second, it may also preverbally as the topic, in which case O is placed after A. In either case, A is always directly after the verb, and no other arguments are allowed between the verb and A.

- (15) wà'c kál ā yì'dé?
 fish carry:PFV ERG.CL1 water
 'The water carried away the fish.'

¹⁰⁵ All examples which are not otherwise marked come from fieldwork trips to Ethiopia between 2011 and 2013. UD is used to refer to examples taken from the Uduk dictionary by Beam & Cridland (1970).

The most prototypical example of an O-Voice construction is that of the experiencer construction, in which O, the experiencer, is high in animacy and affectedness, but low in volitionality and control. A on the other hand could be viewed as the stimulus, low in affectedness, but also (potentially) less animate and volitional than a prototypical Agent. Many ‘dative subject’ clauses would also fall into this category.

- (16) ʾdí būth ā ʾthám
bird hold:IPFV ERG.CL1 wind
‘The bird is cold.’ (lit: the wind holds the bird)

In other instances, however, the A argument may be both animate and volitional, or the arguments are at least equal in animacy, but the speaker chooses to lay greater importance to O. These tend to be less common, but are used for instance when answering a question, or when the topic is carried over from the previous topic of discussion. Thus although O-Voice constructions are commonly used with experiencer constructions, constituent order is generally flexible in Uduk and other types of constructions may also change to O-Voice.

- (17) pólóny shūn ā wàthí?
pumpkin pick:PFV ERG.CL1 man
‘The man picked the pumpkin.’

One important factor to keep in mind regarding transitivity in Uduk is that nearly all verb roots have the possibility of taking either monovalent or bivalent constructions, and there are variations in the verbal agreement patterns depending on the number of arguments. The terms *monovalent* and *bivalent* are used in this article to refer to the overt number of surface arguments in a given construction.

To create a monovalent construction with a verb high in semantic transitivity, the monovalent paradigm is often used to leave the object unspecified when it is clear from the context what is meant. The object is simply left unstated in much the same way that many ergative languages use an antipassive construction to suppress the object. As all paradigms in Uduk are overtly marked, it is not necessary to have a devoted marker for an antipassive; one can simply change the paradigm to create the same meaning.

- (18) áʾdí dhīth gùʾb
3SG sweep:IPFV house
‘S/he is sweeping the house.’

- (19) áʾdí dhīth-íʾd
3SG sweep:IPFV-3SG
‘S/he is (doing) sweeping.’

For verbs which are very low in semantic transitivity, they are able to take bivalent paradigms typically due to changes in meaning associated with paradigm alternations. The verb root *wù*, ‘to die’, would more than likely be considered more of a monovalent verb from a typological perspective. In Uduk, this root can occur with both monovalent and bivalent constructions. When the verb root combines with the word *jíłí?*, ‘sweat’, for instance, a complex predicate is created meaning ‘to perspire heavily’.

- (20) áʾdí wú-ʾd mò
3SG die:PFV-3SG COMPL
‘S/he died.’

- (21) á'dī wú jítí?
 3SG die:PFV sweat
 'S/he was perspiring heavily.'

Despite the somewhat idiomatic nature of the example, the construction itself is pervasive in the language, and nearly all verbs have at least some phrase they can use in order to become bivalent. In a number of instances, one of the arguments is obligatory, such as with the verbs *dream* (obligatory A argument) and *sleep* (obligatory O argument).

- (22) áhā shū'b ā jàn kā ēn'thíny
 1SG dream:IPFV ERG.CL1 dream PREP.CL2 last night
 'I dreamt a dream last night.' (lit: the dream dreamt me last night) (UD)

- (23) á'dī ísh ē
 3SG sleep:IPFV eye
 'S/he is sleeping.'

The word *ē* is an obligatory object of *ísh* when referring to sleeping, and the monovalent paradigm is not possible.

- (24) *á'dī ísh-í'd
 *3SG sleep:IPFV-3SG
 *'S/he is sleeping.' (intended meaning, monovalent paradigm)

Other instances of the verb *ísh* are possible without *ē*; however, the meaning changes.

- (25) wáthí? ísh-í'd kí sāpāth
 man sleep:IPFV-3SG CMPL stillness
 'The man is dead (or lying like the dead).' (lit: the man is lying like stillness) (UD)

- (26) áhā ísh-á á pó cwá
 1SG sleep:IPFV-1SG CL2.LOC back tree
 'I am leaning against the tree.'

3.1 Gender and case

There are two genders in Uduk which are not semantically based categories, provisionally labeled as *Class I* and *Class II*. Gender and case-marking are morphologically conflated in Uduk, and there is some segmental overlap between some of the different categories. Case marking is also dependent on the voice, as seen in table 1 below.

Table 1: Gender and case marking

	<i>A-Voice</i>				<i>O-Voice</i>			
	A		O		O		A	
Class I	∅	ABS	∅	ABS	∅	ABS	(-N) ā, à	ERG
Class II	à	ABS	ā, à ¹⁰⁶	ACC	à	ABS	(-N) mā, mà	ERG

There are two core cases in Uduk, Ergative and Absolutive, and a limited third with Accusative:

- Absolutive, which marks any preverbal argument, as well as Postverbal S and class I O
- Ergative, which marks postverbal A in bivalent constructions
- Accusative, which occurs only with class II O arguments which directly follow the verb

¹⁰⁶ Note that despite being nominal morphology, the Class II marker phonologically cliticizes onto the verb.

In most instances, the primary morphological difference between the two classes is the presence or absence of the proclitic *à*. Absolutive Class I nouns are left unmarked, and Absolutive Class II nouns are marked with *à*.

The Accusative case is only overtly marked with class II postverbal O arguments which directly follow the verb. Class I nouns in O position are marked in the Absolutive, and additionally trigger agreement changes in the verbal paradigm, in that 2nd and 3rd person agreement markers are not expressed. Class II nouns when in direct postverbal O position raise the tone of Absolutive marker from *à* to *ā*, which then moves to a phonologically bound cliticization onto the verb. There is likely a grammaticalization process occurring in which nominal morphology is changing into verbal. This marker is also affected by a tonal rule, *ā* lowering to *à* when directly following a low tone.

A set of the different possible conjugation forms for the verb '*ci'th*' 'cut', with 3SG, are given here for the sake of comparison:

Monovalent

- (27) wāthi? 'cith-i'd
 man cut:PFV-3SG
 'The man was cutting (something), doing cutting.'

Bivalent, class I object

- (28) wāthi? 'ci'th yi'dé?
 man cut:PFV water
 'The man was crossing the water.'

Bivalent, class II object

- (29) wāthi? 'cith-i'd ā yi'd
 man cut:PFV-3SG CL2.ACC skin
 'The man was cutting the skin.'

If additional arguments are added to a bivalent construction, such as with the applicative *ki*, the more animate O generally occurs first, and the verb conjugates according to whichever argument comes directly after it. Pronouns are treated as Class I nouns in terms of how they affect verbal paradigms. The second argument after the verb is always in Absolutive case.

- (30) á'dī 'kwār-á'd ā wārkā?
 3SG write:IPFV-3SG CL2.ACC book
 'S/he is writing a book.'
- (31) á'dī 'kwār-kí é à wārkā?
 3SG write:IPFV-APPL 2SG CL2 book
 'S/he is writing a book for you.'

The Ergative case is marked by a particle rather than a clitic, *mā* for Class II A arguments (replacing the Absolutive *à*), and *ā* for Class I A arguments. Both of these markers are affected by the same tonal rule as with Class II O, in that *mā* and *ā* lower to *mà* and *à* when directly following a low tone in the verb. Verbs ending in a vowel also add an additional epenthetic nasal -N before the A argument. Postverbal Class I A arguments superficially resemble Postverbal Class II O arguments, but cannot be confused due to other changes in the verbal paradigm.

(32) cwá 'cí'th ā wàthí? mò
 tree cut:PFV ERG.CL1 man COMPL
 'The man cut down the tree.'

(33) cwá 'cí'th mā Édwàd mò
 tree cut:PFV ERG.CL2 Edward COMPL
 'Edward cut down the tree.'

One final factor should be mentioned: all verbs discussed up until this point have been active. It is also possible to get stative verbs in Uduk. Stative verbs only allow for a single argument under normal circumstances (and voice thus does not usually play a role). Partarguments invariably trigger bivalent constructions, however. Stative verbs with partarguments are thus a special case, however, and present a challenge in terms of syntactic analysis, as they fall in between ordinary stative verb constructions and other types of bivalent constructions.

4 Introduction to complex predicate constructions in Uduk

As mentioned previously, there are discontinuous constructions in Uduk in which the semantic content of a verbal predicate is carried not only in the verb itself, but rather distributed across multiple words, typically the verb and a body part. Our prototypical example of a bivalent verbal clause was a man cutting down a tree, with *wàthí?* 'man' as A, *cwá* 'tree' as O, and the basic verb '*cí'th*, 'to cut, cross' in the Class I O paradigm.

(34) wàthí? 'cí'th cwá mò
 man cut:PFV tree COMPL
 'The man cut down a tree.'

Now compare this to a variety of partargument expressions which all use the same verb root:

(35) à 'tāsh 'cí'th mè'd
 CL2 soup cross:PFV hand
 'The soup was delicious.' (lit: the soup crossed hand) (UD)

(36) wàthí? 'cí'th 'pén
 man cross:PFV behind
 'The man is late.' (lit: the man crossed behind)

(37) wàthí? 'cith-í'd ā 'cí 'pén
 man cross:PFV-3SG CL2.ACC child behind
 'The man was following the child closely behind.' (lit: the man crossed the child behind)

(38) wàthí? 'cith-í'd ā ŋwá 'bā?
 man cross:PFV-3SG CL2.ACC hen neck
 'The man cut the head off the chicken.' (lit: the man crossed the chicken neck)

(39) wàthí? 'cith-í'd ā yí'd bwà
 man cross:PFV-3SG CL2.ACC skin stomach
 'The man cut the animal skin in half.' (lit: the man crossed the skin stomach)

(40) wàthí? 'cith-í'd ā yí'd ĩs
 man cross:PFV-3SG CL2.ACC pelt body
 'The man trimmed the animal skin.' (lit: the man crossed the skin body)

- (41) à shō'k 'cí'th únī 'twā?
 CL2 rain cross:PFV 3PL mouth
 'The rain hindered them.' (lit: the rain crossed them mouth) (UD)

Syntactically, these are all A-Voice constructions which can be divided into two main types: whether the partargument is occupying the O slot (examples 35, 36), or whether there is a separate overt object and the partargument occurs after the object (examples 37–41).

The first type is parallel to normal verb-noun idioms across the world. Syntactically, a sentence like 'the man hit the road' does have an overt agent and patient in much the same way that the Uduk constructions do, but the arguments themselves are meaningless; the meaning is emergent, obtained only from the whole construction. Compare this to the Ngit examples below (Kutsch-Lojenga, personal communication), in which body parts (as arguments) are also used idiomatically:

- (42) ma mí tsù-dú-nga nóngù
 I PRES.CT mouth-my.LOC-place ask
 'I am thinking.' (lit: I am asking the inside of my mouth)

- (43) dù-du nínza iwà ndì
 head-my destroy:PFV already itself (REFL)
 'I am troubled.' (lit: my head has destroyed itself)

In addition to filling the role of arguments, however, partarguments also can take part in something more closely resembling light verb or coverb expressions, and occur after all other objects. The partargument no longer has any real syntactic role, in a sense, and is more similar to an incorporated discontinuous element of the verb. The semantic content of the construction is distributed across both the verb root and its partargument, but it is not quite as idiomatic as that of the first type. The construction resembles light verb constructions found in a number of Nilo-Saharan (and Afroasiatic) languages, as well as coverb constructions in Australian languages. However, the verb usually does have a basic meaning on its own, so the idea of a semantically empty head verb is not quite the same. Instead, the meaning simply changes with the addition of a partargument. Meaning shifts vary in their transparency; when functioning as predicate classifiers, the meaning is generally clear, but other times it is more idiomatic, and cannot be guessed on the basis of the individual components.

Although there are no monovalent partargument constructions, partarguments do occur in O-Voice in addition to A-Voice. O-Voice partarguments are those in which the default construction is OVA, and the partargument typically fills the postverbal A slot. They tend to be more emotional or experiential in nature, and parallel numerous other idiomatic experiencer constructions. These often only have one actual semantic argument, a semantic S argument which is syntactically O, but they may have up to three partarguments. The stimulus partargument thus occupies something of a dummy syntactic role.

- (44) wàthí? 'kósh mā sū
 man hit:IPFV ERG.CL2 beer
 'The man is drunk.' (lit: the beer hit the man)

- (45) wàthí? 'kósh ā tē?
 man hit:IPFV ERG.CL1 hunger
 'The man is hungry.' (lit: hunger hits the man)

- (46) wàthí? 'kósh ā dù
 man hit:IPFV ERG.CL1 liver
 'The man has a problem with his liver.' (lit: the liver hits the man)

Each of the previous examples could be viewed as an idiomatic experiencer construction, as the meaning is fairly clear and the A arguments in examples 44–45 are not with nouns which ordinarily

occur with partargument expressions. Example 46 does use *dù*, which occurs in a number of partargument expressions; however, in this particular instance it is used in a more literal sense about one's actual liver. In the following example, 'to pass' plus 'mouth' gives a new unpredictable composite meaning, that of passing someone (e.g. in a race or on a road). The construction cannot take an overt semantic object.

- (47) wàhí? pé-n ā 'twā?
 man pass:IPFV-NAS ERG.CL1 mouth
 'The man is passing someone.' (e.g. in a race) (lit: The mouth passes the man)

However, although that particular construction does not occur with a semantic object, it is rather common to find others which do. Furthermore, partargument constructions with more than one partargument are also not rare.

- (48) wàhí? wù-n à īs ē
 man die:IPFV-NAS ERG.CL1 body eye
 'The man is very tired.' (lit: The body dies the man eye)

Further examples of some phrasal verbs are given in tables 2–5, both with and without partarguments. The distinction between idiomatic phrases in which a partargument is used and those in which another nominal argument is used is largely arbitrary, mostly having to do with the frequency in which the noun can take part in the construction. The word *sī*, 'beer', for instance, does not take place in a large number of expressions, and would be thus categorized slightly differently than *bwà*, 'stomach', which takes place in hundreds of combinations.

Table 2: Phrasal verb examples: 'kósh - to hit, kill

A-Voice Constructions		
'kósh _____ īs	hit _____ body	mix solids
'kósh _____ ē	hit _____ eye	mix liquids
'kósh _____ shō'k	hit _____ foot	kick at, kick the air
'kósh _____ mè'd	hit _____ hand	punch at, punch the air
'kósh shō'k ē	hit foot eye	go and come back quickly
'kósh 'kúp	hit head	wash one's hair
A-Voice Idioms		
'kósh cwá īs	hit wood body	row a boat
'kósh _____ jwà	hit _____ complete	kill
O-Voice Constructions		
_____ 'kósh ā 'kúp	_____ hit (by) head	be finished
_____ 'kósh ā shē?	_____ hit (by) tooth	have a toothache, cavity
_____ 'kósh ā dù	_____ hit (by) liver	have a problem in your liver
_____ 'kósh mā bòr	_____ hit (by) chest	have a problem in your chest
O-Voice Idioms		
_____ 'kósh mā kó	_____ hit (by) milk	have malaria
_____ 'kósh mā sū	_____ hit (by) beer	be drunk

Table 3: Phrasal verb examples: wá - to pick, break off something

A-Voice Constructions		
wá 'twá?	pick mouth	break out (of shell, for baby birds)
wá ē	pick eye	be happy after receiving something
wá _____ ē	pick _____ eye	poke out someone's eye
wá _____ 'kúp	pick _____ head	throw something and hit someone in the head
wá _____ gwò	pick _____ word	share a secret; share gossip

Table 4: Phrasal verb examples: wò'th - to help

A-Voice Constructions		
wò'th _____ 'kúp	help _____ head	stop someone on road
wò'th _____ ĩs	help _____ body	stop someone on road
wò'th _____ gwò	help _____ word	intercede verbally
O-Voice Constructions		
wò'th à bwà 'kúp	help (by) stomach head	decide, ponder, change mind

Table 5: Phrasal verb examples: gàm - to find, meet

A-Voice Constructions		
gàm _____ 'kúp	find _____ head	meet, join
gàm _____ ĩs	find _____ body	join, marry
gàm _____ gwò	find _____ word	find out (information)
gàm _____ gwò ĩs	find _____ word body	believe
gàm _____ gwò mè'd	find _____ word hand	agree with
gàm ĩs	find body	hold out, manage

5 Syntax of complex predicate constructions

The syntax of complex predicate constructions is not quite the same as constructions with ordinary arguments, but this depends to some extent on the details. Partarguments do take case marking for instance, as well as triggering verbal paradigm alternations, and they also operate in dependent clauses in the same way that other verbal predicates do.

- (49) màkátō tá-'d ā kó jin 'kōsh-ù ē nyàkká ālbūn
 macchiato COP-3SG CL2.ACC milk REL hit:IPFV-IMPRS eye with:CL2 coffee
 'Macchiato is milk which is mixed with coffee.'

However, although partarguments can become argument-like in certain situations, they still do not bear the same qualities as more typical arguments, even when actually filling an argument slot. There are movement restrictions, for instance, in contrast to normal voice-changing operations which shift the constituent order.

The details of the restrictions depend somewhat on the basic voice of the construction, but they do also share certain things in common.

First, all partargument constructions trigger bivalent paradigms in the verbal predicate. The verb does not recognize the construction as having only a single argument, despite the partargument's lack of semantic value. However, they do not affect the count in terms of Applicatives or Impersonals, and up to 6 arguments/partarguments in a clause have been found.

- (50) á'dī gām-kā? úm gwò 'pén ĩs
 3SG find:IPFV-1SG.ERG 2pl word behind body
 'I've found out from you about it, the truth of the matter.' (UD)

Furthermore, as mentioned earlier, double-partargument constructions are rather common.

- (51) áhā ús-á gwò ē
 I extinguish:IPFV-1SG word eye
 'I'm hiding something, holding something back.'

A second feature in common with both voices involves adverbial insertion. Even though adverbials typically are placed towards the end of the clause, they are fairly flexible in terms of placement. The adverb *shwānénin* refers to a short time ago, and variations in its placement may be seen in the examples below.

- (52) shwānénín áhā 'kósh-á ā ālbūn ē nyàkká kó
recently I hit:PFV-1SG CL2.ACC coffee eye with:CL2 milk
'I mixed the coffee with the milk a little bit ago.'
- (53) áhā 'kósh-á ā ālbūn ē shwānénín nyàkká kó
I hit:PFV-1SG CL2.ACC coffee eye recently with:CL2 milk
'I mixed the coffee with the milk a little bit ago.'
- (54) áhā 'kósh-á ā ālbūn ē nyàkká kó shwānénín
I hit:PFV-1SG CL2.ACC coffee eye with:CL2 milk recently
'I mixed the coffee with the milk a little bit ago.'

However, insertion of adverbials between the verb and the partargument is marginally accepted at best.

- (55) *?áhā 'kósh-á ā ālbūn shwānénín ē nyàkká kó
*?I hit:PFV-1SG CL2.ACC coffee recently eye with:CL2 milk
*? 'I mixed the coffee with the milk a little bit ago.'

For O-Voice constructions, the situation is quite similar.

- (56) shwānénín ālbūn 'kósh ā á'dī ē nyàkká kó
recently coffee hit:PFV ERG.CL1 3SG eye with:CL2 milk
'S/he mixed the coffee with the milk a little bit ago.'
- (57) ālbūn 'kósh ā á'dī ē shwānénín nyàkká kó
coffee hit:PFV ERG.CL1 3SG eye recently with:CL2 milk
'S/he mixed the coffee with the milk a little bit ago.'
- (58) ālbūn 'kósh ā á'dī ē nyàkká kó shwānénín
coffee hit:PFV ERG.CL1 3SG eye with:CL2 milk recently
'S/he mixed the coffee with the milk a little bit ago.'
- (59) *?ālbūn 'kósh ā á'dī shwānénín ē nyàkká kó
*?coffee hit:PFV ERG.CL1 3SG recently eye with:CL2 milk
*? 'S/he mixed the coffee with the milk a little bit ago.'

A third feature in which both voices share in common is that neither type of partarguments can be easily modified by adjectives. That is, both the partarguments functioning as arguments in idiomatic phrases as well as the partarguments which more closely resemble coverb constructions are generally unmodifiable. Thus far only one single instance has been found in which a partargument was modified by an adjective.

- (60) á'dī dhèl ā ē kí dhèl
3SG glitter:IPFV ERG.CL1 eye Cmpl REDUP
'S/he's blind.'
- (61) á'dī dhèl ā ē-n 'dé kí dhèl
3SG glitter: IPFV ERG.CL1 eye-ASS one Cmpl REDUP
'S/he's blind in one eye.'

This appears to parallel the ICV constructions of West Africa, which can also be modified only in rare instances. Essegbey (1999) discusses a situation in which he as a native speaker assumed that ICs could not be modified, until another speaker suggested an example with a modified IC.

In terms of differences between the voices, the main issues revolve around constituent alternations and changes. Movement changes can be somewhat complicated, but a brief overview is nonetheless important.

As mentioned previously, the first place in a clause is reserved for the topic of the construction. In bivalent constructions, this typically is either the object (direct or indirect), or the agent.

- (62) áhā 'kósh-á ā ālbūn ē nyàkká kó
I hit:PFV-1SG CL2.ACC coffee eye with:CL2 milk
'I mixed the coffee with the milk.'

- (63) ālbūn 'kósh-kā? ē nyàkká kó
coffee hit:PFV-1SG.ERG eye with:CL2 milk
'I mixed the coffee with the milk.'

For partargument constructions which do not occupy an argument position, it is not possible to front the partargument. This applies to both A-Voice and O-Voice partarguments.

- (64) *ē 'kósh-kā? ālbūn nyàkká kó
*eye hit:PFV-1SG.ERG coffee with:CL2 milk
* 'I mixed the coffee with the milk.' (intended meaning)

In applicative constructions indirect objects may also be fronted; similar to the previous example, however, it is still not possible to front the partargument.

- (65) áhā 'kósh-kí-ná wàhí? ālbūn ē nyàkká kó
I hit:PFV-APPL-1SG man coffee eye with:CL2 milk
'I mixed the coffee with the milk for the man.'

- (66) ālbūn 'kósh-kí-kā? wàhí? ē nyàkká kó
coffee hit:PFV-APPL-1SG.ERG man eye with:CL2 milk
'I mixed the coffee with the milk for the man.'

- (67) wàhí? 'kósh-kí-kā? ālbūn ē nyàkká kó
man hit:PFV-APPL-1SG.ERG coffee eye with:CL2 milk
'I mixed the coffee with the milk for the man.'

- (68) *ē 'kósh-kí-kā? wàhí? ālbūn nyàkká kó
*eye hit:PFV-1SG.ERG man coffee with:CL2 milk
* 'I mixed the coffee with the milk for the man.' (intended meaning)

However, an unusual property of these constructions is that it becomes possible to front arguments in some instances, such as when one of the arguments of the construction is deleted.

- (69) áhā gām-á é gwò ɪs
I find:IPFV-1SG 2SG word body
'I believe you.'

- (70) é gǎm-kā? gwò īs
 2SG find:IPFV-1SG.ERG word body
 ‘I believe you.’
- (71) *gwò gǎm-kā? é īs
 *word find:IPFV-1SG.ERG 2SG body
 * ‘I believe you.’ (intended meaning)
- (72) gwò gǎm-kā? īs
 word find:IPFV-1SG.ERG body
 ‘I believe.’ (e.g. you, someone unstated)

Constituent changes are also possible if there are two partarguments; in these instances, however, only the first of the two partarguments may move.

- (73) áhā wǒ'th ā bwà 'kúp
 I help:IPFV ERG.CL1 stomach head
 ‘I’m pondering about a decision (and changing my mind).’
- (74) bwà wǒ'th áā 'kúp
 stomach help:IPFV 1SG.ACC head
 ‘I’m pondering about a decision (and changing my mind).’

In most single-partargument A-Voice constructions which do not take objects, it is not possible to front the partargument, as seen below.

- (75) á'dī ísh ē
 3SG sleep:IPFV eye
 ‘He is sleeping.’
- (76) *ē ísh ā á'dī
 *eye sleep:IPFV ERG.CL1 3SG
 * ‘He is sleeping.’ (intended meaning)

However, even these can be adjusted if an adverbial is added.

- (77) kātīlí? ē ísh ā á'dī
 yesterday eye sleep:IPFV ERG.CL1 3SG
 ‘He was sleeping yesterday.’

Stative clauses

In addition to active verb constructions, partarguments can also occur with stative verbs. Stative verbs denote a permanent or temporary state, and the subject of the construction is a patient of the state. The verbal predicate normally consists of a verb root, the particle *á*, and a reduplicant of the verb root. There is no agreement marking on the verb. Monovalent stative verbs have one single core argument S, which is always in the Absolutive case and precedes the verb.

- (78) á'dī mòk á mòk
 3SG be.fat:IPFV NVP REDUP
 ‘S/he is fat.’

Through phrasal verb constructions it is also possible to have multiple arguments with a stative verb. These are mostly, but not exclusively, with partarguments.

If there is a second argument, the construction more closely resembles that of O-Voice constructions, in that the experiencer O is placed in initial position, and A is marked with the Ergative case. The primary difference between ordinary bivalent O-Voice constructions and bivalent stative verbs is that there is an additional complement construction of *kí* plus the verb root after A; that is, the reduplicated element of the stative verb construction remains. The patient of the verb remains in the Absolutive case, but there is an additional argument occurring after the verb.

(79) gù'b 'bíth ā yól kí 'bī'th
house strong:IPFV ERG.CL1 cost CMPL REDUP
'The house is expensive.'

(80) áhā 'bór ā bwà kí 'bōr
1SG good:IPFV ERG.CL1 stomach CMPL REDUP
'I am very happy.'

Constituent order changes with partargument constructions with stative verbs tend to be somewhat more flexible; example 81 shows an O-Voice partargument in an A-Voice construction.

(81) ē dhěl á'dī kí dhèl
eye glitter:IPFV 3SG CMPL REDUP
'(It's that) he's blind.'

There are also a handful of bivalent verbs in which the partargument occupies the O slot, the construction formally resembling ordinary bivalent A-Voice constructions in terms of agreement marking and verbal paradigms. These constructions are challenging to analyze, as they conjugate according to the normal bivalent paradigms, but they only ever occur with the reduplicant complement.

The verb *'thūp* *īs*, 'to be cold', for instance, is an example of a bivalent (A-Voice) stative verb in which it may only occur in its reduplicated form, and non-reduplicated conjugations are not possible (except in limited instances for third person).

(82) mò 'thūp á 'thūp
place be.cold:IPFV NVP REDUP
'It's cold out.'

(83) mò 'thūp-í'd
place be.cold:IPFV-3SG
'It's (become) cold out.'

(84) áhā 'thūp-á īs kí 'thūp
1SG be cold:IPFV-1SG body CMPL REDUP
'I'm cold (or recovering from a fever).'

(85) *áhā 'thūp-á īs
*1SG be cold:IPFV-1SG body
'I'm cold (intended meaning).'

(86) *áhā 'thūp-á
*1SG be cold:IPFV-1SG
'I'm cold (intended meaning).'

6 Partarguments in other word classes

Partarguments do not just occur with verbs. Word classes in Uduk tend to be rather flexible, and thus it is fairly simple to use a verbal construction in another syntactic class. The nominal form of a verb is simply the verb root, adjectives are verbal in nature, and adverbials are mostly made with the complementizer *kí* and the verb root. Thus, there is very little difference between different partargument constructions types, even when not functioning syntactically as the head of a VP. As mentioned previously, stative verbs can occur with partarguments.

- (87) à sáhàn tēntēr ā bwà kí tēntēr
 CL2 dish be.shallow:IPFV ERG.CL1 stomach CMPL REDUP
 ‘The dish is shallow.’ (UD)

Stative verbs also can modify nouns directly; stative verbs with partarguments retain these even when functioning as adjectives.

- (88) Máná sáhàn-à ’kúl bwà yín?
 where dish-ASS deep stomach that
 ‘Where is that deep dish?’ (UD)

It is also possible to create adverbs from partargument constructions. These can be made with both stative and active verbs. The verb *kṵwá?*, ‘to clear’, combines with the partargument *bwà*, ‘stomach’, to give a meaning of to divide or separate.

- (89) kṵwá? áā tō-m pém bwà mò
 clear.IMP 1SG.ACC thing-ASS 1SG.POSS stomach COMPL
 ‘Divide my things for me.’ (UD)

This also can function adverbially.

- (90) Bwày ’kó kā kṵwá-m bwà
 road stay.PL:PFV with.CL2 clear-ASS stomach
 ‘The roads are intersecting, forking.’ (UD)

The verb *dhàl*, ‘to let, leave’, often takes nominal arguments when used for the negative imperative. For partargument constructions, the partargument is placed first, with the complementizer *kí* between it and the verb.

- (91) dhàl-kí bwà kí ’thōth
 leave:IMP-ITV stomach CMPL worry
 ‘Don’t be needy.’

Further examples of nominal compounding are discussed in section 7 below.

7 Semantic correlations

Despite rather opaque semantics synchronically, patterns of general correlation do exist for the use of the partargument constructions. It is generally impossible to guess the entirety of the meaning based on seeing a verb and its partargument, but looking at a selection of different verbs and partarguments does yield some interesting results. In this section, we focus primarily on the partargument *bwà* and its main uses.

Bwà is one of the most prolific partarguments in Uduk, on its own meaning belly or stomach. There are three main ways *bwà* is used.

The first semantic use is as a classifier, for division into smaller parts (in particular dividing something in half). This typically occurs with A-Voice constructions. Some examples are given in table 6.

Table 6: Examples of division verbs with the partargument *bwà*

cé	to open	cé bwà	to separate after meeting, disperse in different directions
'cīmī'th	to cut into small pieces	'cīmī'th bwà	to divide into small parts
dhër	to walk quickly	dhër bwà	to cut in half lengthwise
'ci'th	to cut, cross	'ci'th bwà	to cut in half lengthwise, divide cross section
kwá?	to clear, wound, break	kwá? bwà	to divide
'twák	to burst, pop, lighten (of light)	'twák bwà	to split asunder (into two)
pwáth	to cut wood for kindling	pwáth bwà	to break into small pieces
tháp	to be bland	tháp bwà	to sort out good/ripe things
cēkēr	to scatter (something)	cēkēr bwà	to scatter (anticausative)

Bwà can also be used as a dummy argument for feelings, often (but not exclusively) in experiencer constructions. *Ā* in the following table refers to O-Voice constructions in which *bwà* functions as the A argument.

Table 7: Examples of feeling verbs with the partargument *bwà*

ó	to say	ó bwà	to want, desire
tó'c	to annoy	tó'c bwà	to trouble
wǎn	to exchange	wǎn bwà	to reconcile
wál	to quarrel	wál bwà	to be unstable, unreliable
dīt	to be heavy	dīt ā bwà	to be heavy-hearted, sad
'kúc	to bend, roll, fold	'kúc ā bwà	to be patient
'bór	to be good	'bór ā bwà	to be happy
kūlūm	to think, plan (something)	kūlūm ā bwà	to contemplate, consider
'kūny	to be pleasant in smell, taste	'kūny ā bwà	to be pleased
shī?	to be bad	shī? ā bwà (or shīn ā bwà)	to be sad, angry, resentful, sorry
tá'c	to toss, vote	tá'c ā bwà	to feel insulted or offended
'té	to not want	'té ā bwà	to dislike (strongly); to reject

The third main use is with nominal compounds, to refer to large or wide areas, or sometimes the inside of an area. This third use also occurs in a limited fashion with adjectives and stative verbs.

Table 8: Examples dealing with wide areas with the partargument *bwà*

cwá	tree	bwáńycwá	forest, woods
gù'b	house	bwáńgù'b	inside of home
ē	eye	bwám ē	face
'cé	ear	bwáńy 'cé	inner ear
shūsh	nose	bwám shūsh	nostril
shō'k	foot	bwám shō'k	sole of foot
à pó?	back	bwámpó?	beyond
à bōr	chest	bwámbōr	front, in front of
à bī	upper arm	bwámbī	armpit
à pí?	soot	bwápi?	field of burned grass
à shō	grass	bwáshō	bush, wilderness
à wāsh	dark, red soil	bwáwāsh	place of red soil
gwátá?	flintstone	bwáńgwátá?	fireplace
tī	already	bwántī	old field planted once
di	to be somewhere; to stay, remain	di bwá	to be empty
'kúl	to be deep (water)	'kúl ā bwá	to be deep (of things)
tár	to be flat, level	tár ā bwá	to be shallow, flat (of things)
tēntēr	to be shallow (water)	tēntēr ā bwá	to be shallow (of things)

8 Conclusion

Partargument constructions are a fascinating area of both semantics and syntax in Uduk. The verbal predicate incorporates primarily body parts in order to classify arguments or create idiosyncratic changes in meaning. They bear a distinct resemblance to both light verb constructions and noun-verb idioms, yet remain distinct from them in a number of properties. Documentation of such a system is likely of interest to any typologist interested in noun categorization as well as argumenthood, due to unusual syntactic properties of these constructions.

West African Kwa languages appear to parallel Uduk's partargument constructions in a number of ways, and a more direct comparison would be interesting to pursue, especially regarding voice alternations and argumenthood of inherent complement constructions.

Other potentially interesting areas to explore would be in grammaticalization and historical development. Gumuzic languages form a relationship with Koman languages at a higher level; despite the more distant genetic relationship, however, there is a surprising similarity in terms of predicate classification. Gumuz makes use of noun incorporation for predicate classification, similar to Uduk. The semantics of the classification appears to follow a similar logic in many instances, such as using 'head' for containers (including liquids in containers) and 'eye' for liquids which are not (necessarily) in containers. 'Eye' in particular shows a surprising parallelism, being used in both languages for sharp objects, liquids, and clothing.

However, the incorporation is much more tightly bound syntactically in Gumuz than in Uduk, as seen in the examples below (taken from Ahland 2010):

- (92) áχ^wa ʔaf-əc
clothes wash-eye
'Wash the clothes!'
- (93) ʔaf-əc məχa
wash-eye blood
'Wash off the blood!'

- (94) ʔaf-ók'w li-c-úá
wash-head POSS-eye-2S
'Wash your face!'
- (95) ʔaf-ik'wós mankia
wash-tooth metal.spoon
'Wash the spoon(s)!'

Ahland (2010) concludes on Gumuz incorporated nouns that "...in the absence of historical data on the Gumuz language or thorough grammatical descriptions of the languages thought to be most closely related to Gumuz, it is difficult to ascertain whether or not Gumuz represents an older system perhaps in decay. If such is the case, one would expect to find evidence of such a system in closely related Nilo-Saharan languages."

As evidence has now been presented on such a system, deeper explorations and comparisons could shed light on how predicate classification can arise in languages, as well as how the earlier Nilo-Saharan system may have looked like, in particular in terms of boundedness. More research and documentation on other Koman languages, however, is vital to understanding any type of historical development.

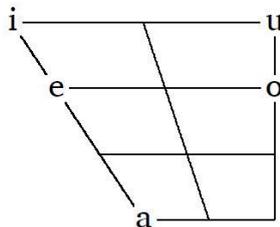
Appendix 1: Phonology

As published data on Uduk is primarily restricted to overviews by Lionel Bender (see e.g. Bender 1983, 1985, 1994), as well as Thelwall's discussion of the phonology (Thelwall 1983), it is worth a brief overview, particularly when new information has been discovered by the author leading to some changes in the phonological description.

Table 9: Consonant inventory¹⁰⁷

	Bilabial	Dental	Alveolar	Post-Alveolar	Palatal	Velar	Glottal
Unaspirated	p p ^w	t [t ^w]	t		c c ^w	k k ^w	ʔ
Aspirated	p ^h p ^{hw}	t ^h	t ^h [t ^{hw}]		c ^h c ^{hw}	k ^h k ^{hw}	
Ejective	p'	t'	t' t' ^w		c' c' ^w	k' k' ^w	
Voiced	b b ^w	d	d d ^w		ʃ ʃ ^w	g g ^w	
Implosive	ɓ ɓ ^w		ɗ ɗ ^w				
Fricatives	(ɸ)		s	ʃ ʃ ^w		(x)	h
Nasals	m [m ^w]		n		ɲ [ɲ ^w]	ŋ ɲ ^w	
Lateral			l				
Trill			r				
Approximant					j	w	

Table 10: Vowel inventory



¹⁰⁷ Consonants in parentheses occur either allophonically or as a result of grammatical processes, and are not phonemes. Consonants in brackets are considered marginal phonemes.

Uduk has a complex phonology, with an extensive consonant inventory and a heavy reliance on tone. Length contrasts are phonemic but rare for both consonants and vowels, and there is no phonemic ATR contrast in the language. There are four basic tonemes in Uduk: L (à), M (ā), H (á), and LH (ǎ); these exist both on the surface and underlyingly. There is an additional rare toneme of HL (â), found in a handful of lexemes; the contours HM (ǎ) and MH (ǎ) primarily exist as a result of grammatical processes, and have not been shown to be contrastive in morphemes roots. Although there are three tone levels in Uduk, due to a complex system of consonant-tone restrictions, minimal tonal sets are extremely limited, and cannot occur in noun or verb roots alone. Only grammatical morphemes can give a full contrast between all three tone levels, in much the same way Mundang has been described (see e.g. Elders 2000 for details).

These restrictions show up in the way of depressor consonant effects, and also create rules in the verbal morphology, important for understanding some of the case marking and agreement alternations. In particular, there is one rule affecting tonal sequences important for the syntax: if a sequence would occur such that an L tone is followed by M, M instead lowers to L. This occurs in both word roots and with a number of grammatical particles, such as the Ergative markers and one of the gender markers. This does not affect compounds or across word boundaries, grammatical particles aside.

Abbreviations

1	1 st person	ICV	Inherent Complement Verb
2	2 nd person	IMP	Imperative
3	3 rd person	IMPRS	Impersonal
ACC	Accusative	IPFV	Imperfective
AGR	Agreement	ITV	Itive
APPL	Applicative	NAS	Nasal
ASS	Associative	NVP	Nonverbal Predicate
CL1	Class I Gender	OBJ	Object
CL2	Class II Gender	PST	Past
CM	Class Marker	PFV	Perfective
CMPL	Complementizer	PL	Plural
COMPL	Completive	POSS	Possessive
CON	Construct	PREP	Preposition
COP	Copula	PRES	Present
CT	Continuous	REDUP	Reduplicated
DEF	Definite	REL	Relative marker
DET	Determiner	SCR	Subject Cross-Reference Marker
ERG	Ergative	SG	Singular
IC	Inherent Complement	V	Verb

References

- Ahland, Colleen A. 2010. Noun incorporation and predicate classifiers in Gumuz. *Journal of African Languages and Linguistics* 31.2: 159–203.
- Ahland, Colleen A. 2012. A grammar of Northern and Southern Gumuz. PhD thesis, University of Oregon.
- Alsina, Alex, Joan Bresnan, & Peter Sells (eds.). 1997. Complex Predicates. Center for the Study of Language and Information, Stanford, CA.

- Amberber, Mengistu. 1995. Functional Verbs in Predicate Formation: Event-Type Hierarchy and Grammaticalization. Berkeley Linguistic Society, I-II.
- Ameka, Felix. 2009. Likpe. In Gerrit J. Dimmendaal (ed.), *Coding Participant Marking: Construction Types in Twelve African Languages*, pp. 215–237. Amsterdam/Philadelphia: Benjamins.
- Beam, Mary S., Betty Cridland & Rasha Angwo. 1963. Gwon this ki 'twam pa mo [Uduk New Testament]. Sudan Interior Mission.
- Beam, Mary S. & Cridland, Elizabeth A. 1970. Uduk-English Dictionary (Linguistics Monograph Series, 4). Sudan Research Unit, University of Khartoum.
- Bender, M. Lionel. 1983. Proto-Koman phonology and lexicon. *Afrika und Übersee* 66: 259–297.
- Bender, M. Lionel. 1985. Gumuz, Koman, Mao and Omotic. In Russel G. Schuh (ed.), *Papers from the Fifteenth Conference on African Linguistics*. Los Angeles: University of California.
- Bender, M. Lionel. 1994. Comparative Komuz grammar. *Afrika und Übersee* 77: 31–54.
- Bender, M. Lionel. 1997. *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom.
- Elders, Stefan. 2000. Grammaire Mundang. PhD thesis, Rijksuniversiteit te Leiden.
- Essegbey, James. 1999. *Inherent Complement Verbs Revisited: Towards an Understanding of Argument Structure in Ewe*. Wageningen: Ponsen en Looijen.
- Greenberg, Joseph H. 1963. *The Languages of Africa*. Bloomington: Indiana University.
- James, Wendy. 1979. *'Kwanim pa: The Making of the Uduk People. An Ethnographic Study of Survival in the Sudan-Ethiopian Borderlands*. Oxford: Oxford University Press.
- James, Wendy. 1988. *The Listening Ebony: Moral Knowledge, Religion and Power among the Uduk of Sudan*. Oxford: Clarendon Press.
- James, Wendy. 2007. *War and Survival in Sudan's Frontierlands: Voices from the Blue Nile*. Oxford: Oxford University Press.
- Neudorf, Susanne. 2013. The Use of Body Part Lexemes in Berta. Paper presented at the Nilo-Saharan Linguistics Colloquium, 22–24 May, 2013.
- Stevenson, Roland. 1944. Notes towards a grammar of the Uduk language. Unpublished Manuscript.
- Thelwall, Robin. 1983. Twampa phonology. In Lionel M. Bender (ed.), *Nilo-Saharan Language Studies*, pp. 323–335. East Lansing: African Studies Center, Michigan State University.
- Tucker, Archibald N. & Bryan, Margaret A. 1956. *The Non-Bantu Languages of North-Eastern Africa*. Oxford: Oxford University Press.
- Tucker, Archibald N. & Bryan, Margaret A. 1966. *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa* (2nd edition). Oxford: Oxford University Press.

THE PHILOSOPHY OF WALKING: MOTION AND THE VERBS OF WALKING IN CHERANG'ANY (KALENJIN)¹

Angelika Mietzner

1 Introduction

Cherang'any is a Southern Nilotic language, being classified as a dialect of Marakwet within the Nandi branch of Kalenjin (Rottland 1982). Survey work by the author indicates that Cherang'any is probably not a dialect of Marakwet, but rather a separate Kalenjin language, classified within the Nandi branch.

Motion events in Kalenjin languages have been touched upon by Mietzner (2009, 2012). This contribution to the description of motion will help to improve the descriptive status of Kalenjin languages and also highlight the inherent semantics of a special field of motion verbs, the so-called verbs of walking.

This paper is an approach to explaining the rich lexicon of these verbs in relation to cultural attitudes towards walking in a society where walking has an extremely positive association.

The term *verb of walking* in this paper is used in order to emphasize the motion of physical locomotion on two feet in an upright position, and thus to distinguish it from the motion verbs in general, meaning a motion by any means. The verbs of walking all have semantic features of, for instance, the intention of the walking process, the emotion shown by the motion or the process of walking at a certain speed. They are syntactically expressed by underived lexical verb stems or may be modified by derivational morphemes, prototypically denoting motion away from or towards the deictic center.

A conspicuously rich lexicon of verbs of walking is used in a society that is known to the world as the “running nation” – the Kalenjin.²

In this paper, firstly, a short introduction to the social and geographical background of the Cherang'any is given, followed by a discussion of some theories from different scientific disciplines dealing with the concept of running.

Afterwards, the syntax of motion events will be described, outlining the wide range of the Cherang'any walking lexicon and also explaining the lexical verbs of walking, as well as the derivational suffixes that turn a non-motion verb into a verb of walking.

Finally, an attempt will be made to explain the phenomenon of Kalenjin people being among the leading runners of the world and the lexicalization of verbs of walking will be discussed. So, a description of Kalenjin society and their attitude towards walking will be the starting point of demonstrating that the culture of walking has had an important impact on the lexicon.

The conclusion is drawn in the last paragraph, giving a theory of verbs of walking embedded in the field of ethnosemantics.

¹ I would like to thank the *Deutsche Forschungsgemeinschaft* (DFG), who is generously sponsoring the project *Die Marakwet-Sprachen (Süd-Nilotisch): Deskription und Vergleich*. I would also like to thank my language assistant Benjamin Kimeli Rotich who has put intensive thought into the idea of this article, and Jacob Suter. I am thankful to Alexandra Aikhenvald, Andrew Chelimo, Helga Schröder, Anne Storch and Helma Pasch for their fruitful comments on previous versions of this paper.

² Everybody who is familiar with athletic events knows the Kalenjin as one of the runners spearheading the list of the world's best athletes. (The gold and bronze medals of the marathon discipline within the 2012 Olympic Games were won by Kalenjin runners).

2 The importance of walking in the Cherang'any culture

2.1 Social and geographical environment

The Kalenjin live in the Rift Valley Province in Western Kenya. The area is mountainous. The Marakwet, under which Cherang'any is classified, originally settled in the escarpment of the Rift Valley between Tot in the North and Iten in the South, where the escarpment goes down to the Kerio valley. The Cherang'any split from the Marakwet and moved further west to the Cherang'any hills, which are a series of hills that rise to a peak of 3,500 meters. The climatic conditions in the hills are the reason for the area being highly fertile. Roads and paths are rough; some of the fields are on steep slopes.

Walking and especially running is an important social aspect in the life of the Cherang'any. It is considered socially desirable by many members of Kalenjin-speaking groups. This may have given rise to professional running, in which some athletes are successful and extremely popular, getting a disproportionate number of medals in marathon competitions and also in other running disciplines. Running is taken very seriously as a profession within Kalenjin society.

2.2 Social and historical background

The link between the Kalenjin and athletics may of course be a coincidence. People see the act of walking and running as something good and necessary.³ Often can one see primary school pupils running and jumping together with their friends on their way to and from school. This is a part of the culture and is considered a positive habit.

In former times, the daily work of Cherang'any people could be carried out only during daytime due to dangerous game such as buffaloes, hyenas, lions or leopards that lived in the Rift Valley. The limitation of work to the daytime forced the people to cover distances at a greater speed in order to return home before sunset, and additionally the terrain was risky and rough. This necessity of being fast is explored further in § 3.3.4 (speed of motion).

The hunting of game was conducted by the Cherang'any men and stealth was crucial in order not to scare the animals. Therefore, verbs of walking in a certain manner were used to describe the motion conducted. These verbs will be given in § 3.3.2 (manner).

The circumcision ceremonies and the idea of getting married very soon after circumcision led to the necessity for the boys to present themselves in a showy way in order to attract the girl's interest and to present themselves to her parents as an appropriate candidate for their daughter. This was done by walking around, showing off and presenting oneself, as will be described in § 3.3.1 (telic verbs). Manner and speed are only two of four inherent semantics that are expressed with the verbs of walking. The full range will be shown in § 3.3.

The distances that have to be covered in a day's work for Cherang'any people are quite considerable. The fields belonging to a household are not always near the house; cattle – the economic basis of life – have to be brought to different places for grazing or salt licks. Bee-keeping, which is another part of agriculture, also requires covering some distance when building bee-hives, controlling them and harvesting the honey.

In the following, motion will be described in order to show its syntactic behavior and the marking of goal, source, figure and ground. Afterwards the verbs of walking will be discussed.

3 Motion

Cherang'any is a VSO language. The agglutinating verb is highly interesting, having up to seven morphological slots, including tense, conditional, negative, subject prefix, an impersonal marker, the stem, ten possible derivational suffixes, an aspect marker and an object or referential marker. Some of these morphemes can appear together, others are mutually exclusive.

³ When asking my co-worker after a long working day whether I could give him a lift back home, he told me that he had not done his limit of (walking) exercise for that day and would prefer to walk back home.

Cherang’any is a tonal language, having both lexical and grammatical tone. Case is marked tonally on nouns, as Cherang’any is a marked nominative language.

3.1 Motion-evoking derivational suffixes

In Cherang’any, motion can also be described by lexical motion verbs, as well as by extended non-motion verbs, when the suffixation of derivational morphemes triggers a motion reading. Directional suffixes in Kalenjin languages have been described by Creider (2002) and Mietzner (2009, 2012).

Two directional morphemes denote the directions ‘towards’ (ventive) and ‘away from’ (itive) the deictic center. Itive and ventive can be modified by a mobilitive morpheme *-aan-*, including the motion towards (mobilitive-ventive) or the motion away from (mobilitive-itive) a deictic center⁴. These verbal suffixes are *-u-* (ventive, motion towards), *-ta-/te-/tA-* (itive, motion away), *-aanu-* (mobilitive-ventive, doing something while moving towards) and *-aate-/aati-* (mobilitive-itive, doing sth. while moving away). These morphemes – with variation – are found in all Kalenjin languages and related languages like Datooga (Kiessling 2007). Furthermore the applicative suffix *-ci/-cini* can express a motion that includes a goal or the crossing of a boundary (Mietzner 2008: 160).⁵ The examples in Table 1 show the various uses of itive, ventive and applicative.

Table 1: Cherang’any verbs with directional and mobilitive suffixes

itive	wir-tá throw-ITV	‘s/he throws away’ (from deictic center)
ventive	wir-u throw-VEN	‘s/he throws towards’ (deictic center)
mobilitive itive	tíl-állí cut-MOB.ITV	‘s/he cuts while going’ (from deictic center)
mobilitive ventive	tíl-állnu cut-MOB.VEN	‘s/he cuts while coming’ (towards deictic center)
applicative	wir-cí throw-APPL	‘s/he throws into/onto’

The itive and ventive suffixes can be attached to motion verbs in order to indicate a direction, but they can also be attached to non-motion verbs, triggering no motion but rather a change in the semantic function, such as, for instance, a comitative, as example (1) shows.⁶

- (1) ḡgì ám-tè kimyèèt ḡwéék
COND eat-ITV ugali:NOM vegetable
‘Please eat the ugali with the vegetables.’

Mobilitive itive and mobilitive ventive suffixes can evoke motion when suffixed to a non-motion verb (2), (3).

- (2) kɔ-síḡjild-állti / kɔ-síḡjild-állnú cìito
P2-spit-MOB.ITV P2-spit- MOB.VEN person
‘While going away / coming, the person spit saliva.’

⁴ The vowel harmony system in Cherang’any and in other Kalenjin languages (Rottland 1982: 81ff.) is non-symmetric, which means, that it is dominant vs. recessive. The dominant feature is the [+ATR] and the recessive one is [-ATR] feature. Vowels are: [+ATR] [i] [u] [e] [o] [ɛ] [ɔ] and [-ATR] [i] [u] [e] [o] [a]. Vowel harmony is phonetically difficult in case of /a/ and /o/. The articulation of [ɛ] and [ɔ] has converged in Cherang’any.

⁵ These suffixes undergo manifold diverse phonological processes that can result in a total omission of, for example, the ventive suffix, or evoking a change of vowel quality and tone.

⁶ There are more semantic extensions with the itive suffix for other Kalenjin languages (Mietzner 2009: 180ff.).

- (3) nám-λλτέ
hold- MOB.ITV
'walk away while holding something/someone'

Table 2 presents non-motion verbs with the mobilitive morpheme. This means of evoking motion is very productive and the table is a selection of semantically highly interesting verbs.

In some of the mobilitive verbs in the table, the suffix has an intentional or purposive connotation. This is obvious in, for example, *wàkààté* 'walk away while making a lot of noise to draw many people's attention to something grave' or *tyémààté* 'walk away while slashing/attempting to overcome something challenging to you'.

Table 2: Non-motion verbs with mobilitive suffixes

wàk-ààté	'walk away while making a lot of noise' (intention: to draw many people's attention over something grave) (> <i>kèè-wác</i> 'cry out')
síngwàl-λλτέ	'walk away while limping' (derogatory) (> <i>ke-síngwàl</i> 'limp')
mwáit-ààté	'walk to several places while giving people information' (> <i>kèè-mwai</i> 'say, speak')
milt-ààté	'roll/push someone who tries to stand up and get away (while both are going away)' (as a policeman conducting a rough arrest) (> <i>kèè-mil</i> 'roll')
tyém-ààté	'walk away while slashing/attempting to overcome something challenging' (> <i>kèè-tyém</i> 'try')
mís-λλτί	'walking while holding a light that goes on and off at intervals' (> <i>kíi-mís</i> 'put off light')
kí-cút-λλτ	'to be dragged carelessly' (derogatory) (> <i>kí-cúút</i> 'pull')
kúrt-λλτέ	'walk while calling out to people' (intention: to convene/assemble) (> <i>kèè-kúr</i> 'call')
képòt-λλτ	'walking and repeatedly getting lost while trying to reach an unfamiliar destination' (> <i>kè-pót</i> 'be lost')
kàs-ààté	'walk while keeping ears open so as to gather information' (> <i>kèè-kás</i> 'listen')
kwòŋ-λλτέ	'walk absentmindedly (> <i>kwòŋ</i> 'be absent')
sís-ààté	'walk silently' (> <i>kèè-sís</i> 'be/keep quiet')

3.2 Motion verbs and events

3.2.1 Marking of goal and source

In a motion event, the goal is marked with the preposition *ókói* 'to, up to, until', which is a locative and temporal preposition and triggers the unmarked absolutive form of the noun. *ókói* is used as 'to, towards', implying that the goal will be reached (4). If the situation and the goal are clear and specific, *ókói* can be omitted (5) so that in these two sentences no difference is evoked by the use of *ókói*. If the motion to or towards the goal is incomplete, i.e. the goal will not be reached, the verb *kùtókìgíéi* 'to head towards' must accompany the motion verb (6).

- (4) kà-wééc=kéy **ókói** ñàŋàséét
P1-return=refl prep village
'They returned to the village.'
- (5) kà-wá ñètèt ñàŋàséét
P1-go boy:NOM village
'The boy went to the (specific) village.'

- (6) kÌ-wá ñètètèt kù-tókcì=kéy ñàjàsétèt
 P3-go boy:NOM 3-head_towards=REFL village
 'The boy went towards the village.'

Any motion verb can be specified with *kùtókcìkékéy* except for *cun* 'fall to the ground', where the direction towards the ground is naturally implied. The use of the applicative suffix *-ci* would express the falling to a different ground from the actual one, e.g. into water.

The prototypical function of the instrumental *-e-* is to mark a noun or a noun phrase as the instrument being used by the agent to carry out the event that is described by the verb. But the instrumental suffix can – if not expressing an instrumental situation – also introduce a location role with verbs that are not motion verbs.

Most commonly, the location role is used when marking the goal, as can be seen in (8), (9).

- (7) ákèt-é-é tùgà kèèwa-ńín
 graze-INST:LOC-ASP cows flat_place-DEM2
 'The cows are grazing **in** a place.'

- (8) órèrèn-é-é òinó kityó
 play-INST:LOC-ASP river only
 'She is playing **at** the river.'

The *source of a motion event* is not marked if a verb like *wéc* 'turn away from, move away from, remove' is used, as this already implies the leaving of a source (9), (10).

If the motion verb does not imply a departure from a source, the source must be marked with the verbal instrumental suffix *-e-* (11), (12).⁷

- (9) kò-wéci=kéy cèbyósétèt ím wèli mí mààt
 P2-move_away=REFL woman in place COP fire
 'The woman moved away from the fire.'

- (10) kò-wéci=kéy ñètúindò kù-ñèt-é-é péék
 P2-move_away=REFL lion:NOM 3-stay-INST:LOC-ASP water
 'The lion walked away from the water.'

- (11) kò-cún-ùn-é-é bókúsit páàs
 P2-fall-VEN-INST:LOC-ASP box:NOM bus
 'The box fell down from the bus.'

- (12) kò-púrùr-ùn-é-é kítít tàrítètèt
 P2-fly-VEN-INST:LOC tree bird:NOM
 'The bird flew from the tree (towards speaker).'

3.2.2 Marking of the ground NP

Looking at the marking of the ground NP, it again becomes obvious that Cherang'any puts emphasis on the lexicalized semantics of verbal constructions.

⁷ Non-motion verbs can also mark the source from where the action is performed:

kù-ál-è hotèli supermarket nde kíóski
 3-buy-INST:LOC hotel supermarket or kiosk 'to buy from a hotel'

kò-am-è mèsà ñuik
 3-eat-INST:LOC table vegetables 'They ate the vegetables from the table (yesterday).'

- (21) ká-cùùt kípòkòrèèt kòòt
 P1-enter big_rat:NOM house
 'The big rat enters the room.'
- (22) ka-man-dá lèkwèèt sánj
 P1-go_out-ITV child:NOM outside
 'The child leaves the house.'
- (23) ké-maŋ-ùn-é
 INF-exit-VEN-ASP
 'to come out of an enclosure'

3.3 Verbs of walking

As already documented for Keyo, a related Kalenjin language (Mietzner 2007), Cherang'any also is rich in motion verbs.

Particularly striking is the fact that, within the group of motion verbs, there are verbs denoting just the motion of walking itself. They imply inherently different semantic strategies of walking, like for example an intention in contrast to a non-intention (3.3.1), which is pursued with the act of walking, the manner in which the walking is conducted (3.3.2), a kind of walking expressing an emotion like "walk angrily" (3.3.3), or simply the speed with which the walking is performed.

Worth highlighting are verbs of walking which have a suppletive stem. Suppletive verb stems are not common in Cherang'any. The verbs *-lapat-* (sg) *-rwei-* (pl) 'to run', *-wa-* (sg) *-pa-* (pl) 'to walk' and *kuco* (sg) *ke-pka* (pl) 'to come' are the only three suppletive verb stems found in the language so far. It is significant that suppletive verb stems are very few and all three of them are verbs of walking. This apparently shows the importance of differentiating in this special semantic field, which brings us back to the importance of walking in the culture, to be explained in the next paragraph.

In the following sections the verbs of walking will be explained and listed according to their inherent semantics, which are telicity, manner, emotion and speed.⁹

3.3.1 Telic verbs of walking

Telic verbs of walking indicate that somebody is walking and is pursuing a target by doing so. The telic verbs of walking imply a physiological intention and can mark a goal or a source. *ki-mbiri:s* 'walk and sing (to attract the attention of a girl) while walking' (24) describes the act of walking or strolling around with the purpose of attracting the interest of a girl. The verb could also be translated or interpreted as 'showing off'. *mbiriis* is a gender-specific verb, for it is an act that can only be performed by boys.

- (24) mbìr-iis-é mùrèn ku-péét-i sígíròino
 walk_showing_off-INTR-ASP man:NOM 3-go:pl-ASP Sigiroino
 'Young men are walking and proudly singing as they walk towards the place of Sigiroino.'

A purpose is also indicated with *ke-más* 'walk and hit grass', which is the act of hitting grass in the morning with a stick while walking in order to get the dew out of it (25).

- (25) kú-ri kírók a=kó-más-àtè rɔwɔndéét
 3sg-break stick and=3sg-walk_hit_grass-ITV.MOB dew
 kù-tók-ciní=kéy káà
 3sg-head-APPL=REFL home
 'He broke a stick and used it to hit the grass to scatter the dew (along his way) as he headed home.'

⁹ The translation of the examples is quite problematic. For this reason I based the translation on the explanation of my co-workers, who assured me that walking is a central aspect of the action described.

3.3.2 Manner

There is quite a large number of verbs of walking that express the manner in which locomotion is performed. Some of them are onomatopoeic; others are basic verb stems and can take the itive suffix, which indicates a given direction of the walking. They all describe a very particular manner of walking.

<i>kè-tákták</i>	‘sneak’
<i>kè-turtur</i>	‘walk with loose clothes’
<i>kè-tartarin</i>	‘walk, stagger as if drunk or sick’ (27)
<i>kè-ndelde</i>	‘tiptoe’ (27)
<i>kè-sab</i>	‘walk slowly, silently’ (28)
<i>kì-kway</i>	‘walk in large steps’ (29)
<i>kè-yap</i>	‘walk – rocking regularly from side to side’ (30)
<i>ke-róób</i>	‘walk carefully; balance; sneak along a wall or rope or anything’
<i>kì-iráán</i>	‘run, jump’

- (26) kyá-á-tartárin-ààti ʌ-wéti kopkéric
 P3-1sg-walk_stagger-ITV.MOB 1sg-go hospital
 ‘I was staggering on my way to the hospital.’
- (27) kú-kéndèlɔda okòí kòmòstá ni kí-ìngét
 3-walk_tiptoe until side REL P3-know
 ‘He tiptoed until the side which he knew [...]’
- (28) [...] ku-ɲèèt Kípcèrsiɲ á=kù-sàb-tá kwó-kwà-ís.
 3-wake_up Kipchesiny and=3-walk_silently-ITV P2-go_in_large_steps-INTR
 ‘[...] Kipchesin’y woke up and walked silently with large steps.’
 (trans: to go lone-hunting) (Logop: 12)
- (29) kí-ɲgù-ma-kù-màn-dá sááɲ, kú-kwàɲ á=kù-yép-è íít
 P3-COND-almost-3-come-ITV bush 3-walk_in_large_steps and=3-train-ASP ear
 ‘When he almost emerged from the bush, he walked in large steps and trained his ear.’ (trans: he took long, silent strides while training the ear)
- (30) kù-lè cí ki-ndò-wééc “ó-ɲgè-yáp kírtíni” ...
 3-say people P3-lead-O1pl 2pl-COND-walk_tossing bushes
 ‘The people leading us said “you should walk tossing from side to side (through the?) bushes” ...’

3.3.3 Emotion

The verbs of walking that express emotion constitute only a small class of two verbs. Whereas *cumbiren* ‘walk proudly’ (31) is not modified by a directional suffix, *nerek-te* ‘move away angrily’ (32) is a composition of the verb *nerec* ‘be angry’ and the itive derivational suffix.

- (31) kú-mí sòlwéét=kéy kipsépèr kù-cùmbirén-é ʌ=kù-máryèn-è.
 3-be fool=REF Kipseper 3-walk_proudly-ASP and=3-whistle-ASP
 ‘[...] the fool Kipseper was walking proudly and whistling.’
- (32) kù-néréek-tá kù-rúp-è múréník-cúút àlák.
 3-walk_angrily-ITV 3-follow-ASP men-DEM others
 ‘[...] and he (the leader) moved angrily away, following the other men.’

The suffixing of the itive morpheme is not productive with emotion verbs. Here the polifunctionality of the itive suffix enters into the equation. When attaching the itive to the emotion verb *ke-paipai* 'be happy', the meaning becomes resultative 'do sth. in order to make sth. better, happier, more wonderful'.

3.3.4 Speed of motion

Indication of speed is an important factor with verbs of walking. The suppletive verbal root *-lapat-* (sg)/*-rwei-* (pl) 'run, run away' expresses fast running. Athletes who run professionally are called *lapatiin*.

(33) kú-**lápát** cii tógòl kú-sór=kéy.
3sg-run:sg people all 3sg-rescue=REFL
'Everybody runs to save himself.'

(34) ki-**rwéi** aláké ku-tókci=kéy oinó,
P3-run:pl others 3-head_towards=REFL river
'Some ran towards the river [...]'

cak 'walk fast' can occur with the itive suffix, emphasizing the motion away. However, it is not necessary to attach the itive; itive motion is already implied in *cak*.

(35) kó-**càk-tá** apuun okoi uurwet-aap lómóiwò - kú-mùj.
3-walk_fast-ITV therefore until shade-GEN lomoiwo 3-rest
'He hurried until he reached the shade of the Lomoiwo tree – he rested.'

'run/walk in the fastest manner' has the metaphorical meaning 'to invade'. It can either be used in an itive construction or in a ventive one. *bók* 'invade in the fastest manner' has traces of the deleted ventive suffix, namely the [-ATR] vowel and the high tone, whereas in *puk-te-kei* (36), the itive suffix is retained before the reflexive suffix.

(36) [...] kó-**pòk-tá**=kéy kù-**ηèèt-è** kòmòswèk tógól
3-walk_very_fast-ITV=REFL 3-come_from-INST sides:NOM all
ak kù-kwèèn pòinèt.
and 3-center antelope
'[...] they ran to it very fast from all sides and surrounded the antelope.'

Finally, *ki-cwe-te* 'run away extremely fast' is the fastest way of running away.

4 Development of culture-specific language features

So far we have described the lexicon and grammar of directional and motion-evoking verbal suffixes, verbs of walking and the cultural surroundings, habits and attitudes of the Cherang'any people concerning walking and running.

The two factors – describing grammar and describing culture – and the connections between the two, is what is known as the field of ethnosyntax (Enfield 2002).

4.1 Some thoughts on hypotheses and explanatory models

From the late 1980s onwards 70 to 80% of the winners of international long distance races were from Kenya, which gave rise to many speculations and explanatory models for the reason for the success of Kenyan athletes. These speculations led to interviews, research and statistics on the origin of this success. Many of the models sound reasonable and are based on solid evidence, especially those having to do with the altitude of the dwellers of the Rift Valley. The thin air at this high altitude causes

powerful lungs and hearts in order to compensate for the deficiency of oxygen. This has been realized by long-distance runners from all over the world, who nowadays travel to Kenya to prepare and train at high altitude in order for their bodies to adapt to these conditions and achieve better results in low altitude competitions. Several high altitude training camps are located in Iten and they host sports tourists from all over the world.

Other models discuss how local culture and tradition may have facilitated the physical state. Manners (2007: 40f), for instance, assumes that traditional practices of the Kalenjin have served as a natural act of selection to bring out the strongest runners. He explains that cattle raids in former times, where people had to run up to 60 km per night, made people strong. This, combined with the rough initiation rites, where boys and girls were laughed at when crying or wailing, combined to create a strong and physically well-developed people.¹⁰

Medical research and statistics are not the only explanations that were discussed for the phenomenon but will, however, not be taken into consideration in the case of the *verbs of walking*.¹¹

These explanatory models are far from being what this article wants to reinforce, but they show that the tradition and the discussion of running in the Kalenjin society has a long history.

Whatever theory or hypothesis we look at, it is obvious that the Kalenjin attract the attention of the world due to their results in running competitions.

4.2 The importance of the Cherang'any language

The occurrence of the itive and ventive directional morphemes are a widespread feature within Nilo-Saharan languages, and therefore not a particularly Cherang'any feature. So it is rather the semantics of the verbs of walking and the semantic polifunctionality of the directional morphemes that can be linked with the culture and the culture-specific highlighting of locomotion.

Ethnosyntax, in its narrow sense, refers to the direct encoding of cultural ideas in the semantics of morphosyntax (Enfield 2002: 7). There are cultures like the Ku Waru of Western New Guinea where binary structures are a common feature of the social system. These binary structures are expressed in the language where pairing is always used when speaking of social groups (Rumsey 2002: 261), and where number distinctions include dual marking as a logical effect of the social system of pairing (ibid: 263).

Other cultures display in their language a distinct sense for spatial orientation, like the Belhare in Nepal, who have many lexemes for all kinds of orientation, cardinal directions or other conspicuous features in the landscape (Bickel 1997).

Nilotic shows various ways of encoding the motion of walking. As can be seen in Heyking & Storch (2007: 103ff.), the Western Nilotic languages Luwo and Boor exhibit a large range of ideophones for motion events. The authors describe how the languages do not have very salient verbal derivation or aspect marking in their motion expressions, but they have a highly developed vocabulary of ideophones encoding motion (108).

The argument that we should take a deeper look into other ethnic groups of the African continent, and the languages of groups from which other athletes come from (e.g. Ethiopian languages), is not one that I intend elaborate upon in this article. There is no intention to demonstrate that Cherang'any has more interesting lexical items or syntactic constructions in the semantic field of motion than other languages of the area. The intention is to show that Cherang'any uses its potential to widen its repertoire in this sector. That languages use these strategies in various fields depending on the importance of a social or sociolinguistic parameter has been discussed by Storch (Lüpke & Storch 2013: 125ff). The Cherang'any do not make use of their multilingual opportunities but rather strengthen their own language. For Cherang'any it is not only the development of motion and walking expressions that typifies their language, there are also other aspects of spatial orientation, motion and place marking that point to the fact that "their language is likely to mirror, to some extent, the habitat

¹⁰ The editors of the volume themselves comment on the model of Manners as an imaginative and entertaining hypothesis (Pitsiladis et al. 2007: 3).

¹¹ For further theories discussed see "The Atlantic" <http://www.theatlantic.com/international/archive/2012/04/why-kenyans-make-such-great-runners-a-story-of-genes-and-cultures/256015/> (accessed 1/10/14).

of the speech community and the way in which its speakers live and view the world” (Dixon 2012: 439).

5 Conclusion

The purpose of this paper was to show that Cherang'any displays a great variety of motion expressions, especially with verbs of walking and the directional suffixes that can either direct or cause motion.

Furthermore, the paper has demonstrated that culture and language go indivisibly together and that – in the case of running or walking in Cherang'any society – theories about physical states or habits can be leveraged with ethnosemantic studies.

Following Wierzbicka, who has stated (1979: 313) that “every language embodies in its very structure a certain world view, a certain philosophy”, it could clearly be shown that the philosophy of walking is undoubtedly reflected in the Cherang'any lexicon.

Abbreviations

1 2 3	1., 2., 3. person	LOC	locative
APPL	applicative	MOB	mobilitive
ASP	aspect	NOM	nominative
COND	conditional	O	object
COP	copula	P1,P2,P3	Past 1,2,3
DEM1	proximal Demonstrative	pl	plural
DEM2	medial Demonstrative	PREP	preposition
GEN	genitive	REF	referential
IMPRS	impersonal	REL	relative
INF	infinitive	REFL	reflexive
INST	instrumental	sg	singular
INTR	intransitive	VEN	ventive
ITV	itive		

References

- Bickel, Balthasar. 1997. Spatial operations in deixis, cognition, and culture: where to orient oneself in Belhare. In Jan Nuyts & Eric Pederson (eds.), *Language and Conceptualization*, pp. 46–83. Cambridge: Cambridge University Press.
- Creider, Chet. 2002. The semantics of participant types in derived verbs in Nandi. *Revue Québécoise de linguistique* 31.2: 171–190. <http://id.erudit.org/iderudit/009316ar> (last access 20.4.2015).
- Dixon, Robert M.W. 2012. *Basic Linguistic Theory* (Further grammatical topics, 3). Oxford: Oxford University Press.
- Enfield, N.J. 2002. *Ethnosyntax: Explorations in Grammar & Culture*. Oxford: Oxford University Press.
- von Heyking, Beatrix & Anne Storch. 2007. Motion in Northern Lwoo: Lwoo and Boor comparative case studies. In Angelika Mietzner & Yvonne Treis (eds.), *Encoding Motion: Case Studies from Africa* (Annual Publication in African Linguistics 5), pp. 89–112.
- Kießling, Roland. 2007. Space and reference in Datooga verbal morphosyntax. In: Payne, Doris & Mechthild Reh (eds.), *Advances in Nilo-Saharan linguistics. Proceedings of the 8th Nilo-Saharan linguistics colloquium*, University of Hamburg, August 22-25, 2001. S. 123-142.

- Levinson Stephen C. & David Wilkins (eds.). 2006. *Grammars of Space: Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press.
- Lüpke, Friederike & Anne Storch. 2013. *Repertoires and Choices in African Languages*. Boston/Berlin: de Gruyter.
- Manners, John. 2007. Raiders from the Rift Valley: cattle raiding and distance running in East Africa. In Yannis Pitsiladis, John Bale & Craig Sharp (eds.), *East African Running: Toward a Cross-Disciplinary Perspective*, pp. 40–50. New York: Routledge.
- Mietzner, Angelika. 2009. *Räumliche Orientierung in nilotischen Sprachen: Raumkonzepte – Direktionalität – Perspektiven*. Cologne: Köppe.
- Mietzner, Angelika. 2012. Spatial orientation in Nilotic languages and the forces of innovation. In Angelika Mietzner & Ulrike Claudi (eds.), *Directionality in Grammar and Discourse: Case Studies from Africa*, pp. 165–175. Cologne: Köppe.
- Pitsiladis, Yannis, John Bale & Craig Sharp. 2007. *East African Running: Toward a Cross-Disciplinary Perspective*. London: Routledge Chapman & Hall.
- Rottland, Franz. 1982. *Die südnilotischen Sprachen. Beschreibung, Vergleichung und Rekonstruktion*. Berlin: Reimer.
- Rumsey, Alan. 2002. Aspects of Ku Waru ethnosyntax and social life. In N.J. Enfield (ed.), *Ethnosyntax: Explorations in Grammar & Culture*, pp. 259–286. Oxford: Oxford University Press.
- Wierzbicka, Anna. 1997. *Understanding Cultures Through their Key Words: English, Russian, German and Japanese*. New York: Oxford University Press.
- <http://www.kenyarunners.com/pages/167371/page167371.html?refresh=1112199340095> (last access 20.4.2015).
- <http://www.theatlantic.com/international/archive/2012/04/why-kenyans-make-such-great-runners-a-story-of-genes-and-cultures/256015/> (last access 20.4.2015).

PERSPECTIVES ON NILOTIC VERB COMPOSITION: WHY CAN'T WE AGREE ON THE MAA VERB?¹

Doris L. Payne

At first glance, Nilotic languages appear to display remarkable variation in word structure. Some languages have characteristically short words, while others are described as having a dozen or more position classes in their verbs. For Maa (Eastern Nilotic), there has been a lack of consensus on how the verb is structured. This study intends to further our understanding of issues in Nilotic word formation by looking at the Maa verb against the background of what has been described about other branches of the family.

Differences in Nilotic word shape might be first appreciated by a rough comparison of phonological length. Consider the number of syllables per word in polished translations of two sentences from token languages of each branch of the family. In English, the New International Version of Matthew 6:25 reads: *Therefore I tell you, do not worry about your life, what you will eat or drink; or about your body, what you will wear. Is not life more important than food, and the body more important than clothes?* This English excerpt has 1.26 syllables per word average (range 1 to 3 syllables; n=38 words). In comparison, Dinka Rek (Western Nilotic) appears to have shorter word-length than English, and Maa (Eastern Nilotic) and Datooga (Southern Nilotic) have considerably longer word length.²

- (1) Dinka Rek: 1.18 syllables/word (range 1 to 2 syllables per word; n=38 words). *Këya, alëk we bāk cī ye dieër kē bāk cam, ku kē bāk dek, rin bī wek rēër we pīr, tēdē ke we dieër kāk bāk cej. Wēi aṅuēën tēnē miēth. Ku guōp raan cī ṅuēën tēnē alēth?* (<http://www.christusrex.org/www1/pater/wolf3-pdf/dinka-rek-nt.pdf> by Bible League)
- (2) Maa: 2.54 syllables/word (range 1 to 6, n=30 words). *Enkaraki ina pe ajoki intae, eminyamaliki enkisui inyi, aajo ainyoo inyanya, anaa ainyoo iokioko, anaa too seseni linyi aajo ainyoo inchoishopo. Ai melusoo enkisui alang endaa, nelusoo osesen alang nkilani?* (*Biblia Sinyati* by Bible Society of Kenya)³
- (3) Datooga: 3 syllables/word (range 1 to 5, n=23 words). *Aba sareani geayeeshshineegwa, adoodeawiissi eawegeagwa, aba gichchea ayaaga ghana gichchea aleaha, ghana gichchea ookwealsiina sasanukkeagwa. Heappi eaweega mwang'ala hamitta, nea saseeda amwang'ala hang'weakka?* (<http://www.bible.is/TCCBST/Matt/6> by Bible Society of Tanzania)

Of course, syllables are not morphemes and when one looks at morphological composition, the Nilotic branches might actually be more similar. One typologically striking thing about some Western Nilotic verbs is that, despite a tendency toward one syllable per word, verb words can still be multi-

¹ I am grateful to Leonard Ole-Kotikash, Keswe Ole-Mapena, and many other Maa speakers, without whom this paper could not have been written. Financial and logistical support for research was provided by Fulbright Foundation grants (1993–1994 and 2009–2010), and NSF grants SBR-9616482 (1987–1999) and SBR-9809387 (1998–2004), and by the University of Nairobi, the Nairobi Evangelical Graduate School of Theology, SIL International, the University of Dar es Salaam and Mkwawa University College of Education.

² The writing systems used in these translations do not mark tone. For this exercise, a “syllable” was determined by counting the number of vocalic portions separated by consonantal strings. Thus, Dinka Rek *aṅuēën* was counted as “2 syllables,” and Datooga *geayeeshshineegwa* as 5. English syllables were determined according to the author’s pronunciation.

³ From here on, Maa data is written in a modified practical orthography with acute accent indicating High tone and Low tone unmarked.

morphemic, expressing some or all of the same morphological categories as their Eastern and Southern cousins.

Section 1 of this paper briefly reviews relevant typological concepts. Section 2 notes morphological categories salient in verbs across all three branches of the Nilotic family and gives a brief taste of Nilotic verb typology by briefly reviewing claims about selected languages from each branch. Section 3 gives an in-depth look at Maa verb formation. While Maa verbs may at first glance appear quite agglutinative, multiple factors conspire to make a position class analysis unsatisfactory. Some parts of the verb might be viewed as much more fusional. The category of aspect shows up in multiple dimensions of what I will call the verb core. Mood and person occur outside the core where there is indeed a more agglutinative profile. Though much research remains to be done on all branches, Maa verb structure may display a somewhat intermediate typological profile, between Western fusional and Southern (perhaps) more agglutinative patterns.

1 Morphological typology

Classic morphological typology (Schleicher 1861, Sapir 1921, Greenberg 1960, Comrie 1989) characterizes word types according to two major scalar parameters: the **index of synthesis** and the **index of fusion**.

At one end of the index of synthesis are *isolating* word types (or languages) which tend toward one morpheme (and one syllable) per word, as in Vietnamese and Classical Chinese. At the other end of this scale, *synthetic* word types (or languages) have multiple morphemes per word. Some highly synthetic languages allow *polysynthetic* words with multiple roots or root-like morphemes (and typically many syllables per word).

The index of fusion applies to just synthetic word types. This parameter concerns how separable morphemes are within words. In *agglutinative* words morphemes are easily separable, as in Swahili. It is also comparatively easy to give one gloss to each isolatable piece. Such word types lend themselves well to “item and arrangement” or “position class” analyses in which particular positions in a possible word can be numbered for their fixed relative order, and labeled for the paradigm of morphemes that occurs exclusively in that position. Dimmendaal (1983) describes the Turkana (Eastern Nilotic) verb in this way, as in (4). This position class analysis and the particular verb illustrating it have 12 meaningful morphemes plus an epenthetic (EP) vowel *-a*. One might object that the tonal tense morpheme doesn’t have a strictly linear position; but this point notwithstanding, Dimmendaal presents the morphemes as cleanly separable.

(4) Turkana (Eastern Nilotic; Dimmendaal 1983: 96)

1	2	3	4	5	6	7	8	9	10	11	12	13
pɛ-	k-	a-	`-	is-	ilot	-aan	-a	-kin	-it	-o	-tò	-rì
NEG	MOV.K	PERS	TENSE	CAUSE	ROOT	HABIT	EP	DAT	ASP	VOICE	NUM	SUBJN

‘that they did not force me to do the washing all the time’

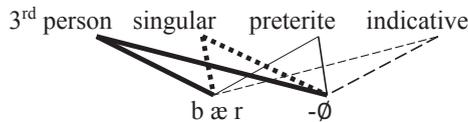
A synthetic *fusional* word structure also has multiple morphemes but morphemes are not so easily separable. For example, Latin noun endings fuse case, gender, and number (Householder 1947), and Latin verb forms fuse aspect-mood, person, and number (Matthews 1972). The Old English partial verb paradigm in (5) demonstrates a clearly fusional problem. The “stem” occurs linearly before the “ending”, but morphological categories of finiteness, person, number, tense, and mood affect changes in both the stem vowel and the ending. In (5), the heading for each row highlights a morphological category that is equally distributed across stem vowel and suffix. That the indicated category is so distributed can be appreciated by contrasting the word *bær* ‘he/she bore’, which has a “zero” ending, with each of the other items in the paradigm in (5).

(5) Old English (partial paradigm)

	bær	‘he/she bore’
FINITENESS	ber-an	‘to bare’
PERSON	ber-e	‘I bear’
PERSON	bæ̅r-e	‘you [sg.] bore’
NUMBER	bæ̅r-on	‘they bore’
TENSE	bir-þ	‘he bears’
MOOD	bæ̅r-e	‘(that) he bore’

Figure 1 thus diagrams *bær* ‘he/she bore’ for the distribution of its inflectional categories across stem versus ending.

Figure 1: Distribution of morphological categories in Old English *bær* ‘he/she bore’
(cf. Anderson 1977)



This particular Old English example is quite dramatic, but the problem is by no means unique. As we will note below, reminiscent fusional word structures can be found in some Nilotic word structures.

2 Nilotic verbs: characteristic morphological categories and morphological typology

With this brief background to morphological typology, we now give a high-level introduction to Nilotic verbs. Much basic research remains to be done on this family, but there has been sufficient description of languages of all three major branches of the family to establish that Nilotic verb forms are far from isolating in terms of the index of synthesis. But this assertion is intriguing given the short word lengths quite typical of Western Nilotic languages, as in Dinka Rek in (1) above.

Typical morphological categories found across all Nilotic branches include:

- **ASPECT:** PERFECT(IVE)-IMPERFECTIVE contrasts are common. Particular languages may have additional aspect and modal contrasts. Though the word “tense” is found in some language-particular descriptions, the term is often misapplied and the family as a whole is more of an aspect-mood (AM) conflation type.
- **DIRECTIONAL:** A contrast is common between TOWARD (in the broader typological literature variously called “ventive”, “hither”, “cislocative”, “adlative”, “centripetal”) and AWAY (“itive”, “thither”, “translocative”, “ablativ”, “allative”, “centrifugal”) categories, which show orientation and/or movement relative to a deictic center. Some languages may include other directional or motion nuances in directional paradigms. In some languages, directionals have semantic extensions into the domain of aspect, pluractionality, and sometimes even number and valence (Dimmendaal 2003, Payne 2011, 2013). Given the potential for spatial-aspectual fusion, Nilotic languages can be said to demonstrate DAM (or even DAMN “directional-aspectual-mood-number”) conflation.⁴

⁴ That is, they are not TAM (tense-aspect-mood) conflating languages.

- **VALENCE & VOICE:** Morphological subcategories across Nilotic languages commonly include:
 - causative
 - applicatives (Dative, Instrumental)
 - middle(s)
 - impersonal passive
 - antipassive (some)
- **PERSON, NUMBER** of core arguments
- **NEGATIVE**
- **DISCOURSE-RELATED CONNECTIVES / TAXIS**

Despite having often similar morphological categories, how the categories are morpho-phonologically instantiated varies across Nilotic languages. The “flavor” of the instantiation is perhaps already suggested by (1) through (4) above: if Western Nilotic languages have short but also morphologically complex verbs, then they might be expected to have more fusional structures; while Eastern and Southern Nilotic might be expected to be more agglutinative. The following briefly sketches some of the morphological patterns noted in the descriptive literature.

Eastern Nilotic

Section 1 above noted Dimmendaal’s (1983: 96) presentation of Turkana verbs, with up to 13 position classes. If correct, then Turkana verbs appear to be highly synthetic and agglutinative. However, descriptions of some other Eastern Nilotic languages begin to raise questions about how agglutinative this branch of the family might be overall. Spagnolo’s (1933) description of Bari suggests that five morphemes can occur on a verb; but he notes “long” and “short” stems and various morphophonemic interactions. This immediately raises the possibility that the morphemes might not be so cleanly separable as one would expect in a prototypical agglutinative language. In section 3 below, we will take up the “problem” of the Eastern Nilotic verb.

Southern Nilotic

Southern Nilotic divides in two main sub-branches, “Kalenjin” versus “Omotik-Datooga” (Rottland 1982). Within the Kalenjin group, Endo may have eleven positions in the verb (Zwarts 2003: 114). For the related language Nandi, Creider & Creider (1989) give no schema of the sort in (4), but they emphasize the “extraordinary productivity” of the verbal system and give paradigms for verb roots or stems with perfective/imperfective aspect contrasts, directionals, ambulatives, mood, causative, associative, inceptive, etc. They further comment that: “Using ... suffixation ... it is possible to form eleven new types of verb... The suffixes can be readily combined with one another...” (Creider & Creider 1989: 85).

Though Nandi verb morphology is highly productive, examination of particular verb forms shows that the effects of ATR harmony, subsequent suffix vowel deletion, and tone processes have started to “pile up” the instantiation of morphological contrasts onto the stem in at least some formations (p. 86). Compare:

(6)	a. kɛ:-pál	b. pól-ú	c. kɛ:-pəl
	INF-dig	dig-VENT	INF-dig.VENT
	‘to dig’	‘he/she/the dig(s) up’	‘to dig up’

Significant descriptive work remains to be done on verb structures in Datooga varieties, but Kiessling (2007) notes about five suffix positions and mentions some co-occurrence complications, which may suggest potential challenges for a straightforward position class analysis.

Western Nilotic

There is internal diversity among the verb structures of the more than twenty Western Nilotic languages. The subfamily divides into two main branches. Within the Luo branch, Anywa verbs typically have from two to perhaps eight morphemes, but Reh (1996: 184) does not suggest a particular position class analysis. However, there are multiple means of effecting morphological contrasts besides concatenatively adding easily-separable morphemes, including consonant gemination; and ATR, vowel length, nasalization, and tone variation.

As already noted, some Western Nilotic verbs tend to have *monosyllabic* but *multimorphemic* stems. This is particularly characteristic of the Dinka-Nuer branch of Western Nilotic. In Thok Reel (Reid 2009) and Agar Dinka (Andersen 1995), the multimorphological structures are produced by up to three degrees of vowel length, vowel quality, voice quality (plain, creaky, breathy voice), consonant alternations, and tone variations. For some detail, Agar Dinka verbs divide into different tone/length classes. Andersen (1995: 44) notes that “the length of the stem vowel of a given verb form is the product of a particular configuration of vowel length at three morphological layers, *viz.* the root layer, the derivational layer, and the inflectional layer.” He describes Agar Dinka verbs as conjugating for at least “topic”, voice/valence, and directional categories. The data in (7) briefly demonstrate the morphophonology of some of the contrasts.

(7) Agar Dinka (̀ = creaky voice, .. = breathy voice)

- | | |
|--|---|
| <p>a. Simple
 ḍḍok à-m̀iit wẹ́ŋ
 boy DECL-pull cow
 ‘The boy is pulling the cow’</p> | <p>b. Centrifugal
 ḍḍok à-m̀iit wẹ́ŋ
 boy DECL-pull.CF cow
 ‘The boy is pulling the cow thither’</p> |
| <p>c. Centripetal
 ḍḍok à-m̀iit wẹ́ŋ
 boy DECL-pull.CP cow
 ‘The boy is pulling the cow hither’</p> | <p>d. Benefactive
 ḍḍok à-m̀iit wẹ́ŋ m̀oc
 boy DECL-pull.BEN cow man
 ‘The boy is pulling the cow for the man’</p> |
| <p>e. Benefactive-antipassive
 ḍḍok à-m̀iit m̀oc
 boy DECL-pull.BEN.APASS man
 ‘The boy is pulling for the man’</p> | <p>f. Antipassive
 ḍḍok à-m̀iit
 boy DECL-pull.APASS
 ‘The boy is pulling’</p> |

With this background, we now examine “the Maa verb”, or better, verb constructions.

3 The problem of Maa verbs

Maa (“Maasai”; Eastern Nilotic of Kenya and Tanzania) verbs may at first seem quite agglutinative. This is certainly suggested by analyses like the following where segmental parsing is fairly straightforward (tonal morphology and allomorphy of the Instrumental notwithstanding):

- | | |
|---|--|
| <p>(8) a. e-te-bel-ishór-iek-i
 3-PF-break-APASS-INST.PF-PASS
 ‘It was used to break (things).’</p> | <p>b. e-shuk-ún-or-é
 3-return.sth-TOWARD-MID-INST
 ‘He will use it to come back.’</p> |
|---|--|

However, Maa verbs defy clear position-class treatment. Tellingly, all published descriptions disagree and/or are vague on verb structure. The presentation in Tucker & Mpaayie’s (1955) grammar was perhaps the first indication of a problem, as they describe Maa verbs partly by presenting certain unanalyzed suffix chunks involving one, two, and three semantic categories per chunk. About thirty years later, Wallace (1981: 75) presented a verb-composition rule that contained seven to eight linear morpheme positions, plus four “if-then” statements. Subsequently Rasmussen (2002: 30) presented a

linear rule as a “significant” revision to Wallace’s analysis, accompanied by several “if-then” statements, plus he cautioned that the “type of base... may impose further restrictions”.

In my own years of field research, I have attempted to construct various analytical charts of verb forms, and can force a count of nine to ten potential linear positions, though they do not ever all co-occur in practice and there are multiple co-occurrence restrictions including some discontinuous elements. I have never been satisfied with any of the charts as definitive, and invariably augment them with arrows and if-then qualifications somewhat different from those of both Wallace and Rasmussen.

Figure 2 gives a first high level overview of the structure of the Maa verb. Some of the connectives are loose clitics while others are phonologically tightly integrated but all comprise a discourse “taxis/mood” paradigm and at most one connective occurs on any given verb. A prefix category that relates to participants (specifying person and number, imperative mood, or an “infinitive prefix” marking just number) precedes what I will call the “Stem”. “Stem” as I use it here may include categories often thought of as “inflectional” including aspect, some mood, and some number.

Figure 2: Maa verb structure: High-level overview

(Connective)=(Negative) – Participants – Stem

An important point to note is that the number of “nine to ten” constructed positions in my various work charts is markedly higher than the number of morphemes found in corpus-attested verbs, which we will look at in a moment. But counting morphemes is first of all synchronically fraught because what counts as a “separate morpheme” is caught up in the gradual process of historical change. For example, in the infinitive verb form *áa-idapash-atá* (INF.PL-scatter-PF.PL) ‘to have scattered, have spread out widely (by a plural subject)’, conjugational paradigms for aspect and person clearly show that the parsed-off pieces *áa-* and *-atá* are synchronically separable from *idapash*. But the remainder presents challenges for counting morphemes. The /i/ of *idapash* is historically separate from the adjectival root *dápásh* ‘broad, wide’, but the root/stem *idapash* ‘scatter, spread out widely’ is almost certainly lexicalized. Further, *dápásh* ‘broad, wide’ is probably related historically to *en=dáp* ‘palm, sole’ and to *a-idáp* ‘to plod’, but it is not clear that the /ash/ of *dápásh* is a synchronically separable piece. It is also not clear how much native speakers actively “parse” all these pieces (even if linguists can do so), nor how transparent the meaning connections are to modern speakers. Given such challenges, the following counts of morphemes in attested verb forms from a text corpus should be interpreted cautiously.

To help evaluate to what extent “nine to ten position classes” characterizes actual performance, a count was made of 854 distinct verb words (not tokens), discounting any strictly tonal morphology. For this count, all stem-initial “Class II” /i/ elements (Section 3.2) were considered to be part of the synchronic “root”. Figure 3 shows that the majority of verb forms contained four or fewer segmentable morphemes.

The initial connectives, NEGATIVE, and participant-related prefix morphemes (cf. Figure 2) are quite separable from the rest of the verb (i.e., they are agglutinative), and there was some question as to how much these elements might constitute the preponderance of morphemes behind the numbers in Figure 3. To investigate this, a second count of 866 distinct verbs (not tokens) was done of morphemes just within the “Stem” of Figure 2. The results in Figure 4 demonstrate that most stems contain just one or two non strictly-tonal morphemes.

The Figure 4 results start to give a rather different understanding of the typical Maa verb stem than might be inferred from average number of syllables per word in the translated Maa passage in (2) above, or from overall total morphemes including the left edge elements. The Figure 4 results arguably also start to make Maa seem a little less “strange” if viewed from the perspective of what is currently known about Western Nilotic verb types.

Figure 3: Maa segmental morphemes per verb word (ignoring strictly tonal morphs)

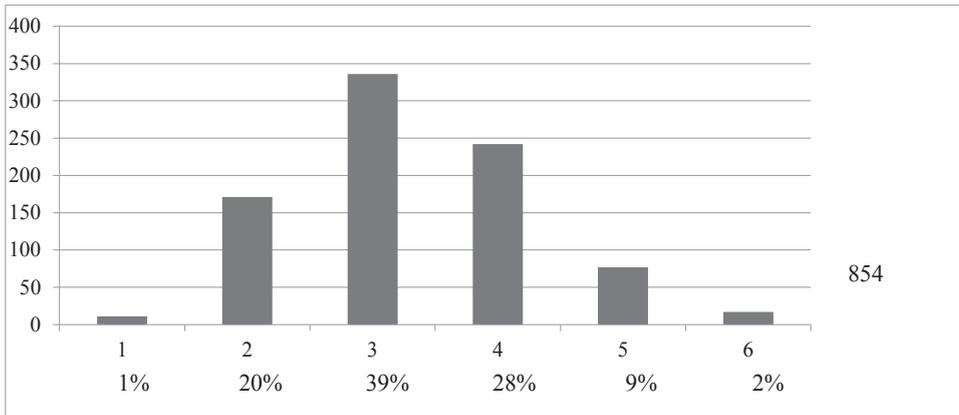
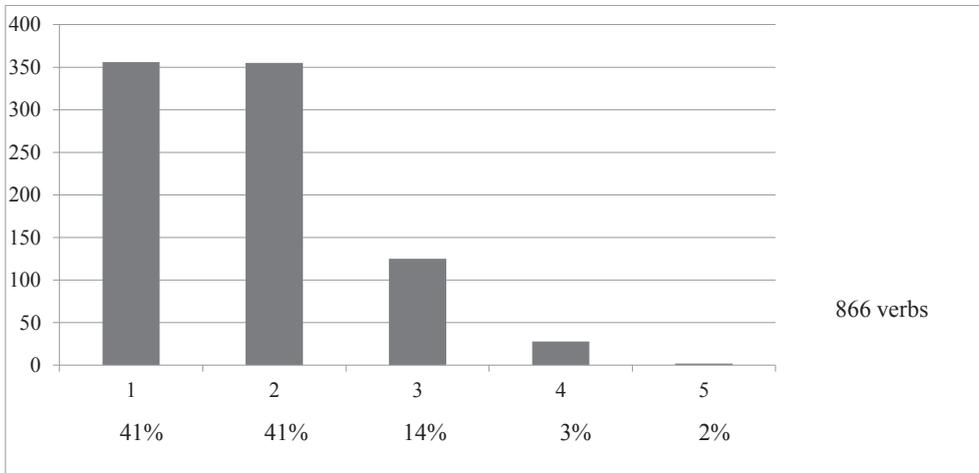


Figure 4: Maa segmental morphemes per verb stem (ignoring strictly tonal morphs)



To check the translated Maa material in (2) above against the profile of the non-translated corpus material, all the verb words in (2) above are parsed in (9). Aside from tonal morphology, the number of morphemes in the stem (bracketed) is indicated. Iterative action in a number of the forms is expressed by stem reduplication, resulting in a higher stem-morpheme count in the particular material in (2) than would occur otherwise.

- | | | | |
|-----|-------------------|---|------------------------|
| (9) | a-[jo-ki] | ‘I say to [you (plural)/them]’ | 2 |
| | em-i-[nyamal-iki] | ‘don’t you worry about it’ | 2 |
| | i-[nya-nya] | ‘you (plural) eat (iterative)’ | 2 (root reduplication) |
| | i-[ok-iok-o] | ‘you (plural) drink (iterative)’ | 3 (root reduplication) |
| | i-[nchop-ishop-o] | ‘you (plural) dress selves (iterative)’ | 3 (root reduplication) |
| | m-e-[lus-oo] | ‘it doesn’t surpass’ | 2 |
| | n-e-[lus-oo] | ‘and it surpasses’ | 2 |

But number of stem morphemes is not what makes the Maa verb so difficult to describe in position class terms. In what follows, I present seven or eight factors that conspire to make Maa verbs challenging to describe. Any one of these factors is not surprising. Taken together, they must be seen

as evidence of the human mind's tendency to co-store even analytically-separable categories (whether or not such may be shown by psycholinguistic measures as "co-lexicalized").

3.1 Tonal and partly-tonal morphology

Tonal morphology is hardly newsworthy in the African context, but it must be observed that tone has a very high functional load in Maa verbs. Here I give just a couple of examples (cf. Rasmussen 2002 for more extensive exemplification of Maa tone morphology). In (10a-b) aspect differences are marked by the tone patterns at the end of the verb plus the presence vs. absence of the prefix *tu-*. In (11), person/number differences are marked by tone and length together.

- | | | | |
|----------------------------|---|------------------------|--|
| (10) a. | áa-tu-duŋ-okiny-íé
3>1SG-PF-cut-DAT-INST.PF
'He has cut for me using it.' | b. | áa-duŋ-okiny-ie
3>1SG-cut-DAT-INST
'He cuts/will cut for me using it.' |
| (11) 1 initial mora | [áísis]
'I will praise him/them.' | 2 initial moras | [ááísis]
'I will praise you (SG).' |
| | [áísis]
'to praise' (by a SG agent) | | [ááísis]
'to praise' (by a PL agent) |
| | | | [ááísis]
'They will praise me.' |

3.2 Arbitrary stem classes

Southern and Eastern Nilotic languages make a distinction between what are called "Class I" versus "Class II" stems (or roots) (Tucker & Bryan 1966). In Maa, the classes largely correspond to verb stems that begin with an initial /i/ or /i/ (Class II), versus those that do not (Class I). However, there are a few irregular verbs that take the morphology of a given class without having this particular phonological shape. Both Class I and Class II roots/stems take the same morphological categories, but the coding elements can be distinct, and may occur in different regions of the verb.

The highlighted morphemes in (12–13) illustrate how Class I versus II verbs differ in marking PERFECT(IVE) and CAUSATIVE morphological categories.

- | | | | |
|------|--|----------|---|
| (12) | PERFECT(IVE) | | |
| | Class I | Class II | |
| a. | Á-tú-duŋ-o o=sínkólio.
1SG-PF-cut-PF MSG-song
'I interrupted the song.' | b. | Á-ísis-a nínché.
1SG-praise-PF them
'I praised them.' |
| (13) | CAUSATIVE | | |
| | Class I | Class II | |
| a. | K-é-ítá-shal enk=ají.
CN2-3-CAUSE-be.wet FSG-house
'It will make the house wet.' | b. | K-á-idúrr-ie in=kishú
CN2-1SG-migrate-CAUSE FSG-cattle
'I will make the cows migrate/move.' |

A (very) few roots are mixed in Class I/II properties. 'Run' takes a Class I CAUSATIVE but a Class II PERFECT(IVE).

- | | |
|--|---|
| <p>(14) Class I CAUSATIVE</p> <p>a. Á-ítá-kuet
1SG-CAUSE-run
'I make her run.'</p> | <p>Class II PERFECT(IVE)</p> <p>b. Á-kúét-a
1SG-run-PF
'I ran.'</p> |
|--|---|

The fact that a given morphological category does not occur in the same position across both verb classes makes it impossible to describe both classes with a single position class template. The fact that a given verb root like 'run' displays some properties of one class but other properties of another further shows speakers' intricate and detailed knowledge of co-occurrence restrictions.

3.3 Lexical aspect classes

Cross-cutting the Class I/II stem classes is a Stative versus Dynamic lexical class distinction – this lexical aspect issue is a fundamental element in understanding how Maa verbs “work”. This distinction is completely orthogonal to the leading /i/ ~ /u/ phonological signal. The distinction is also not a naive semantic one; ultimately, one must look at the morphological potential of different roots/stems. In Kenyan Maa dialects, most Stative versus Dynamic roots (or simple stems) differ in their ability to take directionals, PROGRESSIVE aspect, the INCHOATIVE, and certain nominalization patterns. Table 1 summarizes the morphological potential of the two types. Both may take applicatives (especially the INSTRUMENTAL applicative); but beyond this, the differing possibilities show that trying to force both into a single position class account would predict that many ungrammatical forms should instead be acceptable.

Table 1: Morphological diagnostics of stative versus dynamic verb roots/stems in Southern Kenyan Maa

	Stative Stems/Roots	Dynamic Stems/Roots
Applicatives (DATIVE, INSTRUMENTAL)	+	+
INCHOATIVE <i>-u(n)</i>	+	–
PROGRESSIVE <i>-ita</i>	–	+
Directionals		
<i>-aa</i> AWAY	(+)	+
<i>-o(n)</i> TOWARD	–	+
PERFECT(IVE)/SUBJUNCTIVE <i>(-tV)...-a</i>	'inchoative'	'perfect(ive)'
SITUATION NOMINALIZER <i>-an</i>	+	–
ACTION NOMINALIZER <i>-ata</i>	–	+

To further complicate matters, a few roots ('be bitter, be hard/difficult/strong', 'be generous') have morphological potential of both lexical aspect types. Whether a particular root is in one versus the other lexical aspect type can vary by sense of the root. In (15a) *gol* 'be hard' takes a nominalizer characteristic of Statives. Also indicative of the Stativity of *gol* is the inchoative meaning that results when the PERFECT(IVE) morphology occurs (15b). But in (16) (cited from Northern Samburu Maa), the same root takes PROGRESSIVE morphology which is not usually allowed for Stative verbs. The understanding of what is undergoing progressivity in (16) is not the situation of 'be strong, be hard', but rather the attempt. Overall, the morphosyntactic and semantic data indicate that in most of its occurrences *gol* is lexically Stative; but the tests in Table 1 disallow saying that *gol* conjugates strictly according to the pattern of a Stative verb. *Gol* can co-occur with the PROGRESSIVE, which simultaneously either coerces the sense of 'try to be strong' (16a), or (unusually) yields an inchoative sense (16b).

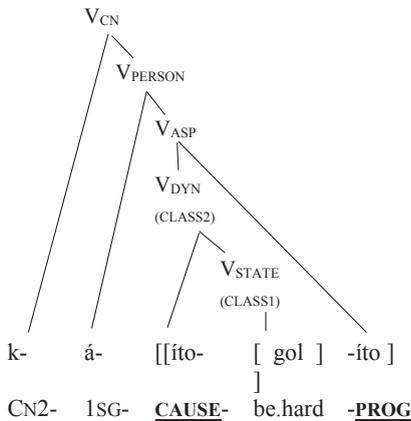
- (15) a. en=gol-ón
 FSG-be.hard-NMLZ.SITUATION
 ‘strength, force’
- b. ε-ta-gol-ó
 TEMP.3-PF-be.hard-PF
 ‘when it becomes difficult’
- (16) a. K-ó-gol-ító.
 CN2-3-be.hard-PROG
 ‘He is trying to be strong.’ (eg. when he has been bereaved). (SN)
- b. e-gol-ító
 3-be.hard-PROG
 ‘He’s becoming stronger.’ (*‘It’s being hard.’)

Another complication is that some roots or stems take only a subset of the normal Stative morphological potential. In both Southern Kenyan Maa and Ilchamus Maa (cf. König 1993) *ishu* ‘be alive’ and *ata* ‘have’ cannot take PROGRESSIVE suggesting they are Stative; but they also cannot take the INCHOATIVE suffix, nor the PERFECT(IVE) with either inchoative or perfect(ive) meaning. Such roots or stems must thus have different word-construction templates from the dominant Stative and Dynamic ones.

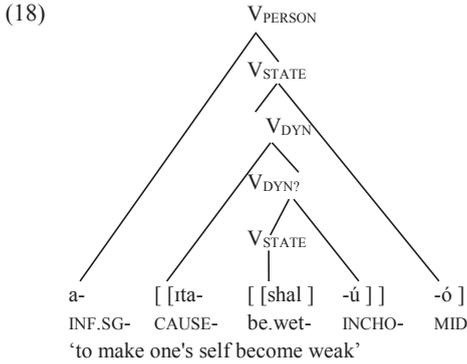
3.4 Stem category-changing morphology

The Stative vs. Dynamic and Class I vs. Class II distinctions give us an initial typology of Maa verb roots (or simple stems). But as might be expected, some morphological categories can derive a given root (or stem) into an alternative stem class. MIDDLES and perhaps the ANTIPASSIVE *-ishɔ(r)* create Stative stems. CAUSATIVES (17, 18) create potentially dynamic stems. In (17) progressivity clearly pertains to the action of ‘making’ or ‘causing’ inherent in the prefix *ito-*, whether or not it may also pertain to the root concept of ‘be hard’ or ‘hardening’.

- (17) K-á-ító-gol-ító.
 CN2-1SG-CAUSE-be.hard-PROG
 ‘I am making it hard/hardening it.’ (S)



In (18), *shal* ‘be wet’ is lexically Stative and as such cannot take a MIDDLE suffix. However, once the INCHOATIVE *-u(n)* and/or the CAUSATIVE occur, a MIDDLE can then occur to re-derive a new Stative stem.



Separately from the Stative/Dynamic lexical aspect issues, the Class I CAUSATIVE morpheme *ita-/ito-* and the somewhat idiosyncratic prefix *ɪ-/i-* change an otherwise Class I root/stem to a Class II stem.

In sum, MIDDLES, perhaps the ANTIPASSIVE *-ɪshɔ(r)*, CAUSATIVES, and the *ɪ-/i-* prefix throw the new stem into specific sets of further morphological possibilities, potentially different from those of the basic root. The result, of course, is that it is impossible to say that a particular verb root takes just one particular word-formation template. The fundamental conclusion here is that any position-class approach which ignores stem-level units is wholly inadequate for Maa.

3.5 Functional category vs. formal instantiation of category

We have already seen in passing that a single, functionally-defined, grammatical category can have markedly differential formal instantiation, depending on stem category in Maa. At least four functionally-defined categories have such behavior in Maa. Here we look more explicitly at these categories and their implications for understanding “the Maa verb”.

A semantically-defined INCHOATIVE category can occur on both (most) Class I and Class II Stative stems. It does not co-occur with Dynamic stems. The INCHOATIVE category is coded in three different ways depending on grammatical aspect.

- If the grammatical aspect is NON-PERFECT(IVE), the strictly +ATR suffix *-u(n)* is used, as in (18). The suffix *-u(n)* is a dedicated INCHOATIVE without other salient senses. Synchronically it contrasts in form, meaning, and distribution with the -ATR TOWARDS suffix *-o(n)/-u(n)* which only occurs on certain Dynamic stems (Table 1). As the two suffixes occur on different lexical aspect stem types, it is arguable whether one can say they are, or are not, part of the same linear paradigm; but the TOWARDS suffix is in the Directional paradigm, which can then be followed by the MIDDLE category; while the MIDDLE can be followed by the *-u(n)* INCEPTIVE.
- With Class I Stative verbs, the discontinuous combination *tV'...-a(k)* indicates INCHOATIVE PERFECT(IVE) aspect (19), and/or INCHOATIVE SUBJUNCTIVE mood (15b). The suffix portion *-a(k)* alone occurs with Class II Stative verbs with the same semantic effect (20).

(19) ɛ-tɔ-bórr-ó en=kíteŋ
 3-PF-be.calm-PF FSG=cow.NOM
 ‘The cow has calmed down/has become calm.’ (W)

(20) í-mbórr-a! (root *ɪbɔrr* ‘be white’)
 2-be.white- PF
 ‘You've become white!’ (e.g. from sickness)

- (22) \acute{A} -ítá-bá-y[†]á εn=kítéŋ n-a-íróbí.
 1SG-CAUSE-reach-AWAY FSG=cow REL.F-F.SG.REL-be.cold
 ‘I will drive the cow to Nairobi.’ (lit: that which is cold)
- (23) \acute{A} -ítá-bá-yie εn=kítéŋ n-a-íróbí.
 1SG-CAUSE-reach-AWAY.PF FSG=cow REL.F-F.SG.REL-be.cold
 ‘I drove the cow to Nairobi.’
- (24) ε-te-beł-áy-ie
 3-PF-break-AWAY-PF
 ‘She broke them (one at a time).’

My point here is not so much what is the “correct” parse of the preceding verb forms. Rather, it is to observe the creeping fusion of once-independent forms in the Maa verb stem which at times makes such analytical parsing decisions seem almost misguided. They are also misguided in that speakers produce and understand the meaning of a whole complex construction (including the tonal patterns) without necessarily assigning one piece of form to one separable atom of meaning.

3.6 Affix co-occurrence conditions

Preceding discussion has highlighted co-occurrence restrictions between stem types and certain affixes (see especially Table 1). For example, Dynamic stems can generally take:

- [1] DATIVE and INSTRUMENTAL applicatives (either, or both together, but in different “slots” of the verb)
- [2] The TOWARD or AWAY Directionals (but not both simultaneously)
- [3] PROGRESSIVE grammatical aspect

But there also are co-occurrence restrictions among affixes separately from stem issues so far discussed. As expected on semantic grounds, the PROGRESSIVE does not co-occur with PERFECT(IVE) nor with the ANTIPASSIVE affix which typically carries “habitual” semantics. These co-occurrence restrictions are not surprising and reveal aspectual incompatibilities.

What is semantically more surprising is that, in major subdialects, a stem carrying the PROGRESSIVE suffix [3] cannot simultaneously take an applicative [1] and/or a Directional [2] affix, and vice versa (though a Directional or the DATIVE Applicative can co-occur with the INSTRUMENTAL Applicative). Thus, the PROGRESSIVE is in paradigmatic contrast with whole complexes of other possible morpheme combinations (see Figure 5 below). Why this is the case, is not yet clear (though the AWAY Directional does have semantic extension into pluractionality, so there may in this instance be some aspectual conflict).

As just noted, the AWAY and TOWARD Directionals and the DATIVE Applicative (which has a semantic directional sense in most of its uses) are in complementary distribution; hence, they would appear to belong to a single paradigm. But they belong to (or create) different stem categories for subsequent PERFECT(IVE) conjugation; that is, there are co-occurrence restrictions between certain semantically directional suffixes and PERFECT(IVE) forms:

- (tV-)...-ie(k)+tone with AWAY (also ANTIPASSIVE and INSTRUMENTAL) stems
- (tV-)...-a(k) with TOWARD and DATIVE (also most other) stems

Such semantically non-transparent co-occurrence restrictions lead to what we might call morpheme complexes, where two or more isolatable pieces must be learned together, becoming co-grammaticalized or co-lexicalized.

3.7 Frozen morpheme complexes or categories: true fusion

A further stage in the scale of fusion are historically old morpheme complexes, where a combination of grammatical or functional category “A” with category “B” does not yield a simple compilation of A+B morphological forms, but a phonologically distinct form C, or perhaps an A+C combination. This is what we find in Maa for Directionals+MIDDLE, Directionals+Aspects, and PERFECT(IVE)+PLURAL combinations.

Forms whose glosses are joined by “+” in Figures 5 and 6 below constitute such frozen forms. Here we talk through just some of the frozen combinations involving Directionals. In (25a) are listed the basic ATR-determined allomorphs of the TOWARD Directional and NON-PERFECT(IVE) MIDDLE. Whether the final consonants /n/ and /r/ occur depends on what particular suffixes might or might not follow them. If TOWARD and NON-PERFECT(IVE) MIDDLE morphological categories co-occur in a verb, the formal instantiation is not something like *-ona(r) or *-ua(r), but is instead -oɲɛ/-oɲie (depending on ATR issues), as in (25b).

- (25) a. -o(n), -u(n) TOWARD
 -a(r), -o(r) NON-PERFECT(IVE) MIDDLE
 b. -onyɛ, -onyie TOWARD+NON-PERFECT(IVE) MIDDLE

In (26a) we see the basic ATR-determined allomorphs of the TOWARD Directional and the PERFECT(IVE) MIDDLE. If TOWARD and PERFECT(IVE) MIDDLE morphological categories co-occur in a verb, the formal instantiation is not something like *-onyɛ, *-unnie, or *-ue, but is instead -uno, as in (26b). (The (26b) form might reflect some older combination of TOWARD with the NON-MIDDLE PERFECT(IVE) basic form -a(k)/-o(k), but this needs further research.)

- (26) a. -o(n), u(n) TOWARD
 -ɛ, -e PERFECT(IVE) MIDDLE
 b. -uno TOWARD+PERFECT(IVE) MIDDLE

The data in (27) show similar frozen morphological combinations involving AWAY. (27a) gives one of the basic synchronic allomorphs of AWAY in a non-middle context. Historically, this possibly developed from a form like /-ara/ (Tucker & Mpaayei 1955, Payne 2013), which accounts for the /r/ element in (27b-c). Again, in (27b) we do not find what might be the synchronically-expected combinations *-ara (from something like -aa(r) AWAY plus -a(r) NON-PERFECTIVE MIDDLE); nor in (27c) do we find *-arɛ (from something like -a(r) AWAY plus -ɛ PERFECT(IVE) MIDDLE).

- (27) a. -áa AWAY
 b. -ari AWAY+ NON-PERFECT(IVE) MIDDLE
 c. -ara AWAY+ PERFECT(IVE) MIDDLE

Instead, the actual phonology of the “frozen” combinations in (27b-c) is reminiscent of what we see in (25b) and (26b): a non-low front vowel occurs for the NON-PERFECT(IVE) AWAY/TOWARD MIDDLE, and a back vowel occurs for the PERFECT(IVE) AWAY/TOWARD MIDDLE. This is just the opposite of what is found in the simple middles not involving directionals, which have a non-low front vowel -ɛ for the PERFECT(IVE) MIDDLE and a back vowel -a(r) for the NON-PERFECT(IVE) MIDDLE. All in all, such frozen complexes, and their seeming synchronic idiosyncracies, give us complete sympathy for Tucker & Mpaayei’s resistance to parsing off many of the morphemes.

4 Conclusion: “The Maa verb”

This paper has hopefully made that case that there is no one structure, template, or constructional pattern that fits all possible verbs of Maa. A single definitive position class analysis is impossible due to the kinds of interacting factors discussed in 3.1 through 3.7. Though many dimensions of Maa verbs do not seem so fusional and individual morphemes with identifiable semantics can often be isolated

using traditional analytical methods, the various co-occurrence restrictions and multiple form-to-meaning relationships indicate that something more like a Word-and-Paradigm framework (Matthews 1972, Anderson 1977) could provide a more accurate technical account of possible word forms. (This says nothing about mental representations that underlie how native speakers comprehend or productively create word forms.)

Nevertheless, to help advance our linguistic understanding of Maa verbs, I offer Figures 5 and 6 for verbs built around Class I Dynamic and Stative verb roots, respectively. These figures do not adequately represent tonal dimensions of morphology. In Figure 5, any MIDDLE element (and possibly the ANTIPASSIVE) derive a Stative stem (which then subsequently belongs to Figure 6). The Directionals and DATIVE applicative are somewhat marginal-to-unacceptable with Statives. Where AWAY does occur with a Stative, it shifts its sense to a plural or plurality-of-situation reading (Payne 2013). The MIDDLE and ANTIPASSIVE forms do not occur with roots that are Stative, unless some other affix has been added that, at an intermediate level, derives a Dynamic Stem such that MIDDLE or ANTIPASSIVE is subsequently felicitous. The MIDDLE and ANTIPASSIVE are, however, included in Figure 6 to show their relative position if they do occur.

In Figures 5 and 6, a heavy black box delineates what appears to be a historically older Core. (The term “Stem” as used in this paper includes this Core plus elements in the final columns of Figures 5 and 6). In addition to a root from the Core, one and only one of the items in bold from columns 1, 2, or 3 is required.

Color is used in Figures 5 and 6 to help communicate the regions of the verb that grammatically or lexically code broad functional domains. Elements that affect argument structure are in red type. Green shading marks Directional or directionally-related morphology. Yellow marks aspectual elements (both lexical and grammatical aspect). Blue marks mood or subjunctive-related morphology. Gray marks number. The coloring reveals that the Core is heavily an aspect domain, while mood (and taxis) are dominantly outside the Core (disregarding tonal morphology).

Figure 5: Class I Dynamic verbs (BOLD = REQUIRED ELEMENT)

1	2	3	4 (Class I)	5	6	7	8	9	
			Direction-Argument-Structure-Aspect- Number (NPF is usually the unmarked case)						
Overlay: 2PL+NPF: Final Mora Reduplication									
INFINITIVE SG a- INFINITIVE PL áa-		CAUSE itV- Creates Class II dynamic Stem		PROG -ita		IMPERS/ PASSIVE/ PLURAL -i			
CN		ARG PERS+ (NUM)							
wha peé (tm-); Disallows aspect & m- Subjn	NEG m-(+ tone)				PF -re(k) (tone; #+/PL-^)		PL+PF -0)ie after <i>AIIAI</i>		
'chain' n-; Disallows m- Subjn					PF SUBIN -a(k)		PL+PF -ie after <i>PF</i>		
'if' ten-					INST/ CAUSATIVE -ie (+/E PF tone; #+/PL-^) (Causative interpretation only for Class II)			IMPERS/ SUBJUNCTIVE -i (tone)	
umil ɔ-; Requires m- Subjn	SUBIN m- (+ tone)								
CN k- (not with subjunctive or imperative)									
		DYNAMIC VERB ROOT		REDUPLICATED ROOT 'iterative'		MID+NPF -a(t)/-o(t) MID+NPF+PL -a(a)/-o(o) MID+PF -e 'became [state]' MID+PF+PL -a(e)/-o(e) 'became [state]'			
		SG IMP Class I IV- (tone) SG IMP Class II i- PL IMP En-		PF IV- SUBIN IV- (tone)		MID+INST -a(é) (NPF) MID+INST -a(e) (PF) Only if transitive stem. AFASS -isho Only if transitive stem.			
		PF+SUB + MID -u(n)-ie AV AS + MID -a(i) Only if transitive stem. AFASS -isho Only if transitive stem.							

- Dynamic roots/stems can, in general, take Progressive, Directionals, Applicatives (Dative, Instrumental).

Figure 6. Class I stative verb (BOLD = REQUIRED ELEMENT)

1	2	3	4 Class I	5	6	7	8	9	10
CN when peé Disallows aspect & m- Subjn 'chain' n- Disallows m- Subjn 'if' ten- until → Requires m- Subjn		INFINITIVE SG a- INFINITIVE PL a-	CAUSE nV- Creates Class If stem. With most roots, creates Dynamic Stem. i- Creates defective Class If stem from some Adjectives PF N-; SUBJN V- (tone)	STATIVE ROOT	AWAY -ia (+/- shift to plurality of situation) BELOWABLE -o(0): very marginal with stative(s) EXHAUSTE -ak(r) (marginal) If Dative occurs, then Middle easily occurs.	MID +NPF+at/ [state] NPF +PF -e 'became' Only some Statives, and only if transitive. MID +NPF+INST -aré (tone) MID +PF+INST -are (tone)	INCHOATIVE -at(r) Only after Stative root, Middle, or Antipassive (not Dative)	INST -ie(k) PF +PF (tone)	IMPERSONAL PASSIVE / PLURAL -i
Overlay: 2PL+NPF: Final Mora Reduplication Direction -Argument Structure: Aspect-Number (NPF is usually the unmarked case)									
			CAUSE -ic(k) (tone) PL +PF -ite PF SUBJN -at(k) (yields INCHOATIVE sense) PL +PF -ia, -ote					CAUSE -ie(k) (Causative interpret only for Class II)	

- Directionals and Dative are often odd with Stative verbs, yielding meanings like 'It is soft to her' (from 'be soft'); but some are acceptable as in 'I will agree with her/I will have love for her' (from 'like'), or 'I am angry towards her' (from 'be angry')
- If Dative occurs, Inchoative does not occur unless there is also a Middle. This suggests that the Dative is preferably restricted to occurring on Dynamic stems; and the Inchoative (re-)creates Dynamic stems. The Inchoative is blocked from occurring on already-dynamic stems.

Abbreviations

APASS	antipassive	MOV.K	“moveable k”
ASP	aspect	NEG	negative
BEN	benefactive	NMLZ	nominalizer
CF	centrifugal	NOM	nominative
CN	connective	NPF	non-perfect(ive)
CP	centripetal	NUM	number
DAM	direction-aspect-mood conflation	PASS	(impersonal) passive
DAT	dative	PERS	person
DECL	declarative	PF	perfect(ive)
DYN	dynamic	PL	plural
EP	epenthetic	PROG	progressive
F	feminine	PSD	possessed
IMP	imperative	REL	relativizer
INCHO	inchoative	SG	singular
INF	infinitive	SUBJN	subjunctive
INST	instrument	TEMP	temporal mode
M	masculine	VENT	ventive
MID	middle		

References

- Andersen, Torben. 1995. Morphological stratification in Dinka: On the alternations of voice quality, vowel length and tone in the morphology of transitive verbal roots in a monosyllabic language. *Studies in African Linguistics* 23: 1–63.
- Anderson, Stephen R. 1977. On the formal description of inflection. *Chicago Linguistics Society* 13: 15–44.
- Comrie, Bernard. 1989. *Language Universals and Linguistic Typology* (2nd edition). Chicago: University of Chicago Press.
- Creider, Chet & Jane Tapsubei Creider. 1989. *A Grammar of Nandi. (Nilo-Saharan Linguistic Analyses and Documentation, 4.)* Hamburg: Buske.
- Dimmendaal, Gerrit. 1983. *The Turkana Language*. Dordrecht: Foris.
- Dimmendaal, Gerrit. 2003. Locatives as core constituents. In Erin Shay & Uwe Seibert (eds.), *Motion, Direction and Location in Languages*, pp. 91–109. Amsterdam: John Benjamins.
- Greenberg, Joseph H. 1960. A quantitative approach to the morphological typology of language. *International Journal of American Linguistics* 26: 178–94.
- Householder, Fred Jr. 1947. A descriptive analysis of Latin declension. *Word* 3: 48–58.
- König, Christa. 1993. *Aspekt im Maa*. Cologne: University of Cologne, Institute of African languages and linguistics.
- Kiessling, Roland. 2007. Space and reference in Datooga verbal morphosyntax. In Doris L. Payne & Mechthild Reh (eds.), *Advances in Nilo-Saharan Linguistics. Proceedings of the 8th Nilo-Saharan Linguistics Colloquium. Hamburg, August 22–25, 2001*, pp. 123–142. Cologne: Köppe.
- Matthews, Peter H. 1972. *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.
- Payne, Doris L. 2011. Nilotic directionals in lexical and syntactic context. Paper presented at the 44th Societas Linguistica Europaea, September 8–11, 2011, Universidad de La Rioja, Logroño, Spain.

- Payne, Doris L. 2013. The challenge of Maa ‘Away’. In Tim Thornes, Erik Andvik, Gwendolyn Hyslop & Joana Jansen (eds.), *Functional-Historical Approaches to Explanation: In Honor of Scott DeLancey* (Typological Studies in Language series, 103), pp. 259–282. Amsterdam/Philadelphia: Benjamins.
- Rasmussen, Kent. 2002. Verb tone in Ilkeekonyokie Maa. MA thesis, University of Oregon.
- Reh, Mechthild. 1996. *Anywa Language: Description and Internal Reconstructions*. (Nilo-Saharan, 11) Cologne: Köppe.
- Reid, Tatiana. 2009. Aspects of phonetics, phonology and morphophonology of Thok Reel. Research dissertation, University of Edinburgh.
- Rottland, Franz. 1982. *Die Südnilotischen Sprachen: Beschreibung, Vergleichung und Rekonstruktion*. (Kölner Beiträge zur Afrikanistik, 7.) Berlin: Reimer.
- Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace and Co.
- Schleicher, August. 1861. *Compendium der vergleichenden Grammatik der Indogermanischen Sprachen. Kurzer Abriss einer Laut- und Formenlehre der indogermanischen Ursprache, des Altindischen, Altiranischen, Altgriechischen, Altitalischen, Altkeltischen, Altslawischen, Litauischen und Altdeutschen*. Weimar: Böhlau.
- Spagnolo, Fr. L. M. 1933. *Bari Grammar*. Verona: Missioni Africane.
- Tucker, Archibald N. & Margaret A. Bryan. 1966. *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press.
- Tucker, Archibald N. & J. Tompo Ole Mpaayei. 1955. *Maasai Grammar, with Vocabulary*. London: Longman, Greens & Co.
- Wallace, B. F. 1981. The morphophonemics of the Maasai verb. In Thilo Schadeberg & Lionel Bender (eds.), *Nilo-Saharan: Proceedings of the First Nilo-Saharan Linguistics Colloquium, Leiden, September 8–10, 1980*, pp. 75–88. Dordrecht: Foris.
- Zwarts, Joost. 2003. *The Phonology of Endo: A Southern Nilotic Language of Kenya*. Utrecht Institute of Linguistics OTS.

Old English data:

<http://www.verbix.com/webverbix/go.php?T1=beran&imageField.x=7&imageField.y=10&D1=23&H1=123> (last access 20.4.2015)

MIXED PIVOT CONSTRAINTS IN TOPOSA CLAUSE CHAINING

Helga Schröder

Clause chaining characterises a combination of one or more finite and non-finite clauses that indicate the sequential order of events and marks foreground information in texts. The non-finite clauses are dependent in tense, aspect, and mood (TAM) to the finite clause. If clauses are linked in such a way, as part of the definition of “clause chaining”, the chained clauses typically represent the foreground information of a text (Dooley 2010: 3). Clause chaining languages are most often found in Papua New Guinea, the Americas, and in Africa in the SOV languages of Ethiopia. This paper deals with the morphosyntactic properties of clause chaining in Toposa, a verb-initial Eastern Nilotic language of South Sudan. The paper demonstrates that Toposa does not follow a switch-reference system for coreferencing arguments in the sentence structure but employs a mixed S/A-S/O pivot constraint. It also discusses how the continuity of referents in clause chains can be skipped in complex sentence structures.

1 Introduction

Clause chaining describes a combination of one or more clauses that show tense, aspect, and mood (TAM) dependency of the non-finite clause on the finite clause and it marks the sequential order of events and typically presents foreground information in texts (see also Dooley 2010: 3). Clause chaining is found widely in Papua New Guinean languages (Elson 1964), in Australian languages (Austin 1979) and North American Indian languages (Longacre 1985, 1990).

Clause chaining often occurs in SOV languages in areas like Papua New Guinea, South America, Korea and Japan among others (Longacre & Hwang 2012: 94). In Africa it is found as SOV variant in Ethiopia (Völlmin et al. 2007). Although it is stated in (Dooley 2010: 8) that clause chaining rarely occurs in SVO languages, Dooley reports that two SVO languages in Africa exhibit signs of clause chaining. Hopper (1979: 213–215) talks about a clause-combining effect in Kiswahili as representing the foreground information of narrative texts.¹ Longacre found clause chaining in Anuak, a Western Nilotic language from South Sudan (1990: 88–90, 2007: 418). I have argued recently that Toposa, a VSO Eastern Nilotic language in South Sudan organises its discourse through clause chaining (H. Schröder 2011).

Clause chaining is often combined with switch-reference (Huang 2000, 2003, Payne 1997, Stirling 1993: 14–18). Switch-reference describes a morphological marking system that indicates on the verb whether conjoined clauses have the same or a different subject or other core arguments. Switch-reference can be found in coordination and subordination relationships, depending on the language.

This paper will demonstrate the clause chaining features of Toposa. Though Toposa does not employ the typical device of switch-reference, conreferential NPs in clause chaining work on a mixed S/A-S/O pivot.

Firstly, I shall look at the morphosyntactic features of Toposa clause chains (section 2), then I shall consider the features of switch-reference that are typical for most clause-chaining languages (section 3). Next I shall deal with the S/A and S/O pivot in Toposa clause chaining (section 4). Finally I shall demonstrate the existence of clause-skipping in evaluating coreference within a chain (section 5), and section 6 has some concluding remarks.

¹ Hopper does not call this feature clause chaining but describes it as a typical clause chaining property (1979: 214).

2 Morphosyntactic properties of clause chaining in Toposa

As agreed upon by many scholars, clause chaining is characterised by non-finite clauses that show inflectional dependencies on the finite clause (Dooley 2010: 3, Payne 1997: 312, Longacre 1990: 11, Myhill & Hibiya 1988: 363). The finite clause stands on its own, carries all the inflectional features like tense, aspect and mood (TAM) in its finite verb and is often referred to as the main clause in a coordinative-subordinative sentence construction. I would suggest labelling the finite clause as the ‘controlling clause’, because it controls the TAM dependencies.

2.1 The status of the non-finite clause

The status of the non-finite clause has been discussed from various angles. One question that scholars are concerned with is whether the non-finite clause is coordinative or subordinative in nature relative to the finite clause. Some authors point out that the non-finite clause is like a coordinative clause (Roberts 1988: 4, 1997: 183, also Haspelmath 1995: 12–17, 2007: 46f) in at least some respects.² Stirling (1993: 195–198) discusses how in Papua New Guinean languages these clauses have been termed coordinative, citing Haiman (1980 and 1983), Comrie (1983), and MacDonald (1983), among others. However, as will also be shown for Toposa, the non-finite clauses show morphosyntactic dependency on the controlling clause in that the chained clauses pick up their tense/aspect/(mood) interpretation from the finite clause. As this morphosyntactic property of the non-finite clause exhibits dependency, some scholars call these clauses “quasi-coordinate” (Haiman & Munro 1983: xii, Stirling 1993: 15). Van Valin & LaPolla label the TAM dependency of the non-finite clause on the controlling clause(s) as operator dependence. For them the clauses with operator dependence represent a hybrid between coordination and subordination and they call this clause linkage “cosubordination” (1997: 455). They argue that semantically it has coordinative effects because the non-finite clause is assertive, but morphosyntactically it is dependent. I shall adopt Van Valin & LaPolla’s view and regard the non-finite clauses as hybrids between coordinative and subordinative clauses. Thus I shall refer to these clauses as *chained clauses* and to the finite clause as the *controlling clause*.³

In the following section I shall deal with the morphosyntactic properties of clause chaining in Toposa, i.e. the operator dependency of tense/aspect, and the participant orientation of the chain.

2.2 Morphological marking, tense operator and post-nuclear orientation

In Toposa, a typical clause chain begins with a controlling clause that is inflected for tense and aspect, whereas the subsequent chained clauses carry the dependency markers *to-/ki-* which signal tense-aspect dependency on the finite verb of the controlling clause. It is also possible that the chain begins with a temporal adverbial that sets the time frame in the controlling clause, as in the following example (taken from M. Schröder 2010: 46):⁴

- (1) Bee koloŋo nuwani, to-loto Kwee ɲa-ki-rap
 it-is-said long.ago very DEP-go jackal INF-DER-search

 ɲa-kee-moogwa, to-ryam-u ɲa-ate ka ɲi-tooni, ...
 F/PL-his-food DEP-find-ALL F/SG-cow of D/SG-person

‘It is said that long ago, Jackal went to search for his food, he found a cow of someone, ...’

² Haspelmath (1995: 12–17, 2007: 46f) gives five tests for identifying coordination: INTERCALATION: if two clauses are coordinate clauses, neither can occur intercalated in the other; TEMPORAL ICONICITY: if two clauses are coordinate clauses and narrate events or states, their linear order must agree with the chronological order of the events or states; CATAPHORIC REFERENCE: if two clauses are coordinate clauses, the first cannot contain a reference whose “antecedent” only occurs in the second; ARGUMENT FOCUS: if two clauses are coordinate, neither can occur as argument focus; EXTRACTION: if two clauses are coordinate clauses, an interrogative expression cannot be “extracted” from the second to occur in sentence-initial position before the first (Dooley 2010: 8). Dooley states that dependent clauses fail these tests and are therefore called quasi-coordinative (ibid. 9).

³ Another term often found in the literature for the chained clause is medial clause (Haiman 1987).

⁴ Note that the examples from this text collection do not show tone. Where tone marking was necessary to indicate case and/or tense, as in examples (4, 5, 6b) and (13), this has been provided. Toposa also distinguishes between \pm ATR vowel sets which have not been differentiated for the purposes of this article. Underlined vowels at the end of words indicate voiceless vowels.

The above sentence represents a typical beginning of an animal fable. The story is set with the formula *bee* ‘it is said’, and the adverbial *koloŋo nuwanj* sets the time frame as in the past. Certain following clauses are chained to the initial clause with the dependency marker *to-* in *toloŋo* ‘he went’ and in *toryamŋ* ‘he found’. Both verbs lack the typical inflection that marks person, tense, and aspect on finite verbs. Examples (2a-b) below demonstrate how the dependency marking would differ in first person singular and plural in personal narratives:

- (2) a. ... a-to-loto
 1P-DEP-go
 ‘I go/went’
- b. ... a-to-lot-o
 1P-DEP-go-PL
 ‘we go/went’

As the first person is marked with an overt prefix preceding the dependency marker of the chained verb, the assumption is that even DEP verb forms have a slot for marking the (coreferential) person on the verb. The zero prefix in the second and third person singular and plural can be regarded as a gap in the paradigm. The following shows the realisation of the coreferential pronouns in the verb of clause-chained clauses:

Table 1: Overview of co-referential integrated pronouns in Toposa

<i>PERS</i>	<i>CO-REF</i>	<i>DEP</i>	<i>root</i>	<i>PL</i>
1	a-	to-	lot	
2	∅	to-	lot	
3	∅	to-	lot	
1	a-	to-	lot	-o
2	∅	to-	lot	-o
3	∅	to-	lot	-o

Example (2a-b) are not finite verbs, they depend on a finite clause. A clause from a first person narrative shall demonstrate this fact:

- (3) Koloŋo ki-riŋ-a dirj isuwa ki-ra ŋi-soroko,
 long ago DEP-bestill-PL truly we.EX DEP-be M/PL-young.men
 a-to-lom-a ŋa-kopo na a-bee-i Nakamoŋo.
 1P-DEP-enter-PL F/SG-place which 3P-call-IMP Nakamoŋo
 ‘Long ago [when] we were still young men, we settled in a place called Nakamoŋo.’

In order to compare dependent verbs with fully inflected ones, consider the following examples (taken from Schröder 2008: 51):

- (4) a. É-múj-i nǎ-kírîŋ.⁵
 1P-eat-IMP F/SG-meat
 ‘I am eating meat.’
- b. Ì-múj-i nǎ-kírîŋ.
 3P-eat-IMP F/SG-meat
 ‘He is eating meat.’
- c. È-múj-î nǎ-kírîŋ.
 1P-eat-IMP F/SG-meat
 ‘I was eating meat.’
- d. È-mùj-í nǎ-kírîŋ.
 3P-eat-IMP F/SG-meat
 ‘He was eating meat.’

As these data show, the normal finite Toposa verb is marked for tense, aspect, person, and number. The tense system is the past and non-past type. Tense in Toposa is marked by the tone pattern that extends over the entire verb and varies according to verb class, person, number, and tense category. In addition to the tone pattern, a tense prefix *a-* occurs in the third person singular and plural in the past tense. In example (4c) and (4d), this tense prefix *a-* has become fused with the person agreement prefix *i-*, resulting in *e-* (for a more detailed description of the fusion of the past marker in TO and KI-class verbs see Schröder 2008: 53). Note how the tone pattern on the verb for first-person singular changes from HHL in (2a) to LHF in (2c) to mark the change from non-past to past. Similarly, the third-person changes from LHL in (2b) in non-past to LLH in (2d) in the past.

Additionally, Toposa has two aspects: imperfective and perfective. Imperfective aspect is indicated by the suffix *-i*, as shown in the above data.⁶ The perfective aspect is indicated by the suffix *-itî*.⁷

- (5) É-múj-itî nǎ-kírîŋ.
 1P-eat-PER F/SG-meat
 ‘I ate meat.’

As these data show, the two verbs *toloto* ‘they went’ and *toryamutû* ‘they found’ in example (1) do not have the typical tense-aspect inflection of the finite Toposa verb, nor do they employ the typical person agreement prefixes as shown in example (2a-d); they represent non-finite forms. The following examples demonstrate the difference between the inflected and dependent verb forms in Toposa for a KI-class verb (a-d) and a TO-class verb (e-h).

- (6) a. è-mùj-í ‘he was eating (inflected form)’
 3P-eat-IMP
- b. kí-múj ‘he was eating (dependent form)’
 DEP-eat
- c. è-mùj-é-té ‘they were eating (inflected form)’
 3P-eat-IMP-PL

⁵ Note that the personal pronoun is usually not expressed as a free word in Toposa but is integrated as an argument on the verb (Schröder 2008: 111).

⁶ The imperfect marker has an allomorph *-e* before the plural suffix *-te* which is used in second and third person plural. First person plural uses the imperfect suffix *-i* and the plural suffix *-o*.

⁷ Most finite verb forms in Toposa are marked for either imperfective or perfective. Verbs in passive, reflexive and stative verbs are not marked for imperfective and perfective.

- d. kǐ-múj-à ‘they were eating (dependent form)’
DEP-eat-PL
- e. á-lòs-í ‘he was going (inflected form)’
3P-go-IMP
- f. tó-lót ‘he went (dependent form)’
DEP-go
- g. á-lòs-é-té ‘they were going (dependent form)’
3P-go-IMP-PL
- h. tó-lót-ò ‘they went (dependent form)’
DEP-go-PL

The non-finite forms are not some form of infinitive, either. This can be seen from constructions like the following:

- (7) ... to-lotǒ Kwee ǵa-ki-rap ǵa-kee-moogwa,⁸ ...
DEP-go jackal INF-DER-search F/PL-his-food
‘... Jackal went to search for his food, ...’

The verb *ǵakirap* ‘to search’ represents the typical infinitive form, which consists of the feminine singular marker *ǵa-* and the derivation prefix *ki-*.⁹

Example (1) and example (7) also show that the placement of the dependent clause of the chain is post-nuclear, i.e. that the controlling clause precedes the chained clause. The post-nuclear orientation of the chained clauses, as Toposa displays it, is rare (Dooley 2010: 6 and Payne 1997: 321), mainly because most clause chaining languages are SOV, whereas Toposa is a verb-initial language.

3 Typical features of switch-reference

One phenomenon that cross-linguistically occurs in many clause-combining processes of coordination and subordination is switch-reference. In this clause combining type the coordination type is most often the clause chaining and the subordination type is of an adverbial nature. In some languages switch-reference may also relate to other clause types like relative and complement clauses or coordinative clauses that are not clause-chained (Stirling 1993: 15). Switch-reference has formal and functional properties, which I shall briefly outline now.

3.1 Formal properties of switch-reference

Switch-reference systems most typically have the following conditions (Stirling 1993: 6):

1. Switch-reference combines at least two clauses (“locality conditions”), but when combined with clause chaining it usually links multiple clauses (Stirling 1993: 18–23).
2. Switch-reference is typically formally marked on the verb as verb inflection (“realisation conditions”) but can also take other formal properties such as being an independent pronoun or it is marked on the pivot NP (Stirling 1993: 29–30).
3. Switch-reference indicates both syntactic and semantic dependency on the controlling clause. Either the dependent clause is subordinated to the main clause, or the dependent and the controlling clause are in a clause chaining relationship (“dependency conditions”).

⁸ Note that *ǵamoogwa* ‘food’ is plural here in the sense of ‘different bits here and there’.

⁹ In TO-class verbs this derivation prefix drops out when a verbal extension like allative, ablative, or benefactive is added.

4. Switch-reference refers to the subject, i.e. it indicates whether the same subject (SS) or a different subject (DS) occurs between the controlling and the dependent clause (s) (“subject conditions”).

Concerning condition four, Stirling points out that also non-subjects in some languages can be switch-referenced (1993: 25–28).

The following examples (8a-b) taken from Stirling (1993: 14) show the use of switch-reference in an adverbial clause¹⁰ in Mojave, a North American Indian language, where the first clause is subordinated to the second, the main clause.

- (8) a. Nya-isvar-**k** iimar-k.
when-sing-SS dance-T
‘When he sang, he danced.’
- b. Nya-isvar-**m** iimar-k.
When-sing-DS dance-T
‘When he_i sang, he danced_j.’ (Mojave, Munro 1980: 145 (4))

In (8a) the switch-reference markers *-k* ‘SS’ indicates that in both clauses the third person is the same subject. In (8b) *-m* ‘DS’ demonstrates that the third person is referred to is a different person. Other examples, also taken from Stirling (ibid.) show switch-reference of the clause chaining type in Usan, a Papuan language:

- (9) a. Ye nam su-**ab** isomei.
I tree cut-SS I:went:town
‘I cut the tree and went to town.’
- b. Ye nam su-**ine** isorei.
I tree cut-DS it:went:down
‘I cut the tree and it fell down.’ (Usan, Haiman & Munro 1983: xi, (3–4))

In (8a-b) the first person subject is marked as *-ab* ‘SS’ and *-ine* ‘DS’. Note also how in (8a-b) the first adverbial clause is dependent on the second independent clause; example (9a-b) exhibit a controlling clause and a chained clause in a clause chaining combination, they indicate sequential actions.

The examples (8a-b) and (9a-b) also demonstrate the four formal conditions of switch-reference listed above: They represent the combination of at least two clauses each (condition 1), the switch-reference is marked by a suffix on the verb (condition 2) and they demonstrate the subject condition and dependence condition respectively (conditions 3 and 4).

An example that shows switch-reference in the context of a clause-chaining language with more than two clauses is the following sentence construction from Amele, a Papuan language (Stirling 1993: 15, citing Roberts 1987: 101):

- (10) Ija Malolo uqa na ka jic ana-g na
1SG Malolo 3SG of:POSS car road mother-3SG:POSS at

omo nu sum-ud-i bibil-igin
there for wait-3SG-PRED SIM-be:DUR-1SG:DS

ne-**ce**-b tobo-**co**-min belo-w-an.
come:down-DS-3SG climb:up-DS-1SG go-1PL:DUR-YESTP
‘While I waited for Malolo’s car there at the main road, he came down, I climbed in, we two went off.’

¹⁰ Adverbial clauses serve an adverbial function, i.e. they either modify a verb phrase or they modify an independent clause. They are often understood as adjuncts, i.e. they are not arguments of a clause. Adverbial clauses add some information to the independent clause, and they are always optional (Payne 2006: 297).

The above clause chain consists of one semantically adverbial clause, two dependently expressed sequential actions and the controlling clause at the end. The switch-reference is marked by DS suffixes on the verb: *-igin* (1 SG), *-ce* (3SG) and *-co* (1SG). The tense operator *-an* ‘yesterday past’ occurs in the controlling clause, which is the final clause of the chain.

3.2 Functional properties of switch-reference

Switch-reference has two functions: Firstly, it demonstrates one way of tracking referents in discourse, and in doing so over a long stretch of discourse, it fulfills its second function: indicating the syntactic pivot function (Van Valin & LaPolla 1997: 287–288, also Stirling 1993: 7).

A pivot concerns whether participant reference in clause linkage works on a nominative-accusative or an ergative-absolutive basis. In ergative-absolutive systems, the conjoining NPs in clause-linkage strictly follow an S/O pattern. In this case the intransitive S is a co-referent of the O argument of the second clause. In the nominative-accusative system they follow an S/A pattern. In this case the S argument of the intransitive clause conjoins with the A argument of the transitive clause (see Dixon 1997: 154).¹¹ In the Usan example (9a) above the clause-chaining works on a S/A pivot because the combination expresses coreference between the A of the transitive clause and the S of the intransitive clause. In (9b) the clause chain represents a S/O pivot as it expresses coreference between the O of the transitive clause, and the S (intransitive argument) of the intransitive clause.¹²

4 The co-referential pivot in Toposa clause chaining

Even though switch-reference often accompanies clause-chaining languages as stated by Roberts (1988: 62), Stirling (1993: 14), Dooley (2010: 5), Huang (2000: 289–290), and Longacre & Hwang (2012: 94), Toposa does not employ switch reference for tracking referents in discourse.¹³ In this section I shall look at how reference tracking in Toposa clause chaining works instead.

The tracking of referents in Toposa clause-chains works on a pivot constraint. Pivot constraints operate in the following way: If languages are constrained by an S/A pivot,¹⁴ the coreferential NPs are always in S or A function. If a common NP is in O function, the clause has to be passivised to satisfy this pivot. Languages, however, that work on a S/O pivot constrain the occurrence of coreferential NPs to their S and O. If a common NP is in A function, antipassive is used to ensure that that constraint is satisfied.

The coreferentiality in Toposa rests on two major conditions:

1. Formal condition: The coreferent is marked on the verb as an integrated subject pronoun either by an overt morpheme for first person singular and plural, or by a zero morpheme for second and third person singular and plural see table 1.
2. Functional condition: The interpretation of argument coreferentiality between clauses works on a mixed S/A-S/O pivot.

The formal and functional conditions of Toposa coreferentiality will be demonstrated in the next section.

¹¹ Stirling states that the S/A pivot in switch-reference systems is the most common one in contrast to the S/O pivot. The S/A pivot is preferred over the S/O pivot even in languages that have some features of ergativity (1993: 6–7). This observation is confirmed by Dixon (1997: 153). I however know of three Central Sudanic languages spoken in Eastern Congo (Omi, Logoti and Lugbara), where the switch-reference system operates on a S/O pivot, Wright (in preparation).

¹² Stirling (1993: 42) mentions that switch-reference can have various other functions, besides occurring in temporally sequential and adverbial clauses. Two of these are that switch-reference can encode distinctions in temporal relations as well as in logical and epistemological relations.

¹³ Dooley remarks that switch-reference is not confined to expressing sequential actions (2010: 8). He refers to two Quechuan languages (citing Cole 1983: 2,4) and to Huichol (citing Comrie 1983:19), and to American and Australian languages where switch-reference occurs more often in association with subordinate adverbial clauses than with sequential actions (citing Roberts 1988: 62).

¹⁴ When talking about S/A or S/O pivots, S stands for the subject of the intransitive clause, A for the subject of the transitive clause and O for the object of the transitive clause.

4.1 The S/A pivot

In this section I shall demonstrate the S/A pivot relationship between coreferential arguments. Example (10) has been taken from a narrative text (M. Schröder 2010: 6):¹⁵

- (11) A-bu Nakuju, ø-to¹⁶-limoki jikapiti nitikawosoni nibe Napurukucu,
 3P-came God, 3P-DEP-told bird very.clever who.called Napurukucu
 tem, "To-woyiu¹⁷ jawuno, kotere ki-yooliyorotori nituja kopo."
 DEP:said IV-twist rope in.order.to DEP-take people down
 'God came, he told a very clever bird whose name was called Napurukucu (= Orange Starling),
 he said, Twist a rope in order to take people down.'

The following sentences provide an example from an expository-descriptive text (M. Schröder 2010: 6, 113–114):

- (12) S1 Seke na ecamitere nakitasapana nituja, isyawunete mono
 So when is.wanted to.initiate people, begin DM
 jikaŋaka, ki-ryama, to-tukwo jatemari itemokino
 leaders DEP-meet DEP-discuss that is.appropriate
 njide lu eriŋa nesapana, itasapanio.
 children who are.not.yet not.initiated be.initiated
 S2 Ku-wudakisi nai nituja lu ecamito jasapana waapei,
 DEP-gather then people who want initiation one.place
 jikilyoko ka njide luuciki, ta-nyama jikora kode nemono.
 men and children small DEP-eat rams or ox
 S3 Narumworete ŋuna, to-loto nai to-pero naperiti.
 end this DEP-go then DEP-sleep in.sleeping.ground

S1 'When it is wanted (= when the leaders want) to initiate people, the generation-set leaders begin, they meet, they discuss that it is alright to initiate the children (= filial generation) who have not yet been initiated.' S2 'The people who want to be initiated gather in one place, the men and the children [of the new set], they eat goats or an ox.' S3 '[After] the end of that they then go to sleep in the [separate] sleeping-ground.'

These examples show the zero subject marking of the coreferential arguments via the *tV-* or *kV-* prefix on the verb (thereby fulfilling the formal condition) in *øto-limoki* 'he said' in example (8). The subject marking precedes dependency marking of the chained verb.

Referring back to examples (8) and (9), the functional condition of the S/A pivot is also applied in these examples. The sentence in example (11) combines the S of the intransitive controlling clause *Abu Nakuju* 'God came' with a coreferential A of the first chained clause *tolimoki jikapiti* 'he told the bird' with another coreferential A of the second chained clause *tem* 'he said'. In example (12) the sentence S1 has an S in the controlling clause in *isyawunete mono jikaŋaka* 'the leaders begin' and combines it with the co-referential S of the intransitive clause and the coreferential A of a transitive clause *kiryama*,¹⁸ *totukwo jatemari itemokino* 'they meet, they discuss that it is appropriate ...' The clauses in S2/S3 of example (12) combine a S with a co-referential A, with a co-referential S and co-

¹⁵ From here on examples are only partially segmented to show dependency marking strategies, i.e. dependency markers and person agreement markers on verbs (with the exception of example (13), which has been fully segmented to demonstrate case marking in passive constructions.

¹⁶ The zero prefix for person will from now on not be marked, but its morphological presence will be assumed, see the explanation between examples (2) and (3).

¹⁷ Note that the prefix *ki-* in this example is an imperative marker that has the same form as the dependence marker.

¹⁸ Note that in Toposa the verb *kiryama* 'meet' is intransitive: *kiryama ka jikoku* 'he met with [the] child'.

referential S in *Kuwudakisi nai ηituŋa ... , tanyama ηikorae kode jemoŋo, ... , toloto nai, topero naperiti* ‘the people gather, ..., they eat goats or an ox, ..., they go and they sleep in the sleeping-ground’.

4.2 The S/O pivot

The S/O pivot applies when the O of a transitive clause has to be coreferential with a subject argument in another clause. In such a case only the S of the intransitive clause or the O of the transitive clause occurs. This S/O pivot linkage is partly due to clauses that have passive constructions. In Toposa impersonal constructions (as also in most Southern Nilotic languages) where the single lexical argument takes the accusative tone marking, just like an object of a transitive clause would. Consider the following:

- (13) a. È-kèr-é-tè ηá-átùk.
3P-run-IMP-PL F/PL-cow/NOM
‘The cows are running.’
- b. Ì-dés-i ηé-kilé ηá-átùk.
3P-beat-IMP M/SG-man/NOM F/PL-cows/ACC
‘The man is beating the cows.’
- c. Ì-dés-it-àè ηá-átùk.¹⁹
3P-beat-IMP-PAS F/PL-cows/ACC
‘The cows are being beaten or somebody beats the cows.’²⁰

The impersonal construction in (13c) does not show the tone marking for nominative, which would be HLL for *ηáátùk* ‘cows’ as in (13a), rather, it bears the accusative tone marking HHF *ηáátùk* as in (10b).²¹ In other words, in Toposa the single lexical argument of the impersonal sentence is morphologically marked in the same way as the object of the transitive clause (see also H. Schröder 2008: 58–59). Scholars have debated whether the impersonal construction in some Eastern Nilotic languages (Toposa, Turkana and Maasai,) and Southern Nilotic languages (Tugen, Kipsigis, Nandi etc.) is a real passive construction, or an impersonal third person construction (H. Schröder 2012). In the latter case the construction of (13c) would best be rendered as ‘somebody beats the cows’. Irrespective of how the construction is understood, the single lexical argument of the impersonal construction has accusative case marking and requires either an S or another O if conjoined, see the following example taken from M. Schröder (2010: 6):

- (14) ... ki-yooliwun-oe nai ηituŋa, ki-bitibitiuni kopo, ηaberu
DEP-let-down-PAS so people DEP-let.themselves down, women

ka ηide tya ηikecekilyoko.
and children and husband-theirs

‘The people were let down, they let themselves down, the women and children and their husbands.’

In (14) the O of the clause *ηituŋa* ‘people’ is coreferential with the S of the intransitive reflexive clause *kibitibitiuni* ‘they let themselves down’.

The next example demonstrates the S/O pivot in a very interesting way (taken from M. Schröder 2010: 8):

¹⁹ Note that in (13b) and (13c) the personal pronoun marker *i-* is the same for the third person singular and plural. Regarding passive: Toposa has an impersonal construction marked by passive suffixes *-o ~ -ae ~ -oe*.

²⁰ The difference between the person agreement third person plural prefix *e-* of examples (13a) and the *i-* of (13c) is due to different verb classes. The *i-* occurs with KI-class verbs and the *e-* with TO-class verbs.

²¹ Note that the accusative tone pattern of *ηáátùk* ‘cows’ is HHF only before pause and HHL (*ηáátùk*) elsewhere.

- (15) S1 To-limokisi_i nai Dimoru talokaatekece, temasi,
 DEP-tell then Stones brothers.theirs DEP:say
 ‘Kapesi ta-anya njibareng kana wiye.’
 IV-go DEP-look cattle in cattle-camps

- S2 Ki-det-ae ka njalita, ki-reje-e, temar-ae ...
 DEP-beat-PAS with sticks DEP-chase-PAS DEP:say-PAS

S1 ‘Those Stones [generation-set] said to their brothers, “Go and look after the cattle in the cattle-camps!”’ S2 ‘They were beaten with sticks, and were chased, and it was said ...’

In (15) the common O *talokaatekece* ‘their brothers’ is coreferential with another O in S2 *Kidetae ka njalita* ‘they were beaten with sticks’, and another O in another impersonal clause *kirejee* ‘they were chased’. What is intriguing in this construction is that the coreferential argument ‘they’ in S2 does not create any syntactic ambiguity. If the S/A pivot would be applied it would be not clear whether the subject *Dimoru* ‘the stones’ or the object *talokaatekece* ‘their brothers’, introduced in S1 as A and O, are coreferenced in S2 as the ones that get beaten. As has been established in this paper a referent in object position in an impersonal construction needs either another O or S as areferent because of the S/O pivot. So it is clear that the ‘they’ of clause S2 refers to the O of S1 *talokaatekece* ‘their brothers’ and not to the A of S1 *Dimoru* ‘the stones’.

The S/O pivot also must apply when the O of a sentence construction is coreferenced in a non-impersonal construction. In such a case, only the S of an intransitive clause or the O of another transitive clause can follow the initial clause to satisfy the S/O pivot, as the following example shows (taken from M. Schröder 2010: 6):

- (16) To-woyiu nai Napurukucu japatu natikaanikani, to-woi²² loowoi.
 DEP-twisted so Napurukucu tendon-string which.strong DEP-long very
 ‘So Napurukucu twisted a strong tendon-string, it was very long.’

The S/O pivot occurs in example (16) where the object of the transitive clause *japatu natikaanikani* ‘a very strong tendon-string’ is the common argument between the two clauses; and it is coreferential with the S of the descriptive clause *towoi loowoi* ‘it was very long’.

This S/O pivot, which is not based on a impersonal construction, is also found in the following example (taken from M. Schröder 2010: 15):

- (17) ... ku-wara Lokolinjro ka njikeetuŋa jadokari lobooto, to-twonikini²³
 DEP-look Lokolinjro and people.his climbing to.camp DEP-difficult

In example (17) the object *nyadokari lobooto* ‘his climbing to the camp’ is coreferential with the S of the descriptive clause *totwonikini* ‘it was difficult’.²⁴

Dependency marking also occurs in simple coordination that does not indicate the sequential order of events and is not related to foreground information. In such cases the strict locality conditions apply as outlined for switch reference above under 3.1. The sentence construction consists of a controlling clause and one or two (at most) chained clauses:

- (18) A-bu to-osiki njakimar sementiks, ta-lakari njakilo.
 3P-came DEP-give.up reading semantics, DEP-happy more.than
 ‘He gave up studying semantics and felt much happier.’

²² Note that *towoi* ‘it was long’ is a stative verb.

²³ Note that the verb *totwonikini* is also a stative verb.

²⁴ Note that in a S/A pivot as in English the second conjoined clause would have to change into a relative clause like ‘Lokolingiro and his people looked [where] to climb to the camp, [but] which was difficult’. Or the subject would have to be repeated; the clause could not just be conjoined as in ‘Lokolingiro and his people looked [where] to climb to the camp, [but] this climbing was difficult.’

- (19) E-baritj̄ j̄nekilye, to-yar-ite.
 3P-rich man DEP-live-SIM
 ‘The man is rich and (furthermore) successful.’

Note that in case of the logical succession of events in (18) the onset *abu* ‘he came’²⁵ has to occur. In example (19) the suffix *-ite* indicates simultaneity. These two examples show a simple coordinated clause construction where participant reference works on a S/A pivot, in (18) A/S are co-joined and in (19) S/S.

5 Clause-skipping

One irregular phenomenon that can occur in clause chaining languages is that in evaluating coreferentiality within the chain, a clause may be embedded, introducing a new referential argument, and the coreference evaluation just skips the argument of the embedded clause. This is commonly referred to as clause-skipping for languages with switch reference systems (Stirling 1993: 18–20).

Consider the following example from Toposa (taken from M. Schröder 2010: 84):

- (20) To-upu j̄naberu, to-loma j̄nakugworo, to-dejuni, ki-boyikini, to-yai
 DEP-tire woman DEP-enter INF:cry DEP-exhaust DEP-sit.down DEP-be
 ca j̄nibore kaku keŋe, ta-tama, ki-yaŋaki, tem ...
 DM thing back her DEP-think DEP-pant DEP:say
 ‘The woman got tired, she started to cry, she became exhausted, she sat down, the thing
 (referring to an evil spirit) was on her back, she thought, she panted, she said ...’

The long chain of eight clauses relies referentially on the clause with the referent *j̄naberu* ‘woman’. The chain is interrupted by *toyai ca j̄nibore kaku keŋe* ‘the thing was on her back’. Here a new argument is introduced: *j̄nibore* ‘thing’. However the referential chain skips consideration of this referent and in the following clause *tatama* ‘she thought’, co-reference is determined relative to the original referent *j̄naberu* ‘woman’.

Such a referent-skipping can be explained with the foreground and background distinction that is commonly found in discourse. In this regard the clause *toyai ca j̄nibore kaku keŋe* ‘the thing was on her back’ adds background information, e.g. explanatory information, to the foreground information, which is captured in the clause-chained sentences.²⁶ See another example where the referent of the embedded clause is skipped (M. Schröder 2010: 48):

- (21) Ani e-jeketa j̄nakile ka jaate, ta-ratarata Kwee j̄nebu,²⁷
 When 3P-become.good milk of cow DEP-cheat jackal hyena
 to-lepuuni ca jaate, to-ŋoba j̄nakile, ani i-doji j̄negoototo,
 DEP-milk DM cow DEP-drink.up milk when 3P-remain foam
 to-lemu j̄nacoto, ki-yatakinea ...
 DEP-take urine DEP-add
 ‘When the milk of the cow had become good, Jackal cheated [intensive] Hyena, while he continually milked the cow, he drank up the milk, when [only] foam remained, he took urine, he added [that], ...’

In this chain, which has seven clauses, the controlling referent is *Kwee* ‘jackal’ and all the S/A arguments of the chained clauses refer to him. But there is a clause in the chain *ani idoji j̄negoototo*

²⁵ The full function of *abu* ‘he came’ occurring in texts or in single co-ordination is still under investigation.

²⁶ How the clause-chaining device in Toposa discourse is related to the foreground and background distinction was shown in H. Schröder (2013).

²⁷ Here the clause has no co-referential constraint because the second clause introduces two new participants, ‘Jackal’ and ‘Hyena’.

‘when the foam remained’, where the only argument is the NP *jegoototo* ‘foam’. However the NP ‘foam’ does not interrupt the coreference chain, as the referent of this NP is skipped in determining the referent of the following chained clause, which is interpreted as coreferential with *Kwee* ‘jackal’ in *tolemu ηacoto, kiyatakinea, ina Nyebu, tem* ‘he took urine, he added [that], he gave to hyena, he said ...’ From a foreground-background distinction the clause *ani idoji jegoototo* ‘the foam remained’ can be regarded as expressing background information.

6 Conclusion

In this paper I have demonstrated that Toposa conjoined clauses exhibit clause chaining properties. A typical Toposa clause linkage shows a tense/aspect dependency of the non-finite clause to the controlling clause. The order of the dependent clauses in the chain, in line with the nature of a verb initial language structure, is postnuclear because the controlling clause precedes the chained clauses. I also argued that Toposa clause chains do not have a switch-reference system for conjoined NPs but display a mixed S/A-S/O pivot system. The coreferential argument is represented by a zero morpheme on the verb when referring to third person singular and plural. The S/A pivot is evident when the A of a transitive clause combines with another A or a S of the intransitive clause. However, the reference system often switches to a S/O pivot system. In such a case an O is coreferential with either the S of an intransitive or another O of a transitive clause. I also illustrated that coreferentiality evaluation in clause chains can skip a referent of a background clause.

Abbreviations

1P/2P/3P	first/second/third person	INF	infinitive
1PL/3PL/1SG	first/third person plural, first person singular	IV	imperative
		M/SG M/PL	masculine singular/plural
ACC	accusative	NOM	nominative
ALL	allative	PAS	passive
D/SG	diminutive singular	PER	perfective
DEP	dependence marker	PL	plural
DER	derivational affix	POSS	possessive
DM	discourse marker	PST	past
DS	different subject	SIM	simultaneous
DUR	durative	SS	same subject
F/SG F/PL	feminine singular/plural	T	tense
IMP	imperfect	YESTP	yesterdays past
INT	intensive		

References

- Austin, Peter. 1979. Switch-reference in Australian languages. In Pamela Munro (ed.), *Studies of Switch Reference* (UCLA Papers in Syntax 8), pp. 1–54. Los Angeles: University of Southern California.
- Cole, Peter. 1983. Switch reference in two Quechua languages. In John Haiman & Pamela Munro (eds.), *Switch-Reference and Universal Grammar*, pp. 1–15. Amsterdam/Philadelphia: Benjamins.
- Comrie, Bernhard. 1983. Switch-reference in Huichol. In John Haiman & Pamela Munro (eds.), *Switch-Reference and Universal Grammar*, pp. 17–37. Amsterdam/Philadelphia: Benjamins.
- Dixon, Robert M.W. 1997. *Ergativity*. Cambridge: Cambridge University Press.
- Dooley, Robert A. 2010. Exploring clause chaining. *SIL Electronic working papers* 2010-001, 1–20.
- Elson, Benjamin (ed.). 1964. *Verb Studies in Five New Guinea Languages*. Norman: University of Oklahoma, Summer Institute of Linguistics.

- Givón, Talmy (ed.). 1979. *Syntax and Semantics: Discourse and Syntax*. New York: Academic Press.
- Haiman, John. 1980. *Hua: A Language of the Eastern New Guinea Highlands*. Amsterdam/Philadelphia: Benjamins.
- Haiman, John. 1983. On some origins of switch-reference marking. In John Haiman & Pamela Munro (eds.), *Switch-Reference and Universal Grammar*, pp. 105–128. Amsterdam/Philadelphia: Benjamins.
- Haiman, John. 1987. On some origins of medial verb morphology in Papuan languages. *Studies in Language* 11.2: 347–364.
- Haiman, John & Pamela Munro (eds.). 1983. *Switch-Reference and Universal Grammar*. Amsterdam/Philadelphia: Benjamins.
- Haiman, John & Sandra A. Thompson (eds.). 1988 *Clause Combining in Grammar and Discourse*. Amsterdam/Philadelphia: Benjamins.
- Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in Cross-Linguistic Perspective*, pp. 1–55. Berlin: de Gruyter.
- Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Language Typology and Syntactic Description* (Complex constructions, II), pp. 1–51. Cambridge: Cambridge University Press.
- Haspelmath, Martin & Ekkehard König (eds.). 1995. *Converbs in Cross-Linguistic Perspective*. Berlin: de Gruyter.
- Hopper, Paul J. 1979. Aspect and foregrounding in discourse. In Talmy Givón (ed.), *Syntax and Semantics* (Discourse and syntax, 12), pp. 213–241. New York: Academic Press.
- Huang, Yan. 2000. *Anaphora: A Cross-Linguistic Study*. Oxford: Oxford University Press.
- Huang, Yan. 2003. Switch-reference in Amele and logophoric verbal suffix in Gokana: a generalized neo-Gricean pragmatic analysis. *Reading Working Papers in Linguistics* 7: 53–76.
- Longacre, Robert E. 1985. Sentences as combinations of clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description* (Complex constructions, 2), pp. 235–298. Cambridge: Cambridge University Press.
- Longacre, Robert E. (ed.). 1990. Storyline concerns and word-order typology in East and West Africa. *Studies in African Linguistics* (Supplement 10), Los Angeles: The James S. Coleman African Studies Center and Berkeley Department of Linguistics. Berkeley: University of California.
- Longacre, Robert E. 2007. Sentences as combination of clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description* (Complex constructions, 2), pp. 372–420. Cambridge: Cambridge University Press.
- Longacre, Robert E. & Shin Ja Hwang. 2012. *Holistic Discourse Analysis*. Dallas: SIL International Publications.
- MacDonald, Lorna. 1983. Tauya medial verbs. *Language and Linguistics in Melanesia* 14.1–2: 113–137.
- Munro, Pamela. 1980. On the syntactic status of switch-reference clauses: The special case of Mojave comitatives. In Pamela Munro (ed.), *Studies of Switch Reference* (UCLA Papers in Syntax, 8), pp. 1–54. Los Angeles: University of California.
- Munro, Pamela (ed.). 1980. *Studies of Switch Reference* (UCLA Papers in Syntax, 8). Los Angeles: University of Southern California.
- Munro, Pamela. 1983. When ‘same’ is not ‘not different’. In John Haiman & Pamela Munro (eds.), *Switch-Reference and Universal Grammar*, pp. 223–243. Amsterdam/Philadelphia: Benjamins.
- Myhill, John & Junko Hibiya. 1988. The discourse function of clause-chaining. In John Haiman & Sandra A. Thompson (eds.), *Clause Combining in Grammar and Discourse*, pp. 361–398. Amsterdam/Philadelphia: Benjamins.
- Payne, Thomas. E. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press.

- Payne, Thomas. E. 2006. *Exploring Language Structure: A Student's Guide*. Cambridge: Cambridge University Press.
- Roberts, John R. 1987. *Amele* (Descriptive Grammar Series). London: Croom Helm.
- Roberts, John R. 1988. Amele switch-reference and the theory of grammar. *Linguistic Inquiry* 19: 45–64.
- Roberts, John R. 1997. Switch-reference in Papua New Guinea: a preliminary survey. In Andrew Pawley (ed.), *Papers in Papuan Linguistics* 3, pp. 101–241. Canberra: Research School of Pacific and Asian Studies, Australian National University.
- Schröder, Helga. 2008. *Word Order in Toposa: An Aspect of Multiple Feature-Checking*. SIL International and the University of Texas at Arlington Publications in Linguistics 142.
- Schröder, Helga. 2011. Toposa, a clause-chaining language. Paper presented at the workshop *Challenges in Nilotic languages – phonology, morphology and syntax from synchronic and diachronic perspective*. Kyoto University (Kyoto), Japan. 2–4 December 2011.
- Schröder, Helga. 2012. The passive-enigma in Nilotic. Paper presented at the *Workshop on Passive in Southern and (other) Nilotic languages*. 5 December, 2012. University of Nairobi.
- Schröder, Helga. 2013. Clause chaining in Toposa, a pragmatic approach. *Lodz Papers in Pragmatics* 9: 25–44.
- Schröder, Martin C. 2010. *Toposa Traditional Texts*. Nairobi: SIL International. Manuscript.
- Shopen, Timothy (ed.) 2007. *Language Typology and Syntactic Description* (Complex constructions, 2). Cambridge: Cambridge University Press.
- Stirling, Lesley. 1993. *Switch-Reference and Discourse Representation*. Cambridge: Cambridge University Press.
- Van Valin, Robert D. Jr. & Randy J. LaPolla. 1997. *Syntax, Meaning and Function*. Cambridge: Cambridge University Press.
- Völlmin, Sascha, Azeb Amha, Christian Rapold & Silvia Zaugg-Coretti (eds.). 2010. *Converbs, Medial Verbs, Clause Chaining and Related Issues* (Frankfurt African Studies Bulletin, 19). Cologne: Köppe.

tone and tongue root [TR] as ventive morphemes in Endo-Marakwet¹

Chelimo Andrew Kiprop

1 Introduction

Marakwet is a Southern Nilotic Kalenjin language that is spoken in Kenya in the North Eastern parts of Elgeyo Marakwet County in the larger Rift Valley Province. The county has four districts within it, namely: Keiyo South, Keiyo, Marakwet West and Marakwet East. The latter two form the “Marakwet Territory”.

The term “Marakwet” is broadly used to group together all the linguistic varieties found within the Marakwet Administrative District. The historian Kipkorir (2009) argues against such grouping by the colonial regimes and records the different varieties as: Endo, Marakweta, Almo, Kiptani, Borokout and Sengwer/Cheranganyi. Kipchumba (2010), in his studies of Marakwet literature, also records that “Marakwet” is a pool of several linguistic entities or groups that must be treated as separate. He categorises these groups as:

- i. those living towards West Pokot,
- ii. those around central Marakwet, and
- iii. those neighbouring the Keiyo to the South.

Endo-Marakwet is one of the linguistic varieties spoken within the Marakwet administrative unit. The Kenya National Bureau of Statistics (2009) shows Endo as spoken by a population of about 18,181. Zwarts (2003) specifically locates the language as being spoken in the area that spans from around Liter in the north to Arror in the south, along the Kerio River.

According to the Bible Translation Literacy (BTL) manual of (2006), Marakwet “languages” vary according to the location where each is spoken. This manual puts the Marakwet linguistic group into two major clusters, the Northern and the Southern speakers. Rottland (1982) also classifies Marakwet into Southern and Northern branches. Zwarts (2003) further confirms the existence of these sub-dialects and records that:

Originally there were several groups in the area, among which were the Endos in the North and Markwets in the South of the Valley (ibid.: 12).

Rottland (cited in Zwarts, 2003: 30) also comments that as compared to other Kalenjin groups, Marakwet is a vast area with more linguistic variation than any other.

This paper focuses on Endo-Marakwet, the variety that Kipchumba (2010) considers more distinct from the rest. It is spoken by the people living in a small location called *Endow* (the area that stretches between Tot and Liter). This variety is generally referred to as Endo by other speakers. The speakers of this language are also called Endo. For the purpose of this paper, the language is referred to as Endo-Marakwet. When asked where they hail from, speakers of Endo say they come from “Endow” /endaw/. When asked who they are, they also say they are /endaw/.

Specifically, the paper intends to explain the idea that in this language, it is possible to signal direction towards the speaker by use of tone and tongue root, independently of the segmental verbalitive and ventive. Previous discussions in this area have focused on the segmental use of these suffixes to signal direction.

¹ The research on Endo-Marakwet was made possible by funding from the DFG (Deutsche Forschungsgemeinschaft) who sponsored my PhD within the project “*Die Marakwet-Sprachen (Süd-Nilotisch): Deskription und Vergleich*”.

Generally in Kalenjin, motion and direction of the action as encoded by the verb are typically denoted by the use of the itive and ventive verbal suffixes. These morphemes are mainly segmental. They not only signal direction but also encompass elements of action and motion. It is this latter concept that I broadly describe as the “mobilitive itive” and “mobilitive ventive”. This idea is further exploited by Mietzner (2007, 2011). In Endo /-a:nu/ denotes motivated action in the direction of the speaker and is called mobilitive ventive:

- (1) túm-ánú
sing-VEN
‘sing while coming in this direction’ (towards speaker)

In Endo as well, /a:ta/ is used to denote motion in a direction away from the speaker. This mobilitive itive also takes the form /a:te/ in other Kalenjin languages such as Keiyo. This is also discussed by Creider (2002), who represents the suffix in Nandi as /-ta/, e.g.:

- (2) tyen-ate
sing-ITV
‘sing while moving away from speaker’

Creider (2002), in his study of Nandi, explains in detail the semantics associated with verbs formed by means of derivational suffixes, e.g. the dative /-ci/ and the instrumental /-e:/. Mietzner (2011) also demonstrates the phonological conditioning that the itive morpheme undergoes in Cherang’any language. She also discusses the mobilitive morphemes and their realisations in Cherang’any. These segmental morphemes (in Nandi and Cherang’any) occur across other languages in the Kalenjin group and are quite productive because they contribute further to the semantics of direction – they specify directions in relation to the speaker.

Heine (1984) observes that the majority of African languages display several productive verbal affixes and extensions and their sole purpose is to modify the semantic content of the verb root. Discussions by various scholars (Mietzner 2011; Creider 2002) are centred on morphemes that are segmentally realised. This paper will give an alternative view – that motion and direction are also realisable non-segmentally through the use of tone and tongue root alternations.

The paper discusses the use of the default ventive [u] and how it relates to the basic verb in the language. It also describes the infinitive and how the use of this infinitive leads to the omission of the segmental ventive that eventually leads to the alteration of the vowels in the verb root.

2 Ventive in Endo-Marakwet

Dimmendaal (2009: 10), in his description of Tama, explains ventive marking as a derivational notion expressing movement towards the deictic centre. This implies a ventive does not only express direction but also the idea of movement.

Data from Endo-Marakwet reveals that it is possible for the ventive to be realised as a phonological feature. This takes the form of a supra-segmental ventive realised through processes of tone and tongue alterations. This supra-segmental ventive is used in cases where the segmental morpheme [u], which is the ventive in Endo, is omitted. A typical verb in Endo that encodes motion in the direction of the speaker has two parts:

- a) The verb root, which has a vowel with H tone and –ATR or +ATR, and
- b) The ventive suffix, which has a vowel with L tone and +ATR.

Zwarts (2000) further confirms that prefixes in Endo and other African languages are predominantly –ATR.

This paper observes that whenever a segmental ventive is omitted, there occurs a compensatory phonological process. In Endo, prefixing the infinitive /ke-/ or /ke-/ on a verb causes the deletion of the physically realised ventive. This implies that it is not normally possible to have an infinitive prefix and the physical ventive attached to a verb root at one time. Only one can be used:

- (3) ké-àp
INF-bring
'to bring'
- (4) àp-u
take-VEN
'take in this direction' (towards speaker)

The majority of independent verbs in the language usually have H tones on their vowels. Introduction of the infinitive thus leads to the omission of the segmental ventive and vice versa. The omission of the segmental ventive creates phonological demands that lead to the supra-segmental features of tone and tongue root being re-adjusted to take the place of the omitted ventive. Such processes are discussed in the following sections.

3 Endo basic verbs

There are two vowel sets in the language which are distinguished in terms of tongue root. While some bear ATR others are RTR. These are shown below:

+ ATR		-ATR	
i	u	ɪ	ʊ
e	o	ɛ	ɔ
a		ɑ	

Dimmendaal (ibid: 4) further argues that there occurs what he calls “cross-height harmony”, which means that vowels in a given word can either be +ATR or -ATR. He provides data to support the idea that both roots and suffixes can be dominant with either of the tongue root features. Such dominance triggers harmony shift in the neighbouring vowels – either progressively or regressively. However, there also exists what are called opaque affixes, across which vowels do not usually harmonise. The present paper is partly a discussion of the opacity of the infinitive.

Quite a number of the non-derived verb stems are mostly characterised by monosyllabic vowels with H tone and +ATR quality.

These are exemplified below:

- (5) pan [pán] bewitch
 (6) ké:l [ké:l] fry
 (7) luch [lúk] box
 (8) sas [sás] hate
 (9) iit [i:t] peep
 (10) ket [két] drive
 (11) wis [wís] run fast

When the quality of such vowels in the verbs changes from H to L tone and from +ATR to -ATR, the semantic output is also altered. This suprasegmental change usually converts the verbs (with inherent -ATR vowels) to nouns (with +ATR vowels). Thus the above verbs are nominalised as follows (all these nouns are in their plural indefinite forms):

- (12) pan [pàn] sorcerers
 (13) kel [kè:l] people who fry
 (14) luch [lòk] people who box others
 (15) sas [sàs] people who hate others
 (16) iit [i:t] people who peep

4 The suffix [u]

Endo-Marakwet, like other Kalenjin languages, has both ventive and itive segmental suffixes that denote specific directions of the actions in the verbs. The morpheme [u]/[u:] is the ventive suffix in Endo, while [-ta] is the default itive. Once the ventive is suffixed to a verb, an action is understood as being executed in the speaker's direction.

The examples below show this:

- (17) a. mwá:r mwa:r-ù
say say-VEN
'say in our direction' (speaker included)
- b. kút ku:t-ù
blow blow-VEN
'blow in the speaker's direction'
- c. tá:k ta:k-ù
receive receive-VEN
'receive towards speaker'
- (18) a. páł pal-ù:
dig dig-VEN
'dig out in the speaker's direction'
- b. náp nap-ù:
weave weave-VEN
'weave in the speaker's direction'

The data above show that the length of the ventive [u] is dictated by the length of the vowel in the verb root. If the vowel in the verb is long, the ventive is shortened as in example 17 above. If the vowel in the verb is short, however, the ventive quantity is lengthened as in 18 above. This scenario points to the idea that the segmental ventive in Endo is actually a long [-u:] which is phonologically conditioned by the vowel in the verb root.

This argument is further supported by the fact that in Endo, there is another short [u] (with the feature [-ATR], L tone) which is used to mark aspect. This variant is not the ventive. Usually, however, in Endo, the morpheme that marks aspect is [a], whereas in other Kalenjin languages, it is [e], as in constructions like:

- (19) kòt-à
blow-ASP
'in the process of blowing'

[u] can also mark present imperfective aspect. In this case it has two functions – as an aspective marker and as a ventive. Thus:

- (20) pal-u
dig-ASP
'in the process of digging' (in the speaker's direction)
- (21) nàp-ù
weave-ASP
'in the process of weaving' (in the speaker's direction)

This is further illustrated in sentences as follows:

- (22) sàp-é:
basket-DET
'weave the basket'
- (23) nàp-ú sàp-é:
weave-IMP.PRES basket-DET
'(Somebody) in the process of weaving the basket in our direction'

The examples provided in (20)-(23) above help to show that there exists a difference between the pure ventive [o] and the [o] that expresses aspect. The central focus of this discussion is not the segmental ventive [o:]. It is however important to describe it at this point because it is what triggers the vowel quality changes in the verb root, which is the central topic of this paper.

5 Tone

The majority of the verbs in Endo are monosyllabic and have H tones. Creider and Creider (1989) in their study of Nandi establish that there are basically five surface tones. These are also five in Endo. They are H, L, HL, LH, LL. In these languages, tone can be used to differentiate lexical items (sort them into different classes) and may also be realised in the lexicon as morphemes.

It is this latter function of realisation as a morpheme that is relevant for this study. Consider the examples below:

- (24) wérá 'to pass by'
wèrà 'a friend'
- (25) kòwò 'a bone'
kòwó 'has gone'

These minimal pairs show that tone is lexical and a non-segmental morpheme that converts a lexical item from one class to the other.

6 Tongue root

Vowels in Endo can be separated on the basis of having the presence or absence of tongue root (TR). The root of tongue either advances +ATR or retracts -ATR. Though it is not productive in derivational processes, tongue root can be used as one means of determining the lexical class of a word or in certain aspects of the grammar. Below are sets of minimal pairs which belong to different lexical categories based on tongue root as a feature.

a) Lexical class

- | | <u>Verbs</u> | <u>Nouns</u> |
|------|--------------------|----------------|
| (26) | [pɔ:r]
thresh | [po:r]
body |
| (27) | [tun]
in future | [ton]
marry |

b) Grammatical aspects

- | | | |
|------|--|--|
| (28) | a. kà-ápé
ASP.PR-go
'have you gone?' | b. ká-ápé
ASP.PST-go
'had you gone?' |
|------|--|--|

In the examples provided above, (28a) demonstrates the use of the present perfect while (28b) is in the past perfect. This suggests that tongue root is a distinctive feature.

7 Phonological processes

Having considered the separate phonological behaviours of the morpheme [ɔ], the basic Endo verbs and the infinitive, it is prudent to analyse how supra-segmental aspects of tone and tongue root play a part in the realisation of motion/directional semantics.

There are several phonological processes that take place when the segmental ventive suffix [u] is omitted and the infinitive /ke:/ is affixed to the verb. This discussion dwells on the phonological processes affecting the properties of the vowels in the infinitive and the verb root.

Quite naturally, the segmental suffix [ɔ:] and the infinitive prefix [ke:] cannot be used at the same time on a common verb root. They are in opposition and this implies that whenever the infinitive is used, the segmental ventive is blocked from use. However, the supra-segmental ventive may be used. The morpheme [u:] is the segmental ventive. When this morpheme is omitted, phonological factors occur to compensate for it through the non-segmental features of tone and tongue root. When this occurs, these features/qualities float to the vowel in the verb and segmental omission takes place. Such processes are explained in the section that follows.

8 Omission of the segment [u:]

The direction of action in a verb (towards the speaker) is naturally denoted by the segmental morpheme [u:]. An example is provided below:

- (29) àp-ù:
take-VEN
'take towards the speaker'

This morpheme is omitted when the infinitive /ke:/ is introduced. This omission has certain phonological implications:

- The qualities of this morpheme, i.e. –ATR and L tone, are not lost with the segment; instead they attach to the vowel segments found in the verb root. The root has vowels which are +ATR and H in tone. Once these different features (of the vowels in the verb root and from the traces from the ventive morpheme) are brought together, neutralisation occurs. The new features, originally belonging to the omitted ventive morpheme, neutralise those of the verb root. The former take precedence and the result is a verb with vowels that are –ATR and have L tone. Zwarts (2003) further states that this feature (TR) is not segmental but contributes to units of length and changes tone patterns.
- The new surface form of the verb has phonological qualities that were initially contained in the omitted ventive morpheme. This implies that the semantic value of the omitted segmental ventive – i.e. towards speaker – does not disappear with the omitted segment but rather transfers to the verb root. Thus the semantic function of the omitted morpheme is apparently maintained in the supra-segmental features, –ATR and L tone.

The omission of the ventive [u:] morpheme creates a syllabic mismatch and mis-timing between the original (verb + segmental morpheme) and the new form (verb + non-segmental morpheme). The original form in (30a) below has two syllables. (30b) has one syllable due to the omitted ventive:

- (30) a. àp-ù:
take-VEN
'take in the direction of the speaker'
- b. ap
take:VEN

Phonologically, there has to be syllabic harmony, which warrants the need for the infinitive to fill the gap. A new process occurs in which the infinitive is prefixed on the new verb form, as shown below:

- (31) kè-ap
 INF-take: in speakers direction
 ‘to take towards the speaker’

In Endo, as well as in all other Kalenjin languages, it is not possible to have a verb with –ATR and also L tones standing independently. Such verbs need the support of other morphemes, in this case the support of the infinitive.

Phonological alterations affect the suffix and the verb root. By this I mean that the deletion of the ventive suffix leaves its vowel qualities to cross over to the vowel of the verb root. Thus the vowel of the verb root is influenced by the ventive suffix, the changes following the pattern below:

infinitive || verb root vowels ← ventive suffix

Dimmendaal (2009: 167) argues for neutral/transparent affixes, which are morphemes that do not block harmony, versus opaque affixes, which are those morphemes that do block harmony spreading. He further explains that all bound morphemes in Kalenjin are integrated into the vowel harmony system. Their tongue root features change according to the tongue root quality of the neighbouring morpheme. Dimmendaal’s example is as follows:

- | | |
|--|-------------------------------------|
| (32) ke:-pɪ:t
INF-grow
‘to grow’ | ke:-pɪ:t
INF-spray
‘to spray’ |
|--|-------------------------------------|

This suggests that vowels in both syllables must share a common feature with regard to the tongue root. They are either +ATR or –ATR.

The ventive in Endo is +ATR and it triggers harmony shift in preceding vowels to its left.

It has been argued that there exist opaque bound morphemes which do not permit harmony across vowel segments. One such morpheme is the infinitive. In Dimmendaal’s example:

- (33) Ke:-ke.r-ke:
 INF-see-REFL
 ‘see oneself’

The reflexive in this example lacks harmony with the vowel elements in both the verb root and the infinitive. The infinitive and the verb root “see” are neutral with regard to their tongue root value and are not affected by the tongue root quality of adjacent morphemes. In this case, the reflexive is +ATR while the infinitive and the verb roots are –ATR.

In a regressive pattern, influence and harmony can stop at the transparent morpheme if all the rest are opaque, as example 33 above shows. Sometimes however, harmony can pass through several steps /morphemes (as long as they are transparent) but stops as soon as it encounters an opaque morpheme. This paper treats the infinitive as an opaque morpheme that ultimately blocks further harmony processes. Phonologically, this could be attributed to the fact that, since the infinitive already has –ATR and L tone, and the final output of the new verb form also has –ATR and L tone, there is nothing for them to share over – they are already in harmony.

Note that basic verbs in Endo have predominantly H tones on the vowels and so have the appearance of being –ATR. Any addition of an affix distorts phonological qualities of the vowels.

For this reason, for the verbs to maintain the same meaning of action done in the direction of the speaker, the use of the infinitive is required. The infinitive is /ke-/ and in Endo it naturally has the L tone. Since the vowel in the new verb form also contains a L tone (as donated by the [v:] morpheme), there is vowel harmony in terms of tone and tongue root, both of which are major characteristics of the Nilotic languages.

9 Semantic implications

The various phonological processes that the vowel in the ventive morpheme undergoes, and the qualities that are transferred the verb root, have been presented. The new verb form takes the infinitive with vowel features –ATR, L. The result of all these processes is a verb structure that has the following shape:

infinitive + verb root

The infinitive has a vowel with the qualities –ATR, L, whereas the root has a vowel with the qualities –ATR, L.

Unlike previous studies, which discuss the use of segmental morphemes and how these are used to denote the semantic notions of motion and direction, this study reveals that it is possible to express the same semantic features of direction without the addition of a segmental morpheme but through phonological compensatory processes where omitted segments are replaced supra-segmentally.

Consider the examples below:

	<u>simple infinitive verbs</u>	<u>infinitive ventive verbs</u>
(34)	[kè:-còr] INF-steal 'to steal'	[kè:-cór] INF-steal; speaker's direction 'to steal heading in the speaker's direction'
(35)	[kè:-ryèp] INF-crash 'to crash'	[kè:-ryép] INF-crash; speaker's direction 'to crash heading in the speaker's direction'
(36)	[kè:-mò] INF-scare 'to scare'	[kè:-mú] INF-scare; speaker's direction 'to scare heading in the speaker's direction'

The data in the left column above are infinitive verbs which basically convey the actions encoded in the verbs.

The examples in the right column contain the infinitive but the qualities of the vowels have changed. Arising from these alterations, there is extra semantic information provided by the verb – that there is some kind of motion involved in the actions. One example below illustrates this:

- (37) [kè:-còr]
INF-steal; speakers direction
'to steal in the speaker's direction'

To the native speaker, this is not simply stealing but theft that is carried out towards the speaker's direction. This is where the element of motion comes in: that there is an implied direction in which the action is done in the speaker's direction.

Infinitive verbs of this kind in Endo (which do not have segmental ventives) refer to actions that are performed by somebody in the direction of the speaker.

10 Conclusion

Morphemes – and directional morphemes in particular – are a rich field for study, especially in Kalenjin languages. Detailed morphophonological and morphosemantic descriptions abound. Despite such work, there is a need for further discussions about supra-segmental analysis, particularly relating to the absence of segments and the way vowel qualities may be copied from deleted segments or morphemes.

This paper has shown that there is deep analytical gap between segments and supra-segments at the level of semantics. Motion and direction are effectively encoded by both segments and supra-segments but the circumstance in which one or the other is used is a question for both syntax and phonology.

This paper demonstrates the possibility of marking motion and direction on a verb in Endo without the physical inclusion of a segmental ventive morpheme. It suggests a binarity in the semantics of direction, in terms of whether it is segmentally or supra-segmentally marked. Supra-segmental morphology and phonology provides alternative productive processes used to mark direction. Endo effectively makes use of tone and tongue root vowel qualities to indicate motion and direction in its verbs.

Phonetic qualities of an omitted ventive are transferred to the vowels in the verb root, which then takes an infinitive prefix. The new infinitive verbs generated have an encoded direction as part of their semantics. Additional information is added to the action encoded by the verb: the motion and direction of the action is specified. The additional meaning encoded denotes that the action is performed in the process of the actor moving in the direction of the speaker.

Abbreviations

ASP.PR	aspect in the present	NEG	negation
ATR	advanced tongue root	NMLZ	nominalizer
DEF	definite	P1, P2,	past 1, 2, 3
H	high tone	P3	
IND	indefinite	PL	plural
INF	infinitive	PST	past
IPF	imperfective	SG	singular
ITV	itive	VEN	ventive
L	low tone	VSO	verb, subject, object
MOB	mobilitive		

References

- Bible Translation & Literacy. 2007. *Kaaleweenuut akoo Tyaanko chepo Markweeta*. Kenya: Uasin Gishu printers.
- Creider, Chet A. & Jane T. Creider. 1989. *A Grammar of Nandi*. (Nilo-Saharan, 4), Hamburg: Buske.
- Creider, Chet A. 2002. The semantics of participant types in derived verbs in Nandi. *Revue Québécoise de linguistique* 31.2: 171–190. <http://id.erudit.org/iderudit/009316ar>.
- Dimmendaal, Gerrit J. & Anneke Breedveld. 1986. Tonal influence on vocalic quality. In Koen Bogers, Harry Van der Hulst & Maarten Mous (eds.), *The Phonological Representation of Suprasegmentals*, pp. 1–33. Dordrecht: Foris.
- Dimmendaal Gerrit J. 2002. Constraining disharmony in Nilotic: what does an optimal system look like? *Journal of African Languages and Linguistics* 23: 153–181.
- Dimmendaal Gerrit J. & Anneke Breedveld. 2009. Tama. In Gerrit J. Dimmendaal (ed.), *Coding Participant Marking*, pp. 307–331. Amsterdam/Philadelphia: Benjamins.
- Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
- Kenya National Bureau of Statistics. 2009. National census. Nairobi, Kenya.
- Kipchumba, Paul K. 2010. *Oral Literature of the Marakwet*. Eldoret: Infomercial Publishers.
- Kipkorir, Benjamin E. & F. B. Welbourn. 1973. *The Marakwet of Kenya: A Preliminary Study*. Nairobi: East African Literature Bureau.
- Kipkorir, Benjamin E. & F. B. Welbourn. 2009. *Descent from Cherangany Hills*. Nairobi: Macmillan.

- Mietzner, Angelika. 2007. Location, motion and movement in Keyo (Southern Nilotic). *Annual Publication in African Linguistics*, 5: 145–164.
- Zwarts, Joost. 2000. *Number in Endo-Marakwet*. Netherlands: Utrecht Institute of Linguistics OTS.
- Zwarts, Joost. 2004. *The Phonology of Endo*. Netherlands: Utrecht Institute of Linguistics OTS.

THE RELATIONSHIP BETWEEN SYLLABLE WEIGHT AND STRESS IN DHOLUO

Jane Akinyi Ngala Oduor

1 Introduction

Existing literature points out the possibility that placement of stress in Dholuo, a Nilotic language spoken in Kenya along the Eastern shores of Lake Victoria, is subject to syllable weight. This paper then attempts to show the relationship between syllable weight and stress in Dholuo. Since Dholuo has both heavy and light syllables, the discussion starts with an analysis of the relationship between the weight structure and stress pattern of bisyllabic, trisyllabic and other polysyllabic words using metrical grids. The same analysis is carried out briefly at the sentence level to show the similarity between lexical and sentence level stress. The paper shows that stress seems to coincide with syllable weight. So, there is a distinct relationship between syllable weight and stress. Syllables that are stressed in almost all cases are the syllables that are heavy. Stress has the effect of lengthening vowels in Dholuo, thus transforming syllables into heavy ones. The paper also shows that stress in Dholuo is detected in the penultimate position in a vowel final word and in the last syllable in a consonant final word. It is observed that there are Dholuo syllables that appear to be inherently heavy because they have either a CVV or CVVC structure, which make them stressable. Most of these inherently heavy syllables appear to be at the right position, that is, where stress is easily detected in Dholuo. A light syllable in Dholuo has a short vowel and is unstressed in almost all cases. It is noted that very few words have a sequence of light syllables only. At sentence level, the emerging pattern of stress is similar to that of words uttered in isolation. It is concluded that Dholuo is a fixed stress language with stress playing the role of demarcation in that it marks the end of an utterance. Stress is not phonemic in Dholuo.

Dholuo belongs to the Western Nilotic branch of the Nilo-Saharan family of languages. It is spoken mainly in the south western part of Kenya and in the northern part of Tanzania. This language has two main dialects (Oduol 1990). These are the Kisumu-South Nyanza (KSN) dialect spoken in a wider geographical area and Boro-Ukwala (B-U) spoken in a smaller region. The KSN variety is discussed in this paper. It is important to specify the variety under study here because apart from the differences in tone, there is a difference in the realisation of vowel length in the two varieties.

2 Dholuo sounds and syllable structure

Dholuo has 26 consonants, five of which are prenasalised. It has nine pure vowels. Eight of these vowels can be grouped into 4 pairs. Each pair of vowels contains an advanced tongue root [+Advanced Tongue Root] ([+ATR]) and a retracted tongue root [-Advanced Tongue Root] ([-ATR]) partner (Okombo 1982). The vowel /a/ is [-ATR] and does not have a [+ATR] partner.

Since the paper investigates syllable weight in Dholuo, the syllable types in the language are identified. Using C to represent a consonant and V to represent a vowel. The syllable types that exist in Dholuo are CV, V, CV:, V:, CVV CVC,VC, CV:C, V:C and CVVC. It is these syllable types that feature in the illustrations in the sections that follow.

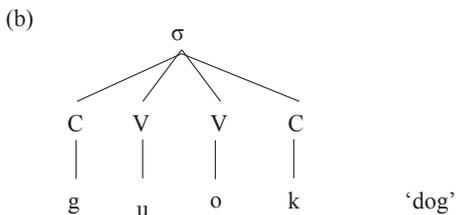
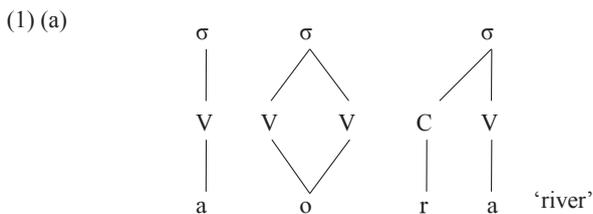
3 Syllable weight

The concept of syllable weight is discussed in this section. Katamba (1989), in his classification of languages according to syllable weight, notes that a language with a heavy and light syllable dichotomy follows one of the two patterns specified below. The first class of languages considers a syllable which consists of a short vowel as light, while a syllable with a long vowel or a long vowel followed by a consonant (or consonants) or a short vowel followed by a consonant (or consonants) is considered to be heavy. In this class, it is noted that a short vowel followed by a consonant is a heavy

syllable. The second class of languages considers a syllable whose rhyme consists of a short vowel to be light, while a syllable whose rhyme consists of a long vowel is heavy. In the second class, the consonants in the rhyme are not considered to be as crucial in defining the weight structure of syllables. Katamba (1989) refers to the first class of languages as type A and the second as type B. Hyman (2003: 5), while noting that the ‘... distinction between “heavy” and “light” syllables has received considerable attention in recent works’, seems to recognise the two patterns or classifications already identified in Katamba (1989). The main difference between the first class and the second class of languages is that in the first class, the existence of the coda or margin may or may not be important in the determination of the weight of a syllable whereas in the second class, the coda or margin is completely irrelevant or does not count.

A study by Oduor (2002) shows that Dholuo, has both heavy and light syllables. In other words, some syllables in Dholuo have a greater weight or quantity than others. She notes that heavy syllables in Dholuo have the following structures CV:, V:, CV:C, V:C, CVV and CVVC. For a syllable to be heavy in Dholuo, the nucleus must branch. In other words, the nucleus must occupy two Vs as seen in the second syllable of (1) (a) and in (1) (b). The light syllables in Dholuo have the following structures CV, V, CVC and VC. The nucleus of each light syllable does not branch. Using Katamba’s (1989) classification of languages according to syllable weight, Dholuo then belongs to the second class of languages, i.e. type B languages because the consonant in the rhyme does not count in the weight of a syllable.

In (1) (a) below, the first and last syllables are examples of light syllables while the second one is heavy. In (1) (b) another example of a heavy syllable in Dholuo is given.



Laver (2000: 517), in discussing syllable weight, states that:

Syllable weight encompasses two types of syllable: ‘light’ syllables and ‘heavy’ syllables. A **light syllable** is one whose rhyme is made up of a nucleus consisting of a short vowel, followed by a maximum of one short consonant ... As a measure of syllable quantity, the phonological length of a light syllable has been called a **mora** (Trubetzkoy 1939). A **heavy syllable** is any other type of syllable, and its phonological length is greater than one mora.

Syllable quantity or weight therefore refers to the phonological length of each syllable. Light syllables have a shorter duration than heavy syllables.

4 Stress

4.1 Stress in languages

This section briefly discusses the nature and location of word stress in languages generally. Kenstowicz (1994) states that in generative grammar, stress is seen as an abstract relation of prominence whose existence is signalled through other prosodic features such as heightened pitch, increased duration and subtle aspects of vowel and consonant quality. He notes that stress is a property of vowels and more generally of syllables. He also asserts that the stress properties of a given syllable can be determined by the way syllables of a word are treated by the grammar of a language and by the perceptions and judgements of the native speaker. Another evidence that can determine the stress properties of a syllable are the adjustments a language may resort to in order to avoid a sequence of stressed syllables.

In placing word stress, languages that pay attention to particular properties of syllables or to a particular position of a syllable in a word are known as fixed stress languages because stress is restricted to a particular syllable (Hyman 1975). In some languages, the distribution of stress takes syllable weight into consideration. Just as Hyman, Kenstowicz (1994) notes that a syllable bears stress because of its location in a word. While the latter states that stress is initial in Finnish and Czech, the former states that it is initial in Hungarian. It is final in French and Turkish and penultimate in Polish. Carr (1999: 88) states that in Modern Greek, there seems to be no generalisations about the stress pattern of words meaning that it is arbitrary. Kenstowicz, 1994 therefore agrees with Carr when he states that stress of a syllable may be an unpredictable property of the lexical items in which the syllable is located. In other words, the make-up of a word determines stress placement. In consideration of the make-up of a word, prefixes, roots and suffixes may also determine stress placement. Hyman (1975) calls such languages free stress languages. According to Carr (1999: 99), ‘... stressed syllables are perceptually more salient than unstressed syllables’. This paper intends to show not only where stress in Dholuo words and sentences is located but also its nature or how it is manifested in words. Therefore, it is these salient parts that the rest of this paper seeks to identify and show how they relate to syllable weight.

4.2 Stress in Dholuo

The earliest literature on stress in Dholuo is found in Stafford (1967). Oduor (2002) states that Stafford discusses stress but not in relation to vowel lengthening. Stafford states that stress is normally on the last-but-one syllable and does not move when suffixes are added. He further states that unlike other suffixes, the plural imperative suffix *-uru* is an exception because it bears stress. According to Stafford, the first syllable of many bisyllabic words is a prefix which is never stressed hence the stress falls on the last syllable. The words he gives as examples are in (2) below. The symbol ‘¹’ indicates the location of stress in the words in (2) and in all the other sections where word stress is marked while ‘-’ indicates syllable boundary.

- | | | | | | | |
|-----|-----|----------------------|---------|-----|-----------------------|---------|
| (2) | (a) | à- ¹ pâ:r | ‘ten’ | (c) | ð- ¹ mbó:ŋ | ‘ankle’ |
| | (b) | à- ¹ dé:k | ‘three’ | (d) | ɲà- ¹ oí | ‘child’ |

He notes that loan words are usually restressed to fit within the stress pattern of Dholuo. An example of a loan word that is restressed is *ká-¹mé:-rà* ‘camera’. In examples (2) (a), (b) and (c), vowel length is indicated in the second syllable of each word, unlike in Stafford (1967). Additionally, tone and syllable boundaries are marked in all the examples, unlike in Stafford. The symbols ¹ - represents a downstepped high tone, ´ represents a high tone, ` a low tone and ^ a falling tone.

Odhiambo (1981) observes that whereas stress has a tendency to lengthen vowels, syllables are stressed depending on their phonological structure. She points out the possibility that placement of stress is subject to syllable weight. Omondi (1982) notes that in Dholuo vowel lengthening is a product of stress and so, every stressed syllable may be lengthened. She states that:

Dholuo word-roots are monosyllabic. Stress is fixed and it falls on the root syllable of the word. For example, in the citation form of the verb when the final vowel is almost

always [o / ɔ] and as such could be regarded as the verb-root. So in a verb like *tedo* (to cook), stress is on the *e* ... (Omondi 1982: 21)

She sees Dholuo as an example of a language, in which according to Lehiste (1970), length or duration is one of the phonetic manifestations of stress. According to Omondi therefore one of the ways in which stress in Dholuo is realised is through length. However, Odhiambo (1981) and Omondi's (1982) discussions are quite general as stress is not at the centre of their studies.

5 General literature on syllable weight and stress in languages

Having stated that C is the onset in a CVC syllable while VC is the core, Hyman (1975: 206) notes that:

In many languages, a syllable whose core consists of a short vowel (V) cannot be stressed and stress must pass to a neighbouring syllable. Such a syllable is said to be *light*. A syllable whose core consists of a long vowel (V:), a VV or VC sequence, or combinations of these, can be stressed and is said to be *heavy*.

Hyman (1975: 206) then quotes Newman (1972) and Allen (1973) in stating that the distinction in syllable weight, described in the preceding quotation, is 'an important phonological variable in the statement of stress placement'. In many languages, only heavy syllables can receive stress. Like Hyman (1975), Katamba (1989) notes that in many languages, syllable weight has an effect on the applicability of certain phonological rules.

Kenstowicz (1994), on his part, notes that phonological quantity in most cases determines the distribution of stress. Therefore, the concept of syllable weight is relevant in the phonology of some languages because the weight of certain syllables determines the behaviour of other phonological phenomena such as stress.

Laver's (2000:517) discussion also goes beyond the mere classification of syllables into heavy and light ones by relating syllable weight to prominence. He states that:

The phonetic realisation of heavy syllables, with their longer nuclear vowels and/ or their more substantial codas, stand out more prominently in perceived flow of speech than do light syllables.

It is thus noted that heavy syllables tend to be more prominent or noticeable than light syllables.

This section has therefore shown that languages have both light and heavy syllables and that syllable weight is relevant in prominence relations between syllables. This paper attempts to illustrate that the assertion made in the literature, for example, by Hyman (1975), Katamba (1989), Kenstowicz (1994) and Laver (2000) is true for Dholuo.

6 Syllable weight and stress in Dholuo

Both short and long words are used to exemplify the relationship between syllable weight and stress in Dholuo. The discussion starts with an analysis of bisyllabic words followed by trisyllabic and other polysyllabic words. Even though Dholuo has very many monosyllabic words, they are not used in the discussions because for a meaningful analysis of stress, words of two or more syllables are required. Though most of the discussions are based on the lexical level, the sentence level is included at the end to show the similarity between word level and sentence level stress.

Metrical grids of the theory of Metrical Phonology as described by Goldsmith (1990) are used in this paper. Metrical grids help to show the relationship between stress placement and syllable weight. In constructing Metrical Grids for Dholuo words, only two rules are used. The two rules are Quantity Sensitivity and End Rule (Final and Initial). In the application of Quantity Sensitivity, a grid mark is placed, on Row 1 (foot row), over any heavy syllable. Quantity Sensitivity alone is enough to show which syllable is stressed, in most Dholuo words. In some words, End Rule Final is applied in addition to Quantity Sensitivity. Therefore, whereas Quantity Sensitivity is applied at the foot level, End Rule Final applies at the word level, as seen in the example in (3).

(3)	phonemic tier	w	ù	ò	n	d	-r	ù	ò	k	‘pretence’
	Row 0 (mora row)		x	x				x	x		
	Row 1 (foot row)		x					x			
	Row 2 (word row)								x		

In (3), on Row 0 (mora row) a grid mark is placed on each mora. Each syllable has more than one mora. This means that both syllables are heavy. Quantity Sensitivity then applies on Row 1 because a grid mark is placed over all the heavy syllables. End Rule Final applies on Row 2 (word row) because a grid mark is placed on the extreme right. End rule usually places a grid mark on the extreme right or left of the domain it is specified for. For Dholuo words with a heavy syllable, End Rule places a grid mark on the right end of a word (final position). End Rule Final helps to show that it is the last heavy syllable of the word that is stressed.

In words with light syllables only, Quantity Sensitivity does not apply. End Rule Initial or Final applies. More explanations on how grids are used are given in the course of the discussions.

6.1 Bisyllabic words

This section is concerned with the relationship between stress and syllable weight in Bisyllabic words. Some bisyllabic words in Dholuo have stress on their first syllable while others have it on the second syllable as seen in the discussions below.

6.1.1 Base forms of bisyllabic words

A majority of these verbs are infinitive verbs ending in -o or -ɔ. These verbs have a ¹CV:-CV syllable structure. The stress pattern of such verbs is shown diagrammatically in (4).

(4)	Phonemic tier	b	è	è	d	ò	‘to sit’
	Row 0 (mora row)		x	x			
	Row 1 (foot row)				x		

In (4) the phonemic tier shows the sound segments that form this word. Row 0 then shows the number of moras in the word. It is noted that the word has two moras in the first syllable and one in the second which means that the first one is heavy unlike the second one. On Row 1, Quantity Sensitivity applies to create a grid mark on the first syllable because it is heavy. The first syllable has a grid mark on the foot row to show that it is stressed. More examples are given in (5). Stress is marked at the beginning of the syllable.

(5)	¹ gǔ:-jǔ	‘to beat’
	¹ ke:-tò	‘to put’

As seen in examples (4) and (5), stress is on the initial syllable. The CV-tier shows that all these initial syllables have a long vowel. Stress is on the syllable with a long vowel and it is this same syllable that is heavy. The second syllable in each word has a short vowel and it is not stressed. It is a light syllable.

Apart from the verbs in (4) and (5), words from other classes (such as nouns, adjectives, etc) have a similar stress pattern. Examples of these non-verbs are given in (6).

(6)	¹ mù:-ðò	‘darkness’	¹ ò:-kò	‘outside’
	¹ mò:-ndò	‘so that’	¹ ɲù:-gú	‘groundnuts’

Some bisyllabic words have a vowel-vowel sequence in the nucleus of the first syllable. Stress is on this initial syllable. Most of them are also are infinitive verbs ending in -o or -ɔ. The metrical structure of these verbs is similar to the one in examples (4), (5) and (6) as seen in (7) below.

(7)	¹ ciě-lò	‘to fry’
	¹ gòě-jǔ	‘to kick’

Just as the verbs given in (7), words from other classes, apart from verbs, have a high vowel-vowel sequence in the first syllable and a short vowel in the second syllable. The first syllable is stressed and is heavy, unlike the second. They have a ¹CVV-CV syllable structure. Consider the examples given in (8).

- (8) 'jíé-¹ló 'python'
'jíé-gó 'jealous'

In this group of words, the second syllable is an open short syllable while the penultimate syllable, which also happens to be the first syllable of the word, has either a long vowel or a vowel-vowel sequence. This is where stress is located.

The next group of words has the structure ¹CVVC-CV. The first syllable of each of these words is stressed and is heavy while the second one of each is light. The examples of these words are given in (9).

- (9) 'lòàŋ-nì 'house fly'
'cièg-nì 'near/ close'

Some bisyllabic words with stress on the initial syllable have the structure ¹CVC-CV. The two syllables have a short vowel but the first syllable is arrested by a consonant, as seen in (10). In constructing metrical grids for words in (10), Quantity Sensitivity and End Rule Final do not apply. It is End Rule Initial that applies at the foot level because the first syllable is articulated with some degree of prominence unlike the second, yet the two syllables are of equal weight.

- (10) 'kùd-nì 'worm'
'θìr-nò 'state of not growing'

As seen in the discussions following example (23), *nì* and *nò* behave like the plural formation suffix *nì* because they do not bear stress.

6.1.2 Prefixes in bisyllabic words

Words containing prefixes are discussed in this section. The first group of words has the structures V-¹CV:C or CV-¹CV:C. The first syllable of each word is a prefix. It is not stressed and is light, while the second one is stressed and is heavy, as seen in example (11).

- (11) à-¹jâ:ŋ 'an abuse' ð-¹kâ:θ 'name of male born when it is raining'
jò-¹tì:c 'workers' à-¹kâ:θ 'name of female born when it is raining'

Some of the prefixes found in the words in this category are '*jâ-*', '*jâ-*', '*jò-*', '*râ-*' and '*à-*' or '*ò-*'. '*jâ-*' is used before certain nouns to denote small or young things (Stafford 1967). '*jâ-*' could also mean 'daughter of'. '*jâ-*' and '*jò-*' are noun forming prefixes. They are prefixed to words to show their work or class (Stafford 1967). Whereas '*jâ-*' denotes singular, '*jò-*' is plural. The prefixes '*à-*' or '*ò-*' are also noun forming prefixes. Sometimes they are used to form proper nouns as seen in the names ð-¹kâ:θ and à-¹kâ:θ. All the prefixes in (11) are not stressed and they form light syllables. The second syllable of each word has a long vowel, is heavy and is stressed.

Bisyllabic words in the second group have the structures V-¹CVVC or CV-¹CVVC. A few examples of such words are given in (12).

- (12) ò-¹ndîèk 'hyena' rà-¹bùòr 'brown'
jà-¹gùòk 'puppy' ò-¹wúór 'a male name'

Just as in (11), some of the first syllables of (12) are prefixes. All the first syllables of the words are light and are not stressed. The second syllable of each word has a sequence of two vowels and it is heavy.

The third group of bisyllabic words with prefixes has the structures CV-¹CVV or V-¹CVV. In this category of words, stress is on the second syllable, as seen in the examples in (13).

- (13) jà-'kùò 'thief'
 ò-'gúé 'lizard'

The fourth group has the structure CV-¹V:C. The first syllable has a single short vowel, while the second one has a long vowel. The second syllable is stressed and therefore it is heavy. The word in this group is given in (14).

- (14) rà-'ú:m 'lid/ cover'

The rule of Quantity Sensitivity applies on the foot row, just as in all the other words with one heavy syllable.

The fifth group of words in this section has the structures ¹CV-CV and ¹V-CV. Once again both syllables have a short vowel and are light. The second syllable being the root of the word is articulated with greater force than the first. End Rule Final applies to create a grid mark on the last syllable.

- (15) jà-'θí 'child'
 ò-'ló 'name of a person'

6.1.3 Suffixes in bisyllabic words

When suffixes are added to replace the vowels -o and -o in the verbs in (4) and (5) stress does not shift. The vowel in the first syllable remains long, while that of the second syllable remains short. The examples are given in (16).

- (16) 'bè:-dí 'you sit (command)'
 'gò:-jé 'beat him/ her/ it (command)'

All the suffixes in (16) are personal pronouns. According to Okombo (1982: 42) these personal pronouns have both the free and bound forms as follows:

- | | | |
|------|------------|-------------|
| (17) | Free forms | Bound forms |
| | án | a- |
| | ín | i- ~ i- |
| | én | e- ~ e- |

It is the bound forms that occur in the words in (16). -a refers to 'me', -i or -i refer to 'you' (second person singular), while -e or -e refer to 'him/ her/ it'.

Stress does not also shift when suffixes are added to the base form of the bisyllabic nouns discussed in (6). The base form of these bisyllabic nouns is ¹CV:-CV. The bisyllabic nouns created as a result of suffixation also have the structure ¹CV:-CV, as shown in (18).

- | | | | | |
|------|----------|-----------------------|------------------------|-----------------------------------|
| (18) | 'wé:-ndó | 'visitor (base form)' | 'wè:-ndé | 'his/ her visitors' |
| | 'ri:-ŋó | 'meat (base form)' | 'rí:- ¹ ŋgí | 'your meat (2nd person singular)' |

It is also noted that when these suffixes are added to some monosyllabic words of the form CV:C or V:C, the resulting words have the syllable structures ¹CV:- CV or ¹V:-CV, respectively. The first

syllable of each word has a long vowel and it is stressed unlike the second, therefore it is heavy, unlike the second. The examples are given in (19).

- | | | | | |
|------|-------|--------------------|-----------|---------------------|
| (19) | jí:ŋg | ‘name (base form)’ | ʼjí:-ʼŋgá | ‘my name’ |
| | ò:r | ‘brother-in-law’ | ʼò:-rá | ‘my brother-in-law’ |

The examples in (19) show that there is resyllabification of the final C in CV:C and V:C sequences, when suffixes are added to form new words. Since a majority of syllables in Dholuo have a CV structure, the resyllabification of the coda consonant in CV:C and V:C syllables occurs to create the preferred syllable structure. Once resyllabification has taken place, the vowel in the root of the word remains heavy and is still stressed. The vowel in the root is stressed unlike the one in the suffix.

When suffixes are added to the verbs to replace the vowels -o and -ɔ in (7), stress does not shift. The first syllable remains stressed while the second one is unstressed, as seen in (20).

- | | | |
|------|---------|---------------------|
| (20) | ʼgòè-jé | ‘kick him/ her/ it’ |
| | ʼθiè-ǎá | ‘treat or cure me’ |

The base form of some of the bisyllabic nouns behave in the same way as the verbs in (20) when suffixes are attached to them. Stress does not shift. All the base forms of these bisyllabic nouns have the structure ʼCVV-CV. The words formed as a result of suffixation also have the structure ʼCVV-CV, as seen in (21).

- | | | | | |
|------|---------|--------------------|----------|-----------|
| (21) | ʼlúé-dó | ‘hand (base form)’ | ʼlúé-ʼtá | ‘my hand’ |
| | ʼtùè-lò | ‘leg (base form)’ | ʼtúé-ndé | ‘legs’ |

Suffixes added to monosyllabic words of the form CVVC result in the formation of words with the structure ʼCVV-CV. The first syllable has a sequence of two vowels and it is stressed. The first syllable is therefore heavy unlike the second. Once again resyllabification takes place to create the preferred syllable structure. The examples are given in (22).

- | | | | | |
|------|------|---------------------|----------|------------------------|
| (22) | θùòl | ‘snake (base form)’ | ʼθúó-ndé | ‘his/ her / its snake’ |
| | rùòθ | ‘lord (base form)’ | ʼrúó-ʼǎá | ‘my lord’ |

The plural formation suffix ʼnǐ’ is yet another suffix that does not bear stress. Stress remains on the root. The last syllable of the words in (10) resembles this suffix.

- | | | |
|------|---------|---------------------|
| (23) | ʼpèl-nǐ | ‘knives’ |
| | ʼlèw-nǐ | ‘clothes, garments’ |

As the words in (24) show, the two syllables of each word are similar in structure. In other words, they both have a CVVC structure. Stress is on the last syllable, which has a CVVC structure and is therefore heavy. If it is to be assumed that in the words in (20) and (21) stress is on the first syllable because of the presence of two consecutive vowels, then the two syllables of (24) should receive equal stress. However, stress is on the last syllable. The rule of Quantity Sensitivity applies to create a grid mark over all heavy syllables. A grid mark is therefore placed over both syllables on the foot row. This means that both syllables are stressable. However, one of them bears stress. Since it is the last syllable that is stressed, End Rule Final applies to create a grid mark over this last syllable.

- | | | |
|------|--------------|----------------|
| (24) | pùòŋj -ʼrùòk | ‘learning’ |
| | jùàg-ʼrùòk | ‘pleas/ cries’ |

Words in this set have the structure CVC-ʼCVVC. They are formed from verbs by the addition of the suffix *rùòk* to the root of verbs which have a reflexive form (Stafford 1967). Some of their reflexive

forms are seen in (41) and (42) (e.g. *pa-kɔ-re* ‘to praise (reflexive)’). The examples of these words are given in (25). These words are formed in the same way as those in (24).

- (25) pòg-¹rùðk ‘difference/ change’
wàr-¹rùðk ‘salvation’

The root of the words in (24) and (25) is not stressed, while the suffix is stressed. As already noted in the literature review, Stafford (1967) claims that stress does not shift when suffixes are added and that only the suffix *-ú:-rí*, which is almost equal in meaning to ‘please’, bears stress. However, it is my observation that apart from the suffix *-ú:-rí*, the suffix *rùðk* which is a noun-forming suffix, bears stress. The second syllable of each of these words is therefore stressed and it is heavy. In (25) the first syllable is not stressed and it is light.

One of the possible reasons why the suffix is stressed and not the root is that the suffix has a sequence of two vowels unlike the root. The articulation of the two consecutive vowels is therefore more prominent and longer than that of the single short vowel in the root. The other possible reason for stress on the suffix *rùðk* is that in Dholuo stress falls on the last heavy syllable as seen in words of more than two syllables in Section 6.2 and 6.3 below.

6.1.4 Bisyllabic compound words

Certain compound words in Dholuo form bisyllabic words. They are divided into three categories. The first category has the structures CVC-¹CV:C and VC-¹CV:C as seen in (26).

- (26) ìc-¹wà:ŋ ‘annoyance/ anger’
wíc-¹wì:l ‘forgetfulness’

In all the words in (26), the first syllable is not stressed and has a short vowel, which means that it is light. The second syllable has a long vowel and it is stressed. It is heavy.

The second category has compound words with the structure CVC-¹CVVC as seen in (27) below.

- (27) wíc-¹kùòt ‘shame / embarrassment’
kòn-¹çìèl ‘half part’

Stress is on the second syllable in this group of compound words. This syllable has a sequence of two vowels and it is heavy. The first syllable is not stressed, it is light since it has short vowel.

The third category has a compound word with the structure CVVC-¹CV:C. Just as example (24), it shows that if two syllables are heavy, it is the last heavy syllable that is stressed. It is the second syllable that is stressed, as seen in (28).

- (28) wùòk-¹tì:c ‘Monday’

6.1.5 Reduplicatives in bisyllabic words

There are reduplicatives that are bisyllabic. In this set, the sounds that appear in the first syllable are repeated in the second, but with vowel lengthening. Consider the examples in (29).

- (29) Íŋ-¹lí:ŋ ‘secretly’
tín-¹tí:n ‘small in body, thin’

There is vowel lengthening in the last syllable because that is where stress is located. It is noted that this is a closed syllable occurring in word final position.

6.2 Trisyllabic words

Trisyllabic words also display a unique stress pattern. A majority of trisyllabic words have stress on the second last or penultimate syllable.

6.2.1 Base forms of trisyllabic words

Words in this group have stress on syllables with either V: or CV: structure. The examples are given in (30) below.

- (30) à-ò:-rà ‘river’ à-’hí:-ná ‘very much’
 jà-’sá:-’jé ‘God’

The words in example (30) have the structures V-’V:-CV, V-’CV:-CV and CV-’CV:-CV. In all these words, the vowels in the penultimate syllable are lengthened. Stress is also on the penultimate syllable. The penultimate syllable is therefore heavy unlike the first and last syllables.

In constructing metrical grids for words in this category, the rule of Quantity Sensitivity is first applied. Since there is only one syllable with two successive ‘Xs’ or moras on Row 0, a grid mark is placed over it on the foot row. This means that it is the only syllable that is heavy and stressed.

Another set of trisyllabic words has stress on its final syllable. In these words, the first two syllables are light while the last one is heavy. Consider the examples in (31).

- (31) kà-là-’tâ:s ‘paper (loan)’
 kù-lù-’ndè:ŋ ‘beetle’

The stressed syllable has the structure CV:C because the vowel in it has been lengthened. Each of the other two syllables has a short vowel and is not stressed. The syllable with the lengthened vowel is heavy. Once again only the rule of Quantity Sensitivity is sufficient to show the stressed syllable of the words in this group.

Words with the structures V-CVC-’CV:C form another set of trisyllabic words. In this set it is the last syllable that is stressed and is heavy. The first two syllables are not stressed and they are light. Only the rule of Quantity Sensitivity is required to show that the last syllable is stressed. The example is given in (32).

- (32) ò-sìp-’tâ:l ‘hospital (loan)’

In all the words in this section stress is either on the penultimate syllable with the structure CV: or on the last syllable with the structure CV:C.

6.2.2 Prefixes in trisyllabic words

Prefixes are found in trisyllabic words as well. They are light and do not bear stress. Stress on the penultimate syllable which is part of the root of the word. The penultimate syllable has a sequence of two different vowels. Words in this category are given in example (34).

- (34) jí-’gúó-gí ‘puppies’
 à-’ðíá-mbó ‘name of a female born in the evening’

Words in this category have either the structure CV-’CVV-CV or V-’CVV-CV. Stress is detected on the only syllable that is heavy, *jí* which is a prefix meaning ‘young ones of’ is the plural form of *jà* discussed in section 6.1.2. The prefix *à* which is a noun forming prefix is also discussed in the same section.

The next three sets of trisyllabic words have the prefix *jà:-dí* which is attached to numbers to indicate the number of times. The first set of words has the structure CV:-CV-’CV:C. In the words in this category, two syllables have a lengthened vowel. The main word stress is perceived in the last syllable with a lengthened vowel, as seen in (35).

- (35) *ɲà:-d̥ɪ-b̥i:c* ‘five times’
ɲà:-d̥ɪ-¹d̥è:k ‘three times’

The second set has the structure CV:-CV-¹CVVC as seen in (36).

- (36) *ɲà:-d̥ɪ-¹ɲòèn* ‘four times’

This word, just like the words in (35), has the main word stress on the last syllable. Only Quantity Sensitivity is needed to show that the last syllable is stressed in (35) and (36).

The third set has the structure CV:-CVV-¹CVVC. The first syllable has a lengthened vowel, while the last two have a sequence of two different vowels each. The three syllables are stressable by virtue of their structure. However, stress is on the last syllable as seen in (37).

- (37) *ɲà:-d̥ɪù-¹c̥ìèl* ‘six times’

In example (37), in addition to Quantity Sensitivity, End Rule Final is needed to show that even though all the syllables are heavy, it is the last heavy one that is stressed.

The next two groups of trisyllabic words have the structures V-CVC-¹CVVC and V-CVV-¹CVVC. The last syllable is stressed and is heavy as seen in (38) below.

- (38) *ð-ɲòŋg-¹rùðk* ‘chameleon’
ð-d̥iè-¹c̥ìèŋ ‘day time’

They both have stress on the last syllable. Whereas only Quantity Sensitivity is needed to show this for the word *ð-ɲòŋg-¹rùðk*, both Quantity Sensitivity and End Rule Final are required for the word *ð-d̥iè-¹c̥ìèŋ*.

6.2.3 Suffixes in trisyllabic words

According to Stafford (1967), the singular imperative of most verbs consists of a shortened form of the verb. The plurals of these singular imperatives is formed by the addition of the suffix *-u:ru*. When this suffix is attached to these verbs, trisyllabic words are formed, as seen in (39).

- | | | | | |
|------|----------------|---------------------|------------------------------|---------|
| (39) | infinitive | singular imperative | plural imperative | meaning |
| | <i>'bè:-dò</i> | <i>bé:d</i> | <i>bè-¹dú:-rú</i> | ‘sit’ |
| | <i>'gò:-jò</i> | <i>'gó</i> | <i>'gò-¹û:-rú</i> | ‘bit’ |

Resyllabification takes place when the suffix *-u:ru* is added to the singular imperative. The singular imperative is monosyllabic and has either a CV:C or CVVC structure. The corresponding plural imperative has either the structure CV-¹CV:-CV or CVV-¹CV:-CV, respectively. The plural imperatives with the structures CV-¹CV:-CV or CV-¹V:-CV are in (39), while those with the structures CVV-¹CV:-CV or CVV-¹V:-CV are in (40). It is noted that stress is on the suffix and not on the root. Their stress pattern is shown in (40).

- (40) *kùà-¹û:-rú* ‘please ask for/ borrow (plural)’
c̥iè-¹lû:-rú ‘please fry’

The next class of words with suffixes consists of a reflexive form of verbs with the structure CV-CV-¹CV. These are trisyllabic words formed by the addition of the suffix *-rɛ* to infinitive verbs. Some of these infinitive verbs are seen in examples (4), (5) and (7). The addition of the suffix *-rɛ* gives rise to a reflexive form of the verb. The infinitive verb, which is bisyllabic, has stress on the initial syllable. The stressed syllable has either a lengthened vowel or a sequence of two different vowels. In cases

where the infinitive form has a lengthened vowel, its reflexive form is articulated with a short vowel, while in cases where the infinitive form has a sequence of two different vowels, the vowel sequence in its reflexive form is articulated rapidly. In other words, the articulation of the CVV sequence in the infinitive verb is more prolonged than in its reflexive form. The examples of these verbs are given in (41) and (42).

- (41) ɲà-dò-^lré 'get to cut (reflexive)'
pà-kò-^lré 'to praise (reflexive)'

All the syllables in the words in (41) are light. The last syllable receives the main word stress. On the mora row, a grid mark is placed over all syllables. End Rule Final, which places a grid mark on the extreme right of the word, is applied on the foot row to show the head of the foot. The rule of Quantity Sensitivity is not required.

The reflexive form of verbs with the structure CVV-CV-^lCV is seen in the examples in (42). Here End Rule Final overrides Quantity Sensitivity because it is the last syllable that is perceived to be stressed probably because of the difference it has in pitch from the rest of the syllables. As already stated the first syllable is articulated rapidly.

- (42) mòò-mò-^lré 'to rush unexpectedly'
rùà-kò-^lré 'to dress (reflexive)'

6.2.4 Reduplicatives in trisyllabic words

The second syllable of each of these reduplicatives is repeated. The repetition gives the word a different shade of meaning as seen in (43). Stress is on the last syllable.

- (43) má-^ljò:m 'soft, fine, smooth' má-jòm-^ljò:m 'soft one'
má-^ltí:n 'young/ small' má-tín-^ltí:n 'small proportions'

6.2.5 Miscellaneous trisyllabic words

The words placed under the miscellaneous group have the structure CV-CVC-CV. All the three syllables have a short vowel each. The metrical structure or stress pattern of these words is given in (44). Each word seems to have its own stress pattern. The pattern perceived is not related to syllable weight. In the word 'θò-mòr-nì, the last syllable behaves like the suffix 'nì' in example (23). The last syllable is not stressed. The second syllable is also not heard as prominently as the first one. The first syllable is articulated with more loudness than the last two. In this word it is noted that all the syllables have a low tone. Consider example (44).

- (44) 'θò-mòr-nì 'safari ants'
ò-^lcûn-nó 'a kind of ant'

The word ò-^lcûn-nó has a low tone on the first syllable while the second and third have a falling and high tone, respectively. The first syllable is not prominent when compared with the last two because it bears a low tone. The second syllable is perceived to be more prominent than the first because of its changing pitch unlike that of the last syllable.

6.3 Polysyllabic words

The majority of words in Dholuo are either monosyllabic, bisyllabic or trisyllabic. Polysyllabic words of four or more syllables are not very many, when compared with monosyllabic, bisyllabic and trisyllabic words.

6.3.1 Basic form of polysyllabic words

Words in (45) have either the structure V-CV-'CV:-CV, CV-CV-'V:-CV, V-CV-CV-'CV:-CV and CV-CV-'CVV-CV. Their stress pattern is given below.

- | | | | | |
|------|------------------|-------------|-----------------|------------|
| (45) | à-bí-'rír:-jók | 'seven' | mà-là-'ír:-kà | 'angel' |
| | è-rò-ká-'má:-nók | 'thank you' | tà-ndà-'wùdè-jà | 'abundant' |

What is common to all these words is that stress is on the penultimate syllable in all of them. The penultimate syllable has either the structure CV: or CVV. The rule of Quantity Sensitivity is sufficient to show that the heavy syllable of each word is stressed.

The structure V-CV-CV-'CV:C is seen in the word in (46). Stress is on the last syllable, which has a lengthened vowel. The first three syllables have a short vowel. The last syllable is heavy, while all the remaining are light.

- | | | |
|------|--------------|------------------------------|
| (46) | à-bá-rá-mà:c | 'wayward/ without any order' |
|------|--------------|------------------------------|

6.3.2 Prefixes in polysyllabic words

There are polysyllabic words in Dholuo that are formed by the addition of the prefix *jà:-dír-* which is used to indicate the number of times. As already seen in examples (35), (36) and (37), this prefix is normally added to numbers. The examples are given in (47).

- | | | |
|------|----------------------|---------------|
| (47) | jà:-dír-'rír:-jók | 'twice' |
| | jà:-dír-'bí-rír:-jók | 'seven times' |

Two syllables in each of the words in (47) have a lengthened vowel. This means that those two syllables are stressable. Stress falls on the last syllable with a lengthened vowel. Therefore, to establish the stress pattern of this word both Quantity Sensitivity and End Rule Final have to apply.

The prefix *kà* in Dholuo is used to form place names. For example, the home of a man named *ngálá* would be *kàngálá*. Polysyllabic words with this unstressed prefixes are in (48). They have the following syllable structures CV- CV-'CV:-CV, CV-CV-'CVV-CV and CV-CVC-'CV:-CV

- | | | |
|------|-----------------|-------------------|
| (48) | kà-mà-'gà:-mbò | 'name of a place' |
| | kà-rà-'cùdè-jò | 'name of a place' |
| | kà-jàm-'ká:-gók | 'name of a place' |

In this group of words, all the syllables are light except the penultimate syllable which is stressed and has a lengthened vowel. The polysyllabic word *jà-sà-'múó-jók* 'lazy person', which contains the prefix '*jà-*' has the same stress pattern as the words in (48). Only Quantity Sensitivity applies.

6.3.3 Suffixes in polysyllabic words

The plural formation suffix *-jé* exists in some polysyllabic words. Like most suffixes it is not stressed. This suffix is found in the word in (49). It is the penultimate syllable that is stressed in this word.

- | | | |
|------|-----------------|---------|
| (49) | ǎá-'rá-'ngè:-jé | 'gates' |
|------|-----------------|---------|

6.3.4 Polysyllabic words with both prefixes and suffixes

The two words in this section with both suffixes and prefixes are *ǎ-jà-wǎ-'ré* 'good morning' and *ǎ-jí-mǎ-'ré* 'good evening'. They have the structure CV-CV-CV-'CV which is almost similar to the one found in the example in (41) above except for the prefix. They also have a stress pattern similar to the one found in the examples in (41). Example (50) shows that the base form of each of the two words is bisyllabic. The prefix *ǎ-* or *ǎ-* and suffix *-ré* or *-ré* are added to them to produce related word

forms that differ slightly in meaning. Some of the meanings given in (50) are drawn from Odaga (2005). (n) refers to the term noun while (v) refers to the term verb.

- | | | | | | |
|---------|------------------------------|----------------------|----|--|--|
| (50) a. | ¹ jà:wò (v) | ‘open, unlock, free’ | b. | ¹ ji:mò (v) | ‘obstruct view, block, cover up, hide’ |
| | jà:wò- ¹ ré (v) | ‘open’ | | | |
| | jà:wò- ¹ ré (n) | ‘dawn’ | | ¹ ji:mò- ¹ ré (v) | ‘obscure, not very clear’ |
| | ò-jà:wò- ¹ ré (n) | ‘good morning’ | | ò- ¹ ji:mò- ¹ ré (n) | ‘good evening’ |

6.3.5 Reduplicatives in polysyllabic words

Reduplicative in Dholuo are also found among polysyllabic words. In some of these words, a portion of or a whole word is repeated. The examples given in (51) have either the CV-CV-¹CV:-CV or CV-CV-CV-¹CV:-CV.

- | | | | | |
|------|--|-------------|--|--------------------|
| (51) | mì- ¹ lí- ¹ mì:-lí | ‘sweet’ | ò-bó- ¹ ɲgó- ¹ bò:- ¹ ɲgò | ‘a type of lizard’ |
| | kà-wì-rà- ¹ wí:-rá | ‘dizziness’ | gá-lá- ¹ gá:-lá | ‘awkward’ |

Stress is on the penultimate syllable which also has a lengthened vowel. It is this penultimate syllable that is heavy. All the other syllables have a short vowel and are therefore light. Unlike the words in (51), the word ò-tíèy-ò-¹tíèy ‘spider’ has the structure V-CVVC-V-¹CVVC which means that it has two syllables with a vowel-vowel sequence. It is the last syllable with this kind of sequence that is stressed. Whereas Quantity Sensitivity is enough to show the stressed syllable of the words in (51), both Quantity Sensitivity and End Rule Final are required to show the stressed syllable of the word ò-tíèy-ò-¹tíèy.

6.4 Vowel lengthening at sentence level (root lengthening)

At the beginning of this paper, it is indicated that the research is based on the lexical level. Accordingly, the sentence is included only as far as it helps to show the positions where the underlying short vowels are realised as lengthened vowels on the surface.

In some Dholuo sentences, a vowel is doubled in the penultimate position of a word that is in utterance-final position. The examples of utterances with such doubled vowels are given in (52) below. In (52) the syllable structure of the utterance káw#á-¹gú:-¹cá ‘take my pot’ is represented diagrammatically. The symbol # represents word boundary while ## represents sentence boundary.

- | | | | | | | | | | | | | |
|------|-----|---|---|---|---|---|----|----------------|----|----------------|---|---------------------------------------|
| (52) | (a) | phonemic tier | k | á | w | # | á | ¹ g | ú: | ¹ c | á | ‘take my pot’ |
| | | Row 0 (mora row) | | x | | x | xx | x | | | | |
| | | Row 1 (foot row) | | | | | x | | | | | |
| | | Row 2 (word row) | | | | | x | | | | | |
| | | Row 3 (sentence row) | | | | | x | | | | | |
| | (b) | ##á- ¹ gú:- ¹ cá # nò:- ¹ tò:-ré## | | | | | | | | | | ‘my pot broke’ |
| | (c) | ##jò:- ¹ tó-ré#né#kè:-ndò # | | | | | | | | | | ‘Kendo broke it by mistake yesterday’ |
| | (d) | ##kè-ndò # jò:- bí:- ¹ ró## | | | | | | | | | | ‘Kendo came yesterday’ |

In (52) (a) and (b) the focus is on the word á-¹gú:-¹cá. Its penultimate syllable is realised with a long vowel because the word occurs in utterance final position. When the word is non-final, as in (52) (b), the vowel remains short. The word nò:-¹tò:-ré in (52) (b) occurs in utterance final position, consequently its penultimate syllable has a lengthened vowel and this is where sentence level stress is detected. The same word has a long vowel in the antepenultimate position as a result of deletion followed by compensatory lengthening (see Oduor 2002). Compensatory lengthening doubling in this word does not determine or affect the location of stress. In (52) (c), the same word is in a non-final position. In this example, the word jò:-¹tó-ré does not occur at the end of the utterance hence the vowel in its penultimate position remains short and does not bear sentence level stress. In the same example, the penultimate syllable of the word kè:-ndò is stressed in utterance final position, hence the lengthening of the vowel in its penultimate syllable. In (52) (d), stress is not on the word kè-ndò and

therefore all its vowels are short. All these words with stress on the penultimate syllable end in a vowel.

The last vowel of a word could also be lengthened in utterance final position as seen in (53).

- | | | | |
|------|-----|-----------------------|---|
| (53) | (a) | ## lúé-té # ɔ-tù:r ## | ‘his/ her hand is broken (impersonal construction)’ |
| | (b) | ## ɔ-tùr # lúé-té ## | ‘his/ her hand is broken (agentive construction)’ |
| | (c) | ## àlót # ɔwàŋ ## | ‘the vegetables are burnt’ |
| | (d) | ## ɔwàŋɔ # àlót:t ## | ‘he has burnt the vegetables’ |

In (53) (a), (c) and (d), the syllable that is stressed is on the extreme right. This means that it is the last vowel that is stressed. It is lengthened because it occurs in a word that is in utterance final position. In all the other positions, in the utterance, the words do not have lengthened vowels. The difference between the words with the lengthened vowels in (52) and (53) is that those in (52) are in words ending in a vowel, while those in (53) are in words ending in a consonant.

Stress in Dholuo is detected towards the end of an utterance, through vowel lengthening. Stress therefore has the effect of lengthening Dholuo vowels. This is similar to Remijsen’s (2001) observation that stressed syllables in Ma’ya are longer than unstressed syllables. For Dholuo, in a vowel final word stress is detected on the open penultimate syllable and in a consonant final word, i.e. a word with a closed syllable as the last syllable, stress is detected on this syllable. This stress pattern is what Okombo (1982) calls root lengthening. He states that, in a word ending in a consonant, in utterance final position, it is the last vowel that undergoes root-lengthening while in a word ending in a vowel, in utterance final position, root-lengthening takes place on the penultimate vowel.

The word *lúé-té* which means ‘his/ her hand’, in examples (53) (a) and (b), shows that the structure of its first syllable does not change whether the word is in utterance final position or any other position. This means that words with syllables that have a sequence of two different vowels underlyingly have the same sequence on the surface. The last word of (53) (b) has stress on its first syllable. This syllable is inherently heavy and also occurs in a position that attracts stress in Dholuo. It is articulated more distinctly in (53) (b) than in (53) (a) where its articulation is quite rapid because it is not the final word of the sentence.

7 Summary and conclusion

From the outline of the stress pattern of Dholuo, it is clear that there is a distinct relationship between syllable weight and stress. Stress seems to coincide with syllable weight. Syllables that are stressed in almost all the examples are the syllables that are heavy. Therefore, a heavy syllable in Dholuo is a syllable that is stressed in almost all cases. For a syllable in Dholuo to be heavy, the nucleus must be occupied by a lengthened vowel hence the representations CV:, V:, CV:C and V:C. The onset and coda are irrelevant. Stress therefore has the effect of lengthening vowels in Dholuo, thus transforming syllables from being light to being heavy. This is evidence to show that Dholuo is a type B language according to Katamba’s (1989) classification.

It is also observed in the discussions that stress in Dholuo is detected in the penultimate position in a vowel final word and in the last syllable in a consonant final word. For a syllable to have a lengthened vowel, it must be in a position where stress is detected. Duration in Dholuo is therefore an effective cue of stress. All the heavy syllables discussed in the preceding paragraph are realised as light syllables in all the other positions where they are not stressed. These heavy syllables are therefore not underlyingly but are derived or acquired on the surface when they occur in the positions where stress is detected in Dholuo.

There are Dholuo syllables that appear to be inherently heavy. These heavy syllables have the structure CVV and CVVC. The structures of these syllables make them stressable. They differ from the syllables with lengthened vowels in that their nucleus is made up of a sequence of two different vowels. Most of these heavy syllables appear to be at the right position, that is, where stress is easily detected in Dholuo.

There are words with two syllables of the same weight, i.e., CVVC-CVVC structure. In such situations, it is the second heavy syllable that is stressed. In such words, the first heavy one is

articulated more rapidly. This means that in Dholuo, it is always the last heavy syllable in a word that is stressed. Dholuo then is a Quantity Sensitive language.

A light syllable in Dholuo has a short vowel and is unstressed in almost all cases. In some Dholuo words with light syllables only, the syllable that seems to be stressed is articulated with greater intensity than the others or has a different pitch from the rest of the syllables in the word. This observation is not solid enough and may require further investigation. Very few words have a sequence of light syllables only.

The discussions of stress in Dholuo words, in sections 6.1, 6.2 and 6.3 above, also show that stress is on the penultimate syllable in words ending in a vowel and on the last syllable in words ending in a consonant. Therefore, the stress pattern of sentences in Dholuo is similar to that of words said in isolation. It is concluded that Dholuo is a fixed stress language with stress playing the role of demarcation. Stress marks the end of an utterance. This paper therefore supports Dimmendaal (2012: 1) assertion that the assumption ‘that metrical structures do no[t] play a role in tonal languages’ is not correct. There is absolutely no evidence in the data to show that stress is phonemic. Dholuo therefore remains a tone language because tone is phonemic.

Abbreviations and Symbols

C	consonant
V	vowel
(v)	verb
(n)	noun
˥	downstepped high tone
-	syllable boundary
˨	high tone
˩	low tone
˧	falling tone
ˈ	stress mark
#	word boundary
##	sentence boundary

References

- Allen, W. Sidney. 1973. *Accent and Rhythm. Prosodic Features of Latin and Greek: A Study in Theory and Reconstruction* (Cambridge Studies in Linguistics, 12). Cambridge: Cambridge University Press.
- Carr, Philip. 1999. *English Phonetics and Phonology: An Introduction*. Oxford: Blackwell Publishing.
- Dimmendaal, Gerrit J. 2012. Metrical Structures: A Neglected Property of Nilotic (and other African Language Families), In Osamu Hieda (ed.), *Challenges in Nilotic Linguistics and More, Phonology, Morphology and Syntax* (Studies in Nilotic Linguistics, 5), pp.1–26. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.
- Goldsmith, John A. 1990. *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.
- Hyman, Larry M. 1975. *Phonology: Theory and Analysis*. New York: Holt Rhinehart and Winston.
- Hyman, Larry M. 2003. *A Theory of Phonological Weight*. Stanford: CSLI Publications.
- Katamba, Francis. 1989. *An Introduction to Phonology*. London: Longman.
- Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.
- Laver, John. 2000. *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge: MIT Press.
- Newman, Paul. 1972. Syllable weight as a phonological variable. *Studies in African Linguistics* 3: 301–323.

- Odaga, Asenath B. 2005. *Dholuo-English Dictionary*. Kisumu: Lake Publishers and Enterprises.
- Odhiambo, Jacqueline H. A. 1981. *Dholuo phonology: a study of the major vowel processes*. MA dissertation, University of Nairobi.
- Oduol, Jacqueline H. A. 1990. *Dholuo dialects: synchronic state and some historical inferences*. PhD dissertation, University of Nairobi.
- Oduor, Jane A. N. 2002. *A study of syllable weight and its effects on Dholuo phonology*. PhD dissertation, University of Nairobi.
- Okombo, Duncan O. 1982. *Dholuo Morphophonemics in a Generative Framework*. Berlin: Reimer.
- Omondi, Lucia N. 1982. *The Major Syntactic Structures of Dholuo*. Berlin: Reimer.
- Remijsen, Bert. 2001. *Word Prosodic Systems of Raja Ampat Languages*. Utrecht: Landelijke Onderzoekschool Taalwetenschap (LOT).
- Stafford, Roy L. 1967. *An Elementary Luo Grammar with Vocabularies*. Nairobi: Oxford University Press.
- Trubetzkoy, Nikolai S. 1939. *Grundzüge der Phonologie, Travaux du Cercle Linguistique de Prague* no. 7, Prague: Cercle Linguistique de Prague; reprinted 1958, Göttingen: Vandenhoeck and Ruprecht; translated in 1949 by J. Cantineau as *Principes de Phonologie*, Paris: Librairie Klincksieck; translated in 1966 by C. A. M. Baltaxe as *Principles of Phonology*, Berkeley: University of California Press.

EASTERN NILOTIC VOWEL HARMONY AND OPTIMALITY THEORY: WHAT CAN THE OPTIMAL SYSTEM LOOK LIKE?¹

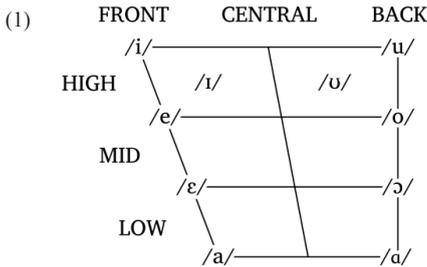
Diane Lesley-Neuman

1 Introduction: Optimality and vowel harmony in the Eastern Nilotic sub-family

Optimality theoretic analyses of Eastern Nilotic (EN) vowel harmony originate from two distinct research traditions. The first comprises theoreticians from outside of the sub-family specialization who seek to develop proposals with universal validity, while linguists respectful of the Africanist descriptivist tradition comprise the second. Some proposals have emphasized the systematicity of spreading processes over the irregularities generally found in surface forms of vowel harmony systems. These irregularities, products of historical change, are deemed here to be part of the grammars of their respective languages and to constitute phenomena to be accounted for theoretically. This minimally entails explanation for the existence of two or more harmony processes within a single grammar, disharmonic domains generated by consonant-generated features and dissimilation rules, variation in harmony set membership according to harmony process type, members of the harmony set doubling as neutral vowels in specific morphemic environments while the morphemes themselves also vary in their status as neutral or active. Contrast is offered between those proposals emphasizing Richness of the Base with those that argue for underspecification and account for surface form irregularities within their constraint-based hierarchies. Models that explain the phonetics-phonology and morphology-phonology interfaces are considered to be the future direction in the theoretical analysis of EN phonology, and should model the historical changes generating phonological phenomena.

As noted by Hall and Creider (1998), Nilotic languages of the Eastern Nilotic (EN) subfamily historically have had 5 x 2 [ATR] harmony systems, which subsequently underwent diachronic change. A 5 x 2 harmony system, as shown in (1), is one in which the five vowels maximally spread within the vowel space and representing the three basic height distinctions of high, mid and low occur in two variants: a ‘tense’ variety bearing a positive value for the feature Advanced Tongue Root [ATR], and a ‘lax’ variety, which bears a negative value for the same. The [+ATR] vowels are /i/, /e/, /a/, /u/, /o/ while the [-ATR] variants are /ɪ/, /ɛ/, /ɑ/, /ʊ/, and /ɔ/. Ideally, the vowels within a given word conform to the same [ATR] specification, and thus ‘harmonize’. Languages such as Bari have maintained this 10-vowel system, while other languages within the EN sub-family have been altered. One common change has been the loss of the [+ATR] variant of the low vowel through vowel mergers, thus creating a nine-vowel system. The mergers are attributed to the minimization of acoustic differences between the [+ATR] vowel and the [-ATR] vowel one degree of height above it from the secondary rise of the highest part of the tongue during backing or fronting processes. Another common change is the enhancement of the [±ATR] distinction through differentiation in voice quality. Especially frequent is the morphologization of the system, where the morphological context determines the participation in harmony by any given vowel. These factors along with local phonetic interactions and phonologized processes produce a more complex picture of EN vowel harmony overall.

¹ Special thanks to David Rood, Bill Raymond, Ricardo Bermudez-Otero, Dick Demers, Gerrit Dimmendaal, Manuela Noske and Andy Cowell who supervised and gave advice for an early version of this work; to Adam Baker, Lise Menn, John Goldsmith, Dave Odden, Grover Hudson, San Duanmu, Peter Avery and the audience at the Michigan State Linguistics Circle for feedback that improved the argumentation, and to Helga Schroeder and an anonymous reviewer who assisted me in improving the final draft. All other errors and omissions are my own.



In this paper, Karimojong, a major language in the Teso-Turkana cluster, will be taken as a prototypical case of Eastern Nilotic [ATR] harmony. A detailed analysis of essential [ATR] harmony data will be followed by theoretical discussion within an optimality-theoretic framework, including the evaluation of competing proposals that include data from the subfamily and outside it.

1.1 Principal harmony domains

Karimojong possesses a nine-vowel harmony system. According to Dimmendaal (2002), the [+ATR] variant of the low vowel has minimally undergone two vowel mergers: with [ɔ] and with [ɑ]. It remains part of the vowel inventory for a small number of lexical items. It is also proposed to be the phonetic result of placing [ɑ] in [+ATR] environments (Lesley-Neuman 2007, 2012).

Each of the two major [ATR] harmony domains is triggered by vowels in privileged positions. Bi-directional root-controlled harmony exhaustively spreads [±ATR] specifications. Dimmendaal (2002: 178) characterizes it as basic to Eastern Nilotic systems, while Lesley-Neuman (2012) claims it to be historically the oldest, effected on the first of three morpho-phonological levels. Simple cases appear in (2). The [ATR] specification of each root vowel spreads leftward to the alternating vowel of infinitive prefix /akɪ-/. Vowels with matching [ATR] specifications are underlined and bolded. Example (2f) shows the low vowel [ɑ] in the root patterning as a [-ATR] harmony trigger, and not serving exclusively as a passive recipient of an [ATR] specification as some models propose.

	[-ATR]		[+ATR]	
(2) a.	akɪ-d <u>ɪ</u> ŋ	‘to pinch’	g. ákɪ-d <u>ó</u> ŋ	‘to castrate’
b.	akɪ-b <u>ó</u> k	‘to swing south’	h. ákɪ-j <u>ú</u> k	‘to smear’
c.	ákɪ-n <u>ɔ́</u> k	‘to light fire’	i. ákɪ-n <u>ó</u> k	‘to go near’
d.	ákɪ-b <u>é</u>	‘to agree’	j. ákɪ-b <u>é</u> lɛ	‘to change’
e.	ákɪ- <u>ír</u>	‘to test’	k. ákɪ-l <u>ík</u>	‘to swallow’
f.	ákɪ-c <u>á</u> p	‘to weed’		

Exhaustive root-controlled spreading is illustrated with the derived infinitive forms in (3). Root [ATR] specifications are supplied bidirectionally to the alternating vowels of recessive morphemes. Harmonizing vowels are marked in bold. Vowels neutral in these examples are active in other environments, such as the vowel [i] from the applicative suffix [-Akin] in examples (3a,c). It remains unchanged regardless of the harmony domain in which it is placed. It can be deduced from the recessive vowel of the infinitive prefix for all forms in (2) and the feature profiles of the root vowels in (2e) and (2k) that across the language, this vowel can otherwise serve as a harmony trigger or a recessive recipient of [ATR] feature information.

(3) Bidirectional root controlled [±ATR] domain

- | | | | |
|----|---|----|--|
| a. | ákɪ-t <u>ɔ́</u> -m <u>ɔ́</u> kɪ- <u>á</u> n- <u>á</u> kin | b. | ákɪ-c <u>ó</u> m <u>ó</u> c <u>ó</u> m <u>ó</u> - <u>ù</u> n |
| | INF-CAUS-handle firmly-FREQ-APPL | | INF-pierce repeatedly-VEN |
| | ‘to cause to frequently handle firmly (for)’ | | ‘to pierce repeatedly (this way)’ |

- c. **áki-zi - dón-ón - ókin**
INF-CAUS-castrate-FREQ-APPL
'to cause to castrate (for)'
- d. **áki-riṃriṃí-ùn**
INF-go around repeatedly-VEN
'to go around repeatedly (this way)'

The second major harmony domain comprises the exhaustive spreading of a suffix-controlled [+ATR] feature. The trigger is an active marker of tense, mood and aspect (TMA), the rightmost affix of inflected verb forms. In (4), past progressive voice forms of the [-ATR] verb *áki-tó-jók* 'to cause to be good' vary in person, number, mood and voice. The [+ATR] feature spreads leftward through the causative prefix [-to-], creating [+ATR] domains marked bold type. It dominates over all other active domains, including localized ones discussed below. The pronominal prefixes, opaque to this spreading process, are among the class of neutral morphemes presented in §1.3.

- (4) Dominant suffix-controlled [+ATR] domain
Infinitive: *áki-tó-jók* 'to cause to be good'
- a. **è-tò-jók-í**
1s-CAUS-be good- IND.ACT.PST.PROG.A.1s
'I was causing to be good.'
- b. **ì-tò-jók-etè**
2p-CAUS-be good- IND.ACT.PST.PROG.A.2p
'You were causing to be good.'
- c. **εke-to-jok-jó**
1s-CAUS-be good-IND.PASS.PST.PROG.A.1s
'I was being caused to be good.'
- d. **εke-to-jok-etei**
1s-CAUS-be good-
NARR.PASS.PST.PROG.A.1s
'...and I was being caused to be good.'

1.2 Local phenomena

This section will discuss three local effects impacting [ATR] specifications. The first is a [-ATR] spreading process generated by the [+ATR] itive suffix /-or/, surfacing as [-ɔr]. A small [-ATR] domain is created by the phonologization of the anticipatory retraction of the tongue for the pronunciation of the rhotic. Vowels in morphemes preceding and following the affix adopt the feature, as shown in (5).

- | (5) | <u>infinitive</u> | <u>itive</u> | <u>freq./itive</u> | <u>applicative</u> | <u>freq./applicative</u> | <u>gloss</u> |
|-----|-------------------|--------------|--------------------|--------------------|--------------------------|-------------------|
| a. | áki-do | áki-dó-òr | áki-dó-ón-òr | áki-dó-(ó)kin | áki-do-ón-ókin | 'produce a child' |
| b. | á-dóre | á-dó-òr | á-dó-ón-òr | á-do-ókin | á-do-ón-ókin | 'fall' |
| c. | á-murò | á-muró-òr | á-muró-ón-òr | á-múr-ókin | á-múró-ón-ókin | 'sacrifice' |

This [-ATR] domain is overridden by the dominant [+ATR] domain triggered by the final TMA marker as shown in (6).

- (6) Infinitive: *-dónódón-ó-ɔr* 'repeatedly castrate (this way)'
- a. **é-dónódón-ó-or-ì** - 'I (will) repeatedly castrate'
- b. **ì-doṇodoṇ-o-or-ì** - 'You (s.) are/were repeatedly castrating.'
- c. **ε-doṇodoṇ-o-or-ete** - 'They are/were repeatedly castrating.'

The second local effect involves consonant-vowel interactions upon suffixation of the frequentive-ventive complex [-*ùnùn*]. Lesley-Neuman (2012) proposes this suffix to be at an intermediate stage of phonological incorporation. When a [-ATR] verb root ends in a vowel, its [ATR] specification spreads to the vowels of this suffix, as shown in (7a,b). This is likewise valid in cases where the vowel extension following the reduplication of the root bears the [-ATR] feature, as exemplified in (7c). The presence of a voiced consonant, particularly a nasal, tends to ensure that the suffix retains its [+ATR] specification as in examples (8a,c). This is attributed to the enlargement of the pharyngeal cavity by tongue root advancement in the effort to maintain voicing (J. R. Westbury, 1977, 1983, personal

communication). Stem-enlarging vowels added after reduplication, the specifications of which are tied to the verb root, appear to erase this effect as shown in (7c) and in (8a,c). In addition, cases exist in which the erosion of the final [n] of the frequentive or the ventive suffixes induces the laxing and even de-voicing of the preceding vowel. Such a case is shown in (7e). Nevertheless, as shown in (7d), a [-ATR] verb root ending in a voiceless consonant can still retain the [+ATR] feature in the suffix.

(7)	<u>infinitive</u>	<u>freq./ventive</u>	<u>gloss</u>
a.	ákí-bú	á-bú-ún-òn	‘return’
b.	ákí-dó	ákí-dó	‘snap one’s fingers’
c.	ákí-tópótópó	ákí-tòpótópó-ún-òn	‘follow repeatedly’
d.	ákí-jók	á-jók-ún-ùn	‘become good’
e.	ádàùn	àdàù	‘he finished’, ‘it is enough’

(8)	<u>infinitive</u>	<u>freq./ventive</u>	<u>reduplicated form</u>	<u>gloss</u>
a.	ákí-dóŋ	á-dóŋ-ún-ùn	á-dóŋ-ún-ùn	‘pinch’
b.	ákí-sób	ákí-sób-ún-ùn	ákí-sóbósóbú-ún-òn	‘create’
c.	ákí-cóm	ákí-cóm-ún-ùn	ákí-cómócómó-ún-òn	‘pierce’

The third local effect is a dissimilation rule preserving featural information in a linguistic environment historically prone to vowel mergers. When [-ATR] vowels [i,ɛ,ɔ,u] occur in stem-final position before suffixes beginning with [a], followed by consonants that are [+cor], their [ATR] values change to those of the corresponding [+ATR] vowels [i,e,o,u]. Exceptions to this rule occur when the [-an] suffix is followed by the applicative suffix [-*akin*]. The effects of this dissimilation rule can be seen in (9), in which the affected vowels are shown underlined and in bold type.

(9)	<u>infinitive</u>	<u>itive</u>	<u>freq./itive</u>	<u>gloss</u>
a.	ákí-ré	á-ré-àr	àr <u>é</u> -án-àr	‘put in a line’
b.	ákí-dó	á-dó-àr	àr <u>ó</u> -án-àr	‘snap one’s fingers’
c.	ákí-bú	á-bú-àr	àr <u>ú</u> -án-àr	‘return’
d.	ákí-mókí	á-mókí-àr	àr okí-án-àr	‘agree’

1.3 Morpheme status

Harmony processes in Karimojong are dependent upon the active or inactive status of each individual morpheme. The active harmony domains presented above correspond to the first and second morpho-phonological levels respectively by their order of presentation. Inactive morphemes are affixed on the third. They are said to be *transparent* in that they do not block spreading when encountered in the midst of an active [ATR] domain. A number of the Level 3 morphemes presented in this section also double as active Level 1 and Level 2 morphemes. This is consistent with the model of gradual historical change, presented in Lesley-Neuman (2012), which also provides a guide to affixes in the language by morpho-phonological level.

1.3.1 Non-alternating pronominal prefixes

The first type of neutral morpheme is the pronominal prefix, seen in (10). The [-ATR] verb [-bilɪbɪl] in (10a) selects the recessive TMA marker [-azi] to express the reflexive voice in future tense. The third person plural pronominal prefix is [ɛ-]. In (10b), this pronominal prefix remains unchanged when an active and dominant TMA marker [(i)jetè] is affixed, changing all of the remaining vowels in the word to a [+ATR] specification.

(10) Non-alternating pronominal prefix

- | | | | |
|----|---|----|--|
| a. | ḙ- bilibilazí
3p-send repeatedly-ACT.REFL.FUT
‘They sent themselves repeatedly’ | b. | ḙ- bilibilijetè
3p-send repeatedly-ACT.IND.FUT
‘They will send repeatedly’ |
|----|---|----|--|

A subset of narrative mood pronominal prefixes alternates with Level 1 and Level 2 processes, manifesting an incipient process of phonological incorporation with the stem. Further examples and discussion are found in Lesley-Neuman (2012: 139–140, 158–159).

1.3.2 Non-alternating frequentives

The first inactive frequentive affix is [-eenen], considered to be an infix applied on Level 3 after Level 1 and 2 processes are completed. The examples in (11) show it retaining its [+ATR] specification despite its presence in [-ATR] domains. In example (11a) TMA marker [-tə̀] receives a [-ATR] specification from the root, despite the intercalation of [eenen]. The bidirectional domain supplies the [-ATR] specification of the causative prefix [-tə̀-]. Example (11b) shows this suffix remaining unaffected when affixed to the infinitive form. Example (11c) shows the [-ATR] specification spreading to the TMA marker [-ḙ] without interference.

(11) Frequentive infix [-éénén]

- | | <u>infinitive</u> | <u>frequentive form</u> | <u>gloss</u> |
|----|-------------------|-------------------------|--|
| a. | ákídóḡ | ε-tə̀-dóḡ-éénéné-tàè | ‘He has frequently caused to pinch.’ |
| b. | ákibú | á-bú-éénén | ‘to return frequently’ |
| c. | ákírè | à-rè-éénén-ḙ | ‘They frequently put (PST) in a line.’ |

The neutral frequentive TMA marker [-itit] appears to have arisen from the reduplication of the TMA marker [-it], presented in §1.3.3. Novelli (1985: 214–215), shows it appearing in a restricted number of verbs. It may occur in isolation or with other TMA markers. Its meaning is roughly equivalent to ‘used to (do)’ when functioning as a synthetic morpheme expressing tense in the present or past perfect. Its phonological neutrality is shown in examples (12a,b) with [-ATR] verb roots remaining unchanged. The affixation of a [+ATR] TMA marker in example (12d) triggers the changes in the specifications the root vowel and the narrative mood pronominal prefixes to [+ATR].

(12) Frequentive suffix [-itit]

- | | | | | | |
|----|---|------------|----|---|------------|
| a. | ì-dəḡ-itit
2s-pinch-ACT.IND.
FREQ.PST.PRF
‘You had frequently pinched’ | ijòḡ
2s | b. | kʔi-dəḡ-itit
SUBJ.2s-pinch-ACT.
FREQ.PST.PRF
‘that you had frequently pinched’ | ijòḡ
2s |
| c. | è-tija-ítít
1s-do-FREQ.IMPF.
‘I used to do’ | | d. | òo-dəḡ-itit-e
NAR.1s-pinch-ACT.FUT
‘...and I will frequently pinch’ | àḙòḡ
1s |

1.3.3 Neutral TMA markers

The neutral TMA marker [it] alone marks the present perfect and past perfect tenses in singular and in first person plural forms. In (13a), the [-ATR] verb [-dəḡ] ‘pinch’ does not change its [ATR] specification upon affixation of this TMA marker; the subjunctive pronominal prefix is neutral and does not alternate in the face of bidirectional root-controlled [-ATR] spreading. In (13b), phonologically active material is affixed as part of the TMA marker, triggering a [+ATR] spreading process that changes the root vowel specification to a positive [ATR] value.

(13) Neutral TMA marker [*it*]

- | | |
|---|---|
| a. kʔiki-dɔŋ-it
SUBJ.1p-pinch-ACT. PRS PRF
'that we have pinched' | b. kʔi-dɔŋ-itó
SUBJ.2p-pinch-ACT. PRS PRF
'that you have pinched' |
|---|---|

The second neutral TMA marker is [-*ere*], a Form B² marker of present and past continuous aspect which can appear as active. It is inactive in contexts where derivational suffixes intervene between the marker and the root vowel. Since these suffixes do not act as harmony blockers in any other instances other than when they are affixed with [-*ere*], the restrictions on spreading are attributed to the weakness of the TMA marker rather than inherent blocking properties of the intervening morphemes. Examples (14a-c) show inactive cases. The alternating pronominal prefix [ɔkɔ-] in (14c) retains its [-ATR] specification due to the inactivity of the [+ATR] TMA marker, while in all three examples, the root vowel specification of the verb [-*dɔŋ*] also remains [-ATR].

(14) Neutral and active forms with [-*ere*]

- | | |
|--|---|
| a. aka-dɔŋ-akin-eré
1s-pinch-DAT-B.ACT.IND. PST PROG
'I was pinching (for)' | b. ε-dɔŋ-éenén-é-eré
3s-pinch-FREQ-B.ACT.IND FUT.
'He will frequently pinch.' |
| c. ókɔ-dɔŋ-ar-eré
NAR.1s-pinch-IT-B. ACT. NARR.PST PROG
'...and I was pinching (away)' | d. aki-dɔŋ-eré
1s-pinch-B.ACT.IND.PST PROG
'I was pinching' |
| e. kʔáka-dɔŋ-eré
SUBJ.1s-pinch-B.ACT.SUBJ. PST PROG
'if I were pinching...' | f. oko-dɔŋ-eré
NARR.1s-pinch-B.ACT.NARR.PST PROG
'...and I was pinching' |

The third type of neutral TMA marker marks the aorist past. The case in (15a), shows a voiceless [+ATR] high front vowel, serving as a synthetic TMA marker of the indicative mood, passive voice and aorist past tense. It does not spread its [+ATR] specification to the root vowel, which remains [-ATR]. A null case is shown in (15b). The TMA marker for the indicative mood, aorist past tense has had its phonological substance completely eroded. The third case, presented for purposes of comparison, is that of aorist past TMA marker [ɔ̃], a recessive voiceless vowel, which bears the [ATR] specification of the root vowel under a Level 1 bidirectional spreading process.

(15) Aorist past TMA markers

- | | | | |
|--|--|-----------------------------------|---------------------------------|
| a. a-dɔŋ-ɨ
3s-pinch-IND.PASS.PST
'He was pinched.' | b. à-dɔŋ
1s-pinch-PST
'I pinched.' | c. à-dɔŋ-ɔ̃
1s-squeeze out-PST | áɛɔŋ
1s
'I squeezed out.' |
|--|--|-----------------------------------|---------------------------------|

1.4 Word formation and harmony

The model for Karimojong word formation that accounts for attested harmony patterns in surface forms requires the inclusion of head-dependent asymmetries generated by the morphological structure, hereafter M-structure, and the phonological structure, or P-structure. A phonological head for each morpheme is marked and headmarking is a determinant of the P-structure. The phonological head is connected to a Prosodic Word (PW) node that is part of what the PW word structure, which in Karimojong, can be recursive. In the verb root, the phonological head is the leftmost vowel. In each affix, it is the vowel closest to the root vowel. For suffixes, the phonological head is the vowel furthest to the left, while for prefixes it is the vowel furthest to the right. When more than one vowel in a morpheme participates in harmony, a branching from the phonological head of the morpheme to the remaining vowels in the morpheme occurs. The [ATR] harmony domain is effected with the

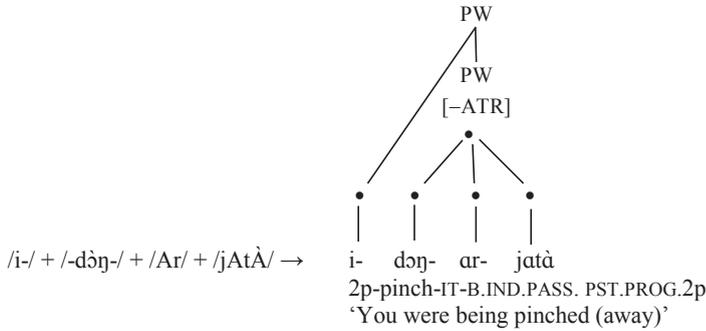
² Form B conjugations are those used when non-canonical word order is employed.

percolation of an [ATR] feature from the head vowel in a privileged position to the PW node, and its subsequent application to the phonological head in each morpheme connected to the PW node.

Nonetheless, another distinction must be drawn between the root-controlled and the suffix-controlled processes. In the case of the former, the root vowel bears an underlying specification which spreads and provides an [ATR] specification to the vowels in the affixes, which are underlyingly underspecified. This is deemed a *feature-filling* process. In the latter, the recessive vowels already bear specifications which are subsequently changed. This is referred to as a *feature-changing* operation.

An example with two instances of branching from the phonological head shows the case of the TMA marker [-jata], marking Tenses 1 and 2 of Form B, Passive Voice, shown in (16). The word [i-dɔŋ-ar-jatà] has two prosodic word nodes, one for the pronominal prefix [i-], which is affixed on Level 3 and does not undergo any vowel harmony processes, and the other for the verb stem, which does. In the latter case, the [-ATR] feature percolates from the root vowel to the PW node, which is linked to the phonological heads of the affixes to the right, which are underspecified for [ATR]. The phonological head of [-jata] has a branched connection leading to both the head vowel and the second and final vowel of the suffix, both of which have the value [a].

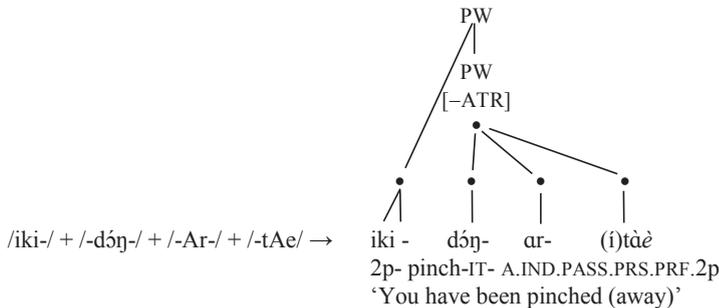
(16) Feature percolation and branching prosodic word nodes



(Novelli 1985: 293)

Where an affix contains one alternating vowel among a set of transparent ones, no branching occurs. Transparent vowels have no connection to the PW node. In (17), the suffix [-tAe] has [-A-] linked to the phonological head and the PW node. Transparent vowel [-e], shown in italics, has no such connection, and does not receive the [-ATR] specification from the root through the PW node. Epenthetic vowel [-i-] does not undergo harmony, and is excluded from determining the phonological head. Pronominal prefix [iki-] 'you', affixed on Level 3, has a separate prosodic word node through which no feature percolation occurs.

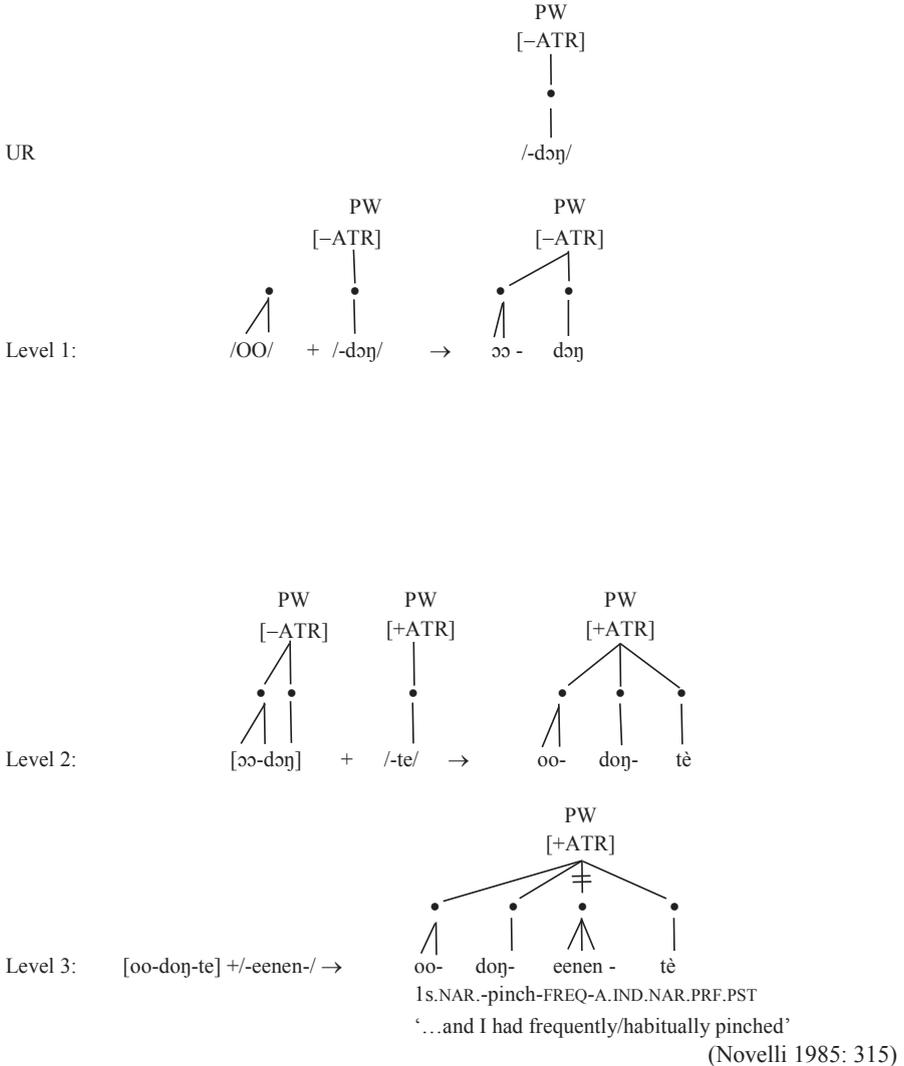
(17) P-structure with transparent vowels



(Novelli 1985: 294)

Inactive and transparent affixes are constituent parts of the prosodic word, but are linked to PW nodes separate from those of active segments. They are affixed in Level 3 processes in which no feature percolation occurs. For example, the transparent infix [-*eenen*] links to the PW node of the root, after Level 1 and Level 2 processes have applied, but is embedded in a recursive PW structure. This process and final configuration of the resulting word can be seen in (18).

(18) Suffix-controlled harmony: alternating pronominal prefix, neutral frequentive suffix



2 Optimality theory

Optimality Theory was originally proposed to address phonological *conspiracies*. A conspiracy occurs when a variety of phonological processes are effected to prevent the appearance of a prohibited form, such as an unattested syllable structure or segmental sequence. This phenomenon cannot be captured by the system of rewrite rules created by Chomsky and Halle (1968), which neglects to describe how phonological processes fit together into a whole system.

The solution proposed by Optimality Theory is an ordered ranking of filters, or violable constraints. An input form acts as a starting point, and a series of candidates for the output or optimal form is created by a system called GEN. The input is purported to be the result of the free combination of linguistic primitives without any language-particular restriction, under a concept termed Richness of the Base (ROTB). The candidates are evaluated under the ranked hierarchy of constraints under a mechanism labeled EVAL which is represented in a tableau. The candidate that best satisfies the ranked hierarchy of constraints is the output form, considered to be optimal. The ranking of constraints is language or dialect-specific, and under the original model the output form is proposed to be produced in one step, in which all of the phonological processes are purported to occur in parallel.

2.1 Pioneering optimality-theoretic proposals for addressing vowel harmony

Three proposals are presented below. The first, Optimal Domains Theory (ODT), concentrates on using the constraints to define harmony domains and categorize them in terms of their scope. The second, Featural Alignment, exploits the machinery of auto-segmental theory while addressing a phenomenon common in African languages with [ATR] harmony: the expression of a morpheme through the action of a dominant [ATR] from an eroded segment in which all other elements are absent from the feature profile. The third proposal, the AGREE family of constraints, creates harmonic domains relying on local agreement between adjacent segments with respect to a harmonizing feature. It had been originally proposed to do away with constraint-based stipulations on the directionality of feature spreading.

2.1.1 Optimal domains (Cole and Kisseberth, 1994)

Optimal Domains Theory (ODT) was proposed as an alternative due to dissatisfaction with Autosegmental Theory (AT). It was felt at the time that AT did not provide a unified treatment of vowel harmony phenomena due to conflicting patterns of *opacity* and *transparency* across systems. *Opacity* is a situation in which an individual segment or morpheme blocks the directional spread of features thereby limiting the extent of the harmony domain. *Transparency* occurs where an individual segment or morpheme remains unchanged by the spreading domain, which appears to pass through it.

The authors propose a framework that accounts for opacity and transparency through different rankings of the same set of constraints. They avail themselves of data from “Idealized Kinande” and “Idealized Pulaar”. As with many early Optimality-Theoretic proposals, underspecification is done away with by the stipulation of full underlying specification for all segments. The limits of a broader harmony domain are established when the constraints of Wide Scope Alignment (WSA) are crucially ranked over those of Basic Alignment (BA). BA stipulates that in a language without harmony, every sponsoring anchor of [F] (Anchor-s), which corresponds to a harmony trigger, is aligned with the edge of an [F]-domain. An anchor sponsors [F] if it is affiliated with [F] in the underlying representation. Non-sponsoring anchors are those anchors that come to be affiliated with [F] in the mapping form underlying to surface form by the operations of GEN. Sponsors may be X-slots or moras in the case of linked features or morphological constituents for so-called floating or morpheme-level features. WSA is the constraint governing the Principle of Extension, which extends the feature [F] over longer stretches of sound to maximize perceptibility and stability of the articulators by minimizing articulatory effort. However, these constraints alone do not force every vowel segment within the domain limits to harmonize; this work is done by another independent constraint EXPRESSION, which is defined in (19).

(19) EXPRESSION: [F] must be affiliated with every anchor in an F-domain.

As noted by the authors, the harmony domain is established when WSA and EXPRESSION are crucially ranked over prohibitions on the insertion of the feature [F], and no harmony domain exists when the reverse is true. This is shown in (20).

(20) Generation of harmony domains via constraint ranking

Harmony: WSA, EXPRESSION >> *Insert [F]

No harmony: *Insert [F] >> WSA, EXPRESSION

Yet, there are cases where feature occurrence will be ranked above EXPRESSION, leaving vowel segments within the defined domain unaffected by harmony. These are the cases of transparency and opacity, the rankings of which are tentatively modeled in (21). I reproduce the tableaux in the original notation of the authors.

(21) Transparency and opacity via constraint ranking

Transparency: [F] Occurrence, WSA >> EXPRESSION

Opacity: [F] Occurrence, EXPRESSION >> WSA

An additional constraint is proposed to address a common feature in [ATR] harmony systems in which there is no longer a [+ATR] low vowel. This proposal can be stated for ODT as in (22).

(22) CLASH: [+low] and [+ATR] do not co-occur in the feature profile of a single segment.

The opaque case is shown with the following tableaux of idealized Pulaar in (23). In both tableaux, the sponsoring anchor is the rightmost vowel. Tableau 1 shows the case of the opaque vowel [a], while Tableau 2 shows the standard harmony domain. CLASH, BA-rt, and EXPRESSION as a group dominate constraints creating a domain aligning the [ATR] feature specification with the left edge, which are nevertheless responsible for creating a [-ATR] domain at the left edge. The opacity of the low vowel is maintained through CLASH, which impedes the low vowel from assuming a [+ATR] feature.

(23) Harmony and opacity with idealized Pulaar

1. UR: C _Y C-a-C _v	CLASH	BA-rt	EXPRESSION	WSA-lf	BA-lf
☞ a. C _Y C-a-(C _v)				*	
b. (C _v C-a-C _v)	*!				*
c. C _Y C-a-C _Y		*!			*
d. (C _v C-a-C _v)			*!		*
2. UR: C _Y C-v					
☞ a. (C _v C-v)					*
b. C _Y C- _Y		*!			*
c. (C _Y C-v)			*!		
d. C _Y C-(v)				*!	

In idealized Kinande, shown in (24), [a] is a transparent vowel. The transparent case is presented in the upper tableau, while the regular harmony case appears in the lower tableau.

While progress was made in the modeling of harmony from an OT perspective, the strategies of ODT do not meet the requirements of the Karimojong case. The above examples deal with idealized data, because, as the authors explain, ‘the harmony systems of actual Pulaar and Kinande involve additional complications’ that do not affect their argument. Such complications are part of the grammars of these languages and should be theoretically accounted for. The surface irregularities in Karimojong from local interactions and dissimilation rules go unaccounted for in this model. Secondly, no effort is made to draw a distinction between feature-filling and feature-changing processes, which in Karimojong accounts for the difference between [a/o] alternation and the phonetic

advancement of [a] in [+ATR] domains. Thirdly, it utilizes CLASH across the board to account for nine-vowel systems while not addressing phonetic advancement.

(24) Harmony and transparency with idealized Kinande

1. UR: CV-k̄a-Cv	CLASH	BA-rt	WSA-lf	EXPRESSION	BA-lf
☞ a. (Cv-k̄a-Cv)				*	*
b. (Cv-ka-Cv)	*!				*
c. Cv-k̄a-Cv		*!			*
d. Cv-k̄a-(Cv)			*!		
2. UR: CV-Cv					
☞ a. (Cv-Cv)					*
b. Cv-Cv		*!			*
c. Cv- (Cv)			*!		
d. (Cv-Cv)				*!	*

2.1.2 Featural alignment (Akinlabi 1994)

One of the most highly productive constraint systems for [ATR] harmony has been Featural Alignment, defined in (25).

(25) Featural alignment

ALIGN (PFeat, GCat): Any occurrence of a prosodic feature is aligned with some grammatical category.

PFeat represents prosodic features that span grammatical categories, such as [ATR], pitch, nasality, roundness, palatalization etc. Featural alignment aligns a free harmonic element with an edge of a grammatical category, or with an entire grammatical category, as is the case of an [ATR] feature and harmony trigger serving as a morpheme where the structure of the original segment has been eroded.

For purposes of illustration, an example from the analysis of Kalabari, a Niger-Congo language with a nine-vowel harmony system, adapted from Akinlabi (1994:12)³ is shown in (30). The constraints employed are defined in examples (26) through (29). In (26), *LO/ATR establishes the nine-vowel harmony system by prohibiting segments that are [+low, +ATR]. Example (27) is a requirement that the [+ATR] feature be parsed within the morpheme. Examples (28) and (29) govern the [ATR] specifications of the vowels near the right and left edges of the stem.

(26) *LO/ATR: If [+low] then not [+ATR]. (Archangeli & Pulleyblank 1994)

(27) PARSE [ATR]: Underlying [+ATR] must be parsed within the morpheme. (Akinlabi 1994)

(28) ALIGN-RIGHT: ALIGN ([+ATR], ROOT, R): Any occurrence of [+ATR] must be aligned with the right edge of the root. (Akinlabi 1994)

(29) ALIGN-LEFT: ALIGN ([+ATR], STEM, L): Any occurrence of [+ATR] must be aligned with the left edge of the stem. (Akinlabi 1994)

³ The tableau in (30) has some differences with Akinlabi’s analysis, namely, in the number of violation marks assigned to ALIGN-RIGHT and ALIGN-LEFT for Candidate (a), as the unparsed [ATR] feature violates these two constraints.

The tableau in (30) presents the evaluation process for [dawo] ‘kola nut’, with an input consisting of a root with underspecified vowels and a free [+ATR] feature to be parsed. Winning candidate (b), simultaneously obeys the mandate to parse the [ATR] feature with rightward and leftward alignment, and the highly ranked co-occurrence restriction on the features [+low] and [+ATR]. Candidate (a) crucially violates the mandate to parse the [ATR] feature and thereby align the feature specifications of the vowels, while the remaining two candidates crucially violate the highly ranked prohibition on segments that are [+low] and [+ATR].

- (30) [ATR] harmony with mid and low vowels in Kalabari
[dawo] ‘kola nut’

/dAwO /+ATR/	*LO/ATR	PARSE-ATR	ALIGN-RIGHT	ALIGN-LEFT
a. dawɔ		*!	*	*
☞ b. dawo				*
c. dæwɔ	*!		*	
d. dæwo	*!			

This system of featural alignment has inherent assumptions that merit discussion. The first is that it has retained from the generative framework use of the underlying form as the input under candidate evaluation, while many OT practitioners at the time proposed doing away with this concept, as did Cole and Kisseberth above. The analysis of Karimojong presented in this paper follows Akinlabi in this respect. The second is that the [+ATR] feature is seen as the sole active feature in an [ATR] specification, while a [-ATR] feature is viewed as a default value. This view is common among researchers that deal with [+ATR] dominant systems exclusively, but it cannot be adopted for the analysis of Karimojong, where the [-ATR] feature in bidirectional root-controlled harmony is active, with even the low vowel [a] functioning as a trigger, as shown in §1.1 example (2). Secondly, Kalabari clearly has a suffix-dominant process, which some researchers have wished to posit as universal. The Karimojong data shows that this is not the case; additional machinery is required to evaluate forms in the language in which root-controlled and suffix-controlled processes act on one form. That requirement is met by the positional faithfulness constraint hierarchy (Beckman 1997, 1999), which will be presented in §3.2. Overall, the featural alignment system addresses domains characterized by exhaustive spreading within defined limits, but makes no provisions for the surface irregularities seen §1.2 and §1.3. Addressing the Karimojong case requires proposing a constraint system compatible with the existence of the intra-morphemic P-Structure in §1.4. To this end, the present paper will propose an adaptation of the AGREE family of constraints (Bakovic 2000), presented in the following sub-section.

2.1.3 AGREE (Bakovic, 2000, 2002)

The AGREE family of constraints assumes that the assimilation of features in harmony processes is strictly local, and has the general format of the constraint in (31).

- (31) General AGREE constraint

AGREE[F]: Adjacent segments have the same value of [F]. (Bakovic 2000: 4)

AGREE constraints are assumed to be left-right symmetrical. In other words, regardless of the input value of [F], there are always two candidates that will satisfy a given AGREE constraint: one where the segments are both [+F], and one where they are both [-F]. These cases of the former are shown in (32a,b), while cases of constraint violation are presented in (32c,d). Whether both segments are [+F] or [-F] is irrelevant as far as the AGREE constraint is concerned, and which of the two forms is ultimately the winner depends on the markedness and faithfulness constraints in the constraint hierarchy.

(32) Satisfaction and violation of AGREE[F]

<p style="text-align: center;">Satisfies AGREE[F]</p> <p>a. <i>x</i> <i>y</i> [+F] [+F]</p> <p>b. <i>x</i> <i>y</i> [-F] [-F]</p>	<p style="text-align: center;">Violates AGREE[F]</p> <p>c. <i>x</i> <i>y</i> [-F] [+F]</p> <p>d. <i>x</i> <i>y</i> [+F] [-F]</p>
---	--

(Bakovic 2000: 6)

In the case of a VCV sequence, the two vowels share the same features within the harmony domain only if the consonant likewise does. For a CVC sequence, in which consonantal features harmonize, the outer consonantal segments share features only if the vowel likewise does. Under AGREE, the reason for the relatively higher degree of incidence of vocalic as opposed to consonantal harmony processes is that consonants can adopt vocalic features without losing their identity as consonants, while vowels seldom adopt consonantal features without losing their identities as vowels.

Masaai and Turkana are the EN languages that Bakovic analyzes. He focuses on the [+ATR] dominance and considers stem control to be residual, disregarding any possible insights to be drawn from lexical phonology and the possibility of distinct morpho-phonological levels. In his analysis of the [a/o] alternation within a nine-vowel system, [a] is considered to be re-paired with [o] in the absence of a [+ATR] variant of the low vowel. Relying on Noske and Dimmendaal, Bakovic considers the low vowel [a] to be opaque; disharmonic forms are produced when a low vowel in the root is subjected to a dominant [+ATR] suffix as a trigger. The sample tableau in (33) shows a basic case of [ATR] spreading under his proposed constraint system. It analyzes the phonology under the morphological concatenation of a [+ATR] root, the frequentive suffix [-*aan-*] and the ventive [-*u*]. The AGREE[ATR] constraint is highly ranked, and the highest-ranked constraint, *[*v,ɣ*] prevents these [+ATR] vowels from surfacing within this [+ATR] domain; this is aided by the lower ranked identity constraints at the right end of the tableau. A constraint conjunction between *[-LO,-ATR] and IDENT-IO [ATR] is designed to penalize the appearance of [o] in the output form, although this conjoined constraint cannot apply across the board in the language as [-ATR] vowels that are [-low] appear throughout the lexicon. The Principle of Globality in OT, which stipulates the global application of constraints throughout the language is thus violated, indicating that the analysis requires further improvement. Additionally, the regularity of spreading advanced by this constraint system makes no allowances for the irregularities found in surface forms due to consonant-generated features and other localized effects such as the [ATR] dissimilation rule seen in §1.2. To be able account for these phenomena, a modification of the AGREE system is proposed in §3.1.

(33)

/tur/ +/AA _n /+/u/	*[<i>v,ɣ</i>]	AGREE [ATR]	*[-LO,-ATR] & IDENT-IO [ATR]	IDENT- IO [ATR]	IDENT- IO [LO]	IDENT-IO [round]
a. tur <u>aa</u> nu		**!				
b. tur <u>æ</u> nu	**!			**		
c. tur <u>ɔ</u> nu	**!			**	**	
☞ d. tur <u>oo</u> nu				**	**	**
e. tur <u>aa</u> nu		*!	*	*		
f. tur <u>oo</u> nu			**!	**		

2.1.4 Headed spans (McCarthy, 2004)

The Headed Spans proposal attempts to describe spreading processes within Optimality Theory by using faithfulness and markedness constraints. Accordingly, every harmonic span has a head segment with a value *a* for a feature [F] that is shared by the other members of the span, which occur in a contiguous string. Directionality is to be expressed through markedness constraints requiring the head to lie on a certain edge of the span. Faithfulness constraints require [*a*F] segments heading spans in the input to head them in the output. Substituting for ALIGN and AGREE constraints in OT are markedness constraints that adjacent spans violate.

The crucial element of the Headed Spans proposal that determines its applicability to the Karimojong case is the principle that each type of harmony domain is defined by a unique constraint ranking. Bidirectional root-controlled harmony would have a set of constraints which can be labeled Constraint Ranking R. Suffix-controlled harmony would be governed by a Constraint Ranking S. Demonstrating that this proposal cannot work for Karimojong is relatively straightforward.

In (34a), bidirectional root-controlled harmony, governed by Constraint Ranking R spreads the [−ATR] feature from the root so that the vowel [a/o] of the suffix is realized as [a]. In (34b), the same suffixation process is carried out on a verb which would permit [+ATR] root controlled harmony. But this is a case in which the consonant-generated [−ATR] feature creates a mini-domain that spreads bidirectionally leftward. This suffix-based process is governed by a different set of constraints – Constraint Ranking S, and the result shows no exhaustive bidirectional spreading of the root [+ATR] specification nor of the [−ATR] feature of the suffix. The phonology of the same morphological suffixation processes would have to be governed by two unique constraint hierarchies, which contradicts the basic assumptions behind the proposal. The Karimojong case requires a common constraint set that can accommodate both phenomena, such as will be seen in §3.5.

- (34) a. aki-dóŋ +/Ar/ → á-dóŋ-àr constraint ranking R
 b. aki-dó+/Ar/ → aki-dó-òr constraint ranking S

For its over-emphasis on the regularity of vowel harmony processes, Span Theory, as proposed, cannot adequately account for the irregularities in the surface forms in Karimojong [ATR] harmony data, which were found as the result of painstaking descriptive data collection.

2.2 Harmonic serialism

Works such as Idsardi (1998), Bermudez-Otero (1999), Kiparsky (2000) and the anthology of Hermans and van Oostendorp (1999) highlight the problematic nature of the parallel derivation concept in OT. As noted by McCarthy (2010), parallel OT is a direct mapping theory between underlying and surface representations without any intermediate steps. Nevertheless, the OT process of candidate evaluation under a hierarchy of ranked violable constraints does not necessarily need to be linked to a direct mapping architecture but to one with indirect mapping—in other words, with intermediate derivational steps. The version of indirect mapping proposed by McCarthy is called Harmonic Serialism. Candidate evaluation is carried out multiple times to generate intermediate forms, with the stipulation that each intermediate form generated constitutes a harmonic improvement over the previous form. This is defined as a candidate incurring fewer violation marks than the previous form, although there are still undefined questions regarding how gradually a derivation is to be developed, given that the definition of the changes effected is based on theoretical artifacts.

Given the scope of the field of phonology, models similar in form are structured over different bases. Derivation in the model presented here is morphologically based, in that the phonological processes come packaged in the processes of affixation. The data result from reconstructed and reconstructable historical changes. Given this connection to physical reality, the model presented in this paper is considered to have explanatory power superior to that of Harmonic Serialism, and detailed argumentation in support of the three morpho-phonological levels can be found in Lesley-Neuman (2012). Since MP levels are verifiable through psycholinguistic experimentation, as in Vannest & Boland (1999), no further attempt to model EN phonology through Harmonic Serialism will be made.

3 Addressing Eastern Nilotic typological features

Cole (2009) argues that representational models of phonology, such as those seen in §2.1.1 through §2.1.4, make faulty predictions about the degree of uniformity found in vowel harmony processes. Under these models, assimilation spreads directionally and exhaustively from the harmony trigger to the boundaries of the domain, due to a mandate to spread features or through the alignment of feature structures with the edges of the domain. All segments are targeted and subject to feature compatibility with no asymmetries within the domain, without any restrictions on the harmony process across the

word. Such models, without additional machinery, cannot accommodate the variation in the [ATR] statuses of segments found in surface forms of the Karimojong verbal complex as seen in §1.2 through §1.4. Furthermore, two separate harmony domains cannot be accommodated within a one-step parallel derivational process, first of all while remaining consistent and true to the historical-cognitive model in which morpho-phonological levels are posited, and secondly because the fixes to parallelism fail to provide adequate and plausible explanations, as will be shown in §3.6. It is the thesis of this paper that only a descriptive but systematic approach to the target languages can provide the amount and quality of data necessary to accurately formulate generalizations upon which theoretical analyses are based, and to discover any exceptions and irregularities to be explained. Likewise, it is imperative that the formulation of generalizations be driven by the data rather than by the artifacts and machinery of theoretical models, so as to avoid mischaracterizations of inherently natural processes.

This whole-language approach to phonology permits the clarification of theoretical issues. Firstly, the constraint systems and evaluation processes presented in §2.1.1 through §2.1.3 assumed differing positions on one essential point: whether underspecification is required to adequately and accurately present input-output relationships in an optimality-theoretic framework. For Eastern Nilotic, this question was answered perhaps definitively by Manuela Noske (2000: 801–804), who first postulated that the alternating suffixes, deemed Level 1 affixes in the present paper, were underspecified underlyingly and received their specifications from the root as the dominant harmony trigger, while domains effected by [+ATR] dominant suffixes, Level 2 processes in the present analysis, were feature-changing operations. Such a conclusion can only be drawn while making sense of the Turkana data as a whole.

Noske (2000) also provides a second example. Previous to its publication, standard theoretical orthodoxy postulated the dominance of feature specifications of the root over those of suffixes, despite the existence of ample evidence of suffix-based harmony processes in African languages. McCarthy and Prince (1995) proposed as a linguistic universal in optimality-theoretic terms for root faithfulness to out-rank suffix faithfulness. The [+ATR] suffix dominance in Turkana, as in Karimojong, a language to which it is historically related, overrides root-controlled processes and even root specifications, a fact that disproves this ill-conceived universal, and which is presented and argued in theoretical terms.

Nonetheless, approaches to fieldwork and organization of the data do impact the ultimate analysis. Eastern Nilotic languages, with potentially large, agglutinated verbal complexes that offer rich sources of expression to accomplished speakers, provide ample data with respect to the phonological behavior of affixes and multiply affixed forms. A position class analysis with as many positions as can be elicited naturally from native speaker consultants (Doris Payne, p.c.) is imperative to properly characterize this behavior and prevent faulty generalizations. Noske (2000), in the effort to argue theoretical points, misses generalizations that could potentially have been made from the data, and presents contradictions both in her analyses and with the data available from well-established sources, some of which are cited in her article. In addition, a very small number of examples are employed to illustrate her points, and some of these are repeated throughout the paper.

For example, her insightful observation regarding feature-filling versus feature-changing processes goes unapplied in the case of the itive suffix [-*Ar*], which shows the classic [a/o] recessive alternation as a feature-filling process, with a variant in which the suffix [-*ɔr*] surfaces in some [+ATR] verbs as a harmony trigger spreading a [-ATR] feature bidirectionally. This is as opposed to the feature-changing processes on the next morpho-phonological level in which other vowels change their specifications to [±ATR] while the [a] vowel remains unchanged in its basic identity or advances slightly under a dominant [+ATR] process. That these facts are common to other Eastern Nilotic languages, and even properly to Turkana, is seen in §1.2 of this present paper, the data for which was taken from Novelli (1985); in Dimmendaal (1983: 110–112) with an additional voiceless vowel for infinitive forms, and in Schroeder and Schroeder (1984) for Toposa. Acoustic evidence of tongue root advancement for feature-changing [+ATR] processes is subsequently provided in Guión, Post and Payne (2004), but the basic [a/o] alternation for Masaai is described by Tucker and Mpaayei (1955). This information is subsequently summarized in Dimmendaal (2002: 172) as part of an analysis of Eastern Nilotic derivational suffixation, which is reproduced below in (35).

(35) Selected Eastern Nilotic derivational suffixes

	Bari	Maasai	Ateso	Toposa	Turkana
Ventive	-u(n)	-o(n), -u(n)	-o(n), -u(n)	-o(n), -u(n)	-o(n), -u(n)
Itive	-ara, -ara	-a(a), -o(o)	-ar, -or, -or	-ar, -or, -or	-ar, -or, -or
Dative	-ki(n)	-ki(n), -ki(n)	-ki(n), -ki(n)	-ki(n), -ki(n)	-ki(n), -ki(n) ⁴

Nonetheless, Noske (2000: 785) presents the itive suffix as alternating between [-ar] and [-or] only, without mentioning the fundamental difference between the two suffixes, i.e. the recessive nature of [-ar] and the dominant nature of the exceptional [-or], and appears to attribute the effects of this exceptional suffix to characteristics inherent to the affected vowels rather than consider the role of feature transfer from the consonants in the environment and its intersection with vowel mergers in progress, as discussed in Hall and Creider (1998). That it cannot be attributed to the inherent characteristics of the vowels can be shown in the analysis of other recessive [a/o] suffixes for the same set of verbs, as done for Karimojong in §1.2 example (5) and which can be postulated from careful inspection of the Turkana data in Dimmendaal (1983: 111–117). That the final vowel [-i/-i] marking infinitive forms of the dative is recessive and follows the [ATR] specification of the root is shown in (36). Itive forms of [+ATR] verbs postulated from Noske's stated rule appear in the second column in (37). Dative forms are taken from Dimmendaal (1983: 116). There is also an ample supply of itive forms of [+ATR] verbs employing the [-o] alternate of the itive [-or]. A sample is presented in (38); those cases for which Noske presents a form with [-or] are noted. As is the case with Karimojong and other Eastern Nilotic languages, the behavior of the exceptional [-ATR] suffix [-or] appears from the phonologization of the co-articulation effect of the rhotic on the vowel, rather than from any inherent characteristic of the vowel across the language.

(36) Infinitive dative and root forms in [±ATR] verbs

<u>root</u>	<u>gloss</u>	<u>infinitive dative</u>	<u>gloss</u>
-sɔb	'do'	a-sɔb-akin-ɨ	'do for'
-tjak	'divide'	a-tjak-akin-ɨ	'share with'
-iboj	'sit'	aki-boj-ikin-ɨ ⁵	'sit down somewhere'
-rem	–	a-rem-okin-ɨ	'hate'

(37) Postulated itive with infinitive dative forms

<u>root</u>	<u>postulated itive</u>	<u>infinitive dative</u>	<u>gloss</u>
-rem	-rem-or	a-rem-okin-ɨ	'hate'
(-bus)	-bus-or ⁶	aki-bus-okin-ɨ	'drop'

(38) Itive forms of [+ATR] verbs

a-limw-ɔrɨ	'confess'
a-buk-ɔrɨ/ a-buk-or	'pour out'
a-ped-ɔrɨ/ a-ped-or	'pierce'
a-bon-oon-ɔrɨ	'go to and fro'
a-tik-oon-ɔrɨ	'rape, violate'
a-k-ilik-ɔrɨ/ a-k-ilik-or	'take down'
ke-duk-ɔrɨ	'...that it had been built'

⁴ A printing error in Dimmendaal (2002) is corrected here.

⁵ Vowel deletions and substitutions frequently occur in dative forms, eliminating the [+ATR] characteristic vowel [o] from the dative affix in this case.

⁶ The specification of the root vowel of this form does not change due to the resistance to [-ATR] spreading of the voiced consonant [b]+[+ATR] high vowel [u] in the root [-bus].

Further comparison of the works of Noske and Dimmendaal confirms the validity of the Lesley-Neuman's (2012) historical model of suffix change in [ATR] harmony participation over time. For example, the subjunctive suffix [-rɛ] appears in Dimmendaal (1983: 111–117) as a neutral affix triggering no spreading of its [-ATR] specification, as shown by the cases of [±ATR] verbs in (39). Examples drawn from Noske (2000: 780) in (40) present it as a dominant trigger for the creation of a [-ATR] domain, changing the [ATR] specifications of the root vowels of verb forms.

(39) Subjunctive forms of [±ATR] verbs

[-ATR]		[+ATR]		
a.	α-møk-à-rɛ	‘mix’	c. α-bun-i-ò-rɛ	‘come’
b.	α-k-irrk-à-rɛ	‘hunt’	d. α-k-i-pej-ò-rɛ	‘welcome (someone)’

(40) Infinitive and subjunctive forms of [+ATR] verbs

a.	α-ki-rem	‘to spear’	ε-rɛm-ε-rɛ	‘(Why) is it speared?’
b.	α-ki-mor	‘to insult’	ε-mɔr-ε-rɛ	‘(Why) is he insulted?’

At issue is Noske's treatment of neutral suffixes in her optimality-theoretic analysis. While declaring neutral suffixes to be in the same prosodic word as the base of suffixation [786], she does not account for their neutrality in her constraint rankings, and provides only a small number of relevant data points in her presentation [781: 35]. This failing leaves unexplained neutral suffixes in other Turkana varieties, such as those presented by Dimmendaal.

Nonetheless, an important issue to the analysis of Turkana is the behavior of low vowels. Both Dimmendaal (1983: 297–299) and Noske (2000: 781) present the gerundive suffix [-e] as a [+ATR] dominant suffix, the spreading of which is not effected where low vowels and low vowel derivational suffixes intervene. Examples are seen in (41), (42) and (43). In (43a-d), the gerundive suffix [-e] changes the root and other vowel specifications of all forms to [+ATR], and in some cases adding consonantal gliding to the forms. In (41d) another vowel change is effected: [i] + [a] → [e]. The impersonal pronominal prefix [-e-] of the gerundive in these examples is [+ATR]. Example (41e) shows a case in which the low vowel [a] of the reduplicated root appears to block [+ATR] spreading, effecting no feature change. The pronominal prefix of this gerundive form is [-e-], which bears a [-ATR] specification. Example (41f) shows the root and gerundive forms of a [+ATR] verb, in which all vowels, including that of the alternating pronominal prefix, bear a [+ATR] specification.

(41) Root and gerundive forms

	<u>root</u>	<u>gloss</u>	<u>gerundive</u>	<u>gloss</u>
a.	-dɛk	‘be sick’	e-dek-e	‘disease’
b.	-imɔj	‘eat’	e-k-imɔj-e	‘way of eating’
c.	-cil-i-cil	‘scratch’	e-k-cil-i-cil-e	‘way of scratching’
d.	-li-i-à-rɛ	‘crawl’	e-lij-è-je	‘way of crawling’
e.	-i-lam-a-lam	‘toss sandals’	ε-k-i-làm-a-lam-e	‘way of tossing sandals’
f.	-itu-tuk	‘advise’	e-k-itù-tuk-e	‘way of advising’

In (42a,b), the [+ATR] specification does not spread due to blocking of the low vowel alternates of the infinitive and the dative. Note that even the high vowel [ɪ] of the dative suffix retains the [-ATR] specification. The alternating pronominal prefix assumes the [-ATR] specification of the root. In (42c), the [+ATR] specification of the gerundive suffix spreads to the dative suffix, endowing it with a [+ATR] specification, while the low vowel in the root retains its [-ATR] specification, which is also supplied to the pronominal prefix. Example (42d) allows the spread of the [+ATR] feature through the dative suffix and effects a change in the specification of the root vowel and the pronominal prefix. Absent a low vowel, spreading is unobstructed. Examples (42e,f) are [+ATR] forms with [-un] ventive suffixes; all vowels bear the same [+ATR] specification.

(42) Root and gerundive forms with derivational suffixes of [±ATR] verbs

	<u>root</u>	<u>gloss</u>	<u>gerundive</u>	<u>gloss</u>
a.	-i-lot-ar	'wash'	ε-k-lot-ar-e	'way of washing'
b.	-inək-akın	'light'	ε-k-ınək-akın-e	'way of lighting'
c.	-na-ı-kın	'give'	ε-na-ıkin-e	'way of giving'
d.	-en	'tie'	en-ikin-e	'way of tying'
e.	-wok-un	'carry'	e-wok-un-e	'way of carrying'
f.	-ido-un	'bear'	e-k-idò-un-e	'way of bearing'

In examples (43a,b), the dominant [+ATR] specification from the gerundive suffix spreads to all of the vowels in the word, and is not blocked by the [-Un] ventive, which contains a high vowel. In contrast, the examples in (43c,d) present cases in which the [α]-alternate of the dative suffix blocks spreading. In the data presented by Noske, the high front vowel [i] of the dative suffix does alternate and assumes the [+ATR] specification of the root, differing from the cases in examples (42a,b).

(43) Infinitive and gerundive forms of [-ATR] verbs

	<u>infinitive</u>	<u>gloss</u>	<u>gerundive</u>	<u>gloss</u>
a.	agjel-ən	'to buy'	e-gjel-un-e	'way of buying'
b.	akı-dək	'to climb'	e-dok-un-e	'way of climbing this way'
c.	akı-lıp	'to pray'	ε-k-ılıp-akın-e	'way of praying'
d.	akı-nək	'to light'	ε-k-nək-akın-e	'way of lighting'

From the disharmonic Turkana examples presented above, it is necessary for an optimality-theoretic analysis to account for the role of low vowels in blocking [+ATR] spreading. Noske, following previous researchers, does so through the application of the constraint *LO/ATR. Disharmonic forms are accounted for by using two alignment constraints, ALIGN-R and ALIGN-L, as shown in (44)⁷. She ensures dominance of the suffix [ATR] specification by assigning a high ranking to the constraint IDENT-IO SUFFIX[ATR], crucially ranking it over IDENT-IO[ATR]. This is part of the machinery of a positional faithfulness hierarchy, explained below in §3.3. The constraint IDENT-IOROOT[ATR] likewise ensures root dominance in forms with recessive rather than dominant suffixes.

(44)

	[-ATR] [-ATR][+ATR] ε- ram - e	IDENT-IO SUFFIX [ATR]	*LO/ATR	ALIGN-L [αATR]	ALIGN-R [αATR]	IDENT-IO ROOT [ATR]	IDENT- IO [ATR]
a.	[+ATR] / \ / ε-ræm- e		*!			*	**
b.	☞ [-ATR] [+ATR] / \ ε-ram- e			*	**		
c.	[-ATR] / \ / ε- ram- ε	*!					*

In addition to not accounting for neutral suffixes, Noske's (2000) analysis presents other problems. The first is the violation of the Principle of Globality in OT. The Principle of Globality states essentially that a constraint ranking for a particular language variety must be applicable to all forms throughout that variety. The second problem that arises is that irregularities stemming from consonant-generated features remain unaccounted for.

⁷ I have corrected a tableau used for an intermediate step in Noske's analysis to represent the final constraint ranking in which ALIGN-L >> ALIGN-R. I have also corrected errors in the assignment of violation marks for these two constraints, which force alignment with the [ATR] specifications that occur at the left and right edges of the stem respectively, differing slightly from similar constraints proposed by Akinlabi above.

The tableau in (45), taken from Noske (2000: 799) employs the high-ranking constraint *HI/RTR to block feature change of a [+ATR] root vowel of the verb [*a-buk-ɔr*] ‘to pour out’ under suffix dominant [-ATR] spreading. The proposed input form is unclear, and the suffixation of the itive is proposed to be a feature-changing operation. Nonetheless, the constraint ranking produces the attested form. In tableau (46), taken from Noske (2000: 801) the same constraint ranking is employed for the ventive form [*a-dɔk-un*] ‘to climb this way’. It is unable to produce the attested form. The constraint *HI/RTR as posited, while expressing the opposition between raising the tongue in the production of high vowels and retracting it for [-ATR] feature, does not address the complex situation of historical change and consonant-generated features in the case of the ventive suffix, as discussed in §1.2. This indicates that the model is in need of revision. This present paper accounts for consonant-generated features by proposing the DOMAIN C constraint family in §3.4. The variability of suffix participation in harmony due to historical change is accounted for by a stratal optimality-theoretic model. Advancing this model requires the modification of the AGREE constraint system, discussed in the following section.

(45) Use of *HI/RTR under [-ATR] suffix control

a-buk-ɔr - to pour out

$\begin{matrix} [+ATR] & [-ATR] \\ & \\ \alpha\text{-buk-} & \text{ɔr} \end{matrix}$	IDENT-IO Suffix [ATR]	*HI/RTR	ALIGN-R [αATR]	ALIGN-L [αATR]	IDENT-IO Root [ATR]	IDENT-IO [ATR]
a. $\begin{matrix} [-ATR] \\ \wedge \\ \text{-b} & \text{ɔ} & \text{ɔr} \end{matrix}$		*!			*	*
b. $\begin{matrix} [+ATR] \\ \wedge \\ \text{-buk-} & \text{or} \end{matrix}$	*!					*
c. $\begin{matrix} [+ATR] & [-ATR] \\ & \\ \text{-buk-} & \text{ɔr} \end{matrix}$			*	*		

(46) *a-dɔk-un* - to climb frequently

$\begin{matrix} [-ATR] & [+ATR] \\ \wedge & \\ \alpha\text{-d} & \text{ɔk-un} \end{matrix}$	IDENT-IO Suffix [ATR]	*HI/RTR	ALIGN-R [αATR]	ALIGN-L [αATR]	IDENT-IO Root [ATR]	IDENT-IO [ATR]
a. $\begin{matrix} [-ATR] \\ \wedge \\ \text{-d} & \text{ɔk-} & \text{un} \end{matrix}$	*!	*				*
b. $\begin{matrix} [+ATR] \\ \wedge \\ \text{-dok-} & \text{un} \end{matrix}$					*	*
c. $\begin{matrix} [-ATR] & [+ATR] \\ & \\ \text{-d} & \text{ɔk-} & \text{un} \end{matrix}$			*	*		

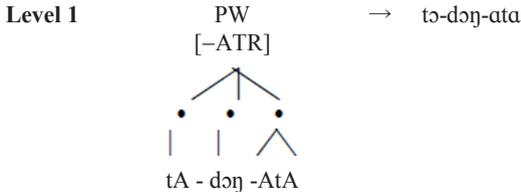
3.1 Modification of the AGREE constraint system

Compliance with the Principle of Globality and the avoidance of ad hoc constraints requires the establishment of a constraint system that is general in scope but adaptable to the special conditions imposed by EN phonological data, such as that shown in §1.4 example (18). This system is based on an addition to the AGREE constraint family to make allowances for the phonological structure internal to the morpheme by enforcing agreement of the heads of prosodic word nodes within that structure. This constraint is presented in (47).

(47) PW head agreement

AGREE [α ATR, Hd, PW]: The [ATR] specification of every phonological head has the same value as its dominant PW node.

An illustration of its effects is shown in the Level 1 process in (48). The [-ATR] specification has percolated from the head of the root vowel to the PW node. The [ATR] specifications of all the heads in the PW must agree in value, and therefore must all be ATR. The resulting output is [tɔ-dɔŋ-ata].

(48) Action of AGREE [α ATR, Hd, PW]

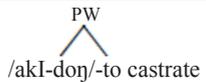
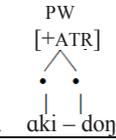
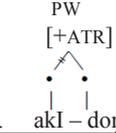
AGREE [α ATR, Hd, PW] is one of several violable constraints within a hierarchy that determines the final shape of output forms. The ranking proposal continues below with the presentation of an additional constraint required for a ranking tableau.

As shown in §1.1, a bidirectional root-controlled harmony domain results from the spreading of [\pm ATR] specification from the root vowel both to the left and right. The basic assumption of this OT analysis is that the recessive affixes receive the [ATR] specification from the root vowel and are underlyingly underspecified, while the harmony trigger is underlyingly specified. Spreading from the root vowel follows from the AGREE constraint in (47), which stipulates agreement between the feature [α ATR] that percolates to the PW node and those of the phonological heads that determines surface specification values. Any optimal candidate with harmony violates the low-ranking constraint IDENT IO [α ATR], defined in (49), when the unspecified alternating suffix input adopts the [α ATR] feature of the root, causing its output feature specification to be distinct from that of the underspecified input.

(49) IDENT-IO [α ATR]: Let β be an output segment and γ the input correspondent of β . If β is [α ATR], then γ must be [α ATR].

For this to be effected, AGREE [α ATR, Hd, PW] must be crucially ranked over IDENT-IO [α ATR]. This is seen in the ranking tableau in (50). Candidate (b) crucially violates AGREE [α ATR, Hd, PW] as the [ATR] value of the head of the alternating prefix does not align with that of the PW node. Candidate (a) is aligned, but violates IDENT-IO [α ATR] as harmonizing the [ATR] value of the prefix with that of the root, violating faithfulness to the input. It is, however, the winning candidate and attested form, as it manifests root control of the [ATR] specification of the PW and spreading of the specification leftward.

(50) AGREE [α ATR, Hd, PW] >> IDENT-IO [α ATR]

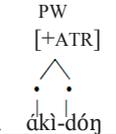
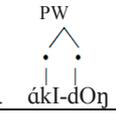
	AGREE [α ATR, Hd, PW]	IDENT-IO [α ATR]
		*
	*!	*

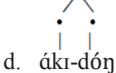
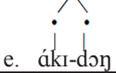
Vowels underspecified for [ATR] never appear in output forms of the Karimojong language, yet exist in input values. Therefore, there is need to posit a constraint requiring a tongue root specification to adequately represent the process within EVAL. That constraint is INTERPRETABILITY [ATR] and it is proposed in (51). It must exist as a highly ranked constraint to block underspecified vowels from the surface form and for constraints governing the harmony process to have proper effect. Its ranking over AGREE [α ATR, Hd, PW] and IDENT-IO [α ATR] is shown in the ranking tableau in (52).

(51) INTERPRETABILITY [ATR]: All vowels must bear a specification for tongue root position. (Pulleyblank 1997)

The ranking tableau in (52) presents the derivation of /akI-donj/ ‘to castrate’, in which the prefix vowel is underspecified. The attested output form, Candidate (a) is indicated to be the optimal form when INTERPRETABILITY [ATR] is crucially ranked over AGREE [α ATR, Hd, PW] and IDENT-IO [α ATR], as the [ATR] specification of the prefix vowel agrees with that of the dominant PW node. Candidates (b) and (c) crucially violate this constraint, as both vowels in the output form have surfaced unspecified for [ATR]. Ranking this constraint below IDENT-IO [α ATR] would otherwise allow unspecified forms to surface in the output. Candidate (d) is disallowed due to a violation of AGREE [α ATR, Hd, PW], in that the [-ATR] specification of the prefix vowel does not agree with the specification of the dominant PW node. Yet as a vowel specified for [ATR], it is unaffected by INTERPRETABILITY. Candidate (e) changes the [ATR] specification of the root vowel to [-ATR], and surfaces with a [-ATR] prefix vowel, incurring two violations of IDENT-IO [α ATR], one more than Candidate (a), the attested output form.

(52) INTERPRETABILITY [ATR] >> AGREE [α ATR, Hd, PW] >> IDENT-IO [α ATR]

/ákI-dónj/ ‘to castrate’	INTERPRET [ATR]	AGREE [α ATR, Hd, PW]	IDENT-IO [α ATR]
	α		*
	**!		*

/ákĩ-dón/ ‘to castrate’	INTERPRET [ATR]	AGREE [αATR, Hd, PW]	IDENT-IO [αATR]
PW [+ATR] 		*!	*
PW [-ATR] 			**!

3.2 Positional faithfulness (Beckman 1997)

In the analysis thus far, the alternation of prefix vowels has been explained through initial underlying underspecification and subsequent agreement with the [ATR] specification of the dominant PW node, which has percolated from the root. To effect a root-controlled process, the constraint ranking must reflect that the identity of the root vowel of the output with respect to the input is maintained.

- (53) IDENT-IO-Position [F]: Let β be an output segment in a privileged position P and α the input correspondent of β. If β is [αF], then γ must be [αF]. “Corresponding segments in a privileged position must have identical specifications for [F].” (Beckman 1999)

The second of the two constraints is simply an IDENT constraint specific to the feature, as shown in (54).

- (54) IDENT-IO [F]: Let β be an output segment and γ the input correspondent of β. If β is [αATR], then γ must be [αATR]. (McCarthy & Prince 1995)

The positional faithfulness ranking schema, which effectively forces all other vowels in the harmony domain to adhere to feature specification of the dominant position is as shown below in (55).

- (55) Ranking schema for positional faithfulness (Beckman 1999)
 IDENT-IO-Position [F] >> IDENT [F]

To deploy the constraints in the positional faithfulness hierarchy in (55) for the Karimojong case, the constraints in (53) and (54) must have [ATR] as [F], and the privileged position should be indicated as the root. Hence faithfulness to the input [ATR] value of the root vowel is more highly ranked than that of any of the affix vowels, which, under this [ATR] harmony model are proposed to be underspecified. Hence the two constraints required for the Karimojong positional faithfulness hierarchy are presented in (56) and (57).

- (56) IDENT-IO ROOT [αATR]: Let β be an output segment in a root and γ the input correspondent of β. If β is [αATR], then γ must be [αATR].
- (57) IDENT-IO [αATR]: Let β be an output segment and γ the input correspondent of β. If β is [αATR], then γ must be [αATR].

In any given optimal form with root-controlled harmony, the root vowel with its underlyingly specified [αATR] value surfaces faithfully, while alternating segments, which are unspecified, violate IDENT-IO [αATR] when they adopt the [αATR] feature of the root on the surface. Since IDENT-IO

[α ATR] is crucially ranked below IDENT-IO ROOT [α ATR], the hierarchy indicates, the other vowels assume the [ATR] specification of the root vowel. In (58) below, winning candidate (a) is the one in which the vowel specification is dominant over those of the alternating affixes. The solution is a positional faithfulness hierarchy (Beckman 1999), in which a highly ranked constraint defines a privileged position with respect to other positions. A positional faithfulness hierarchy utilizes specific formulations drawn from the IDENT family to posit two constraints, one dominant over the other, to produce the desired effect. The first maintains the identity of a given feature of a segment in a privileged between the input and the output, and is proposed as shown in (54). The proposal requires the segment in the privileged position to be specified in the input.

(58) AGREE [α ATR, Hd, PW] >> IDENT-IO ROOT [α ATR] >> IDENT-IO [α ATR]

/áki-dóŋ-An/-to castrate frequently	AGREE [α ATR, Hd, PW]	IDENT-IO ROOT [α ATR]	IDENT-IO [α ATR]
<p>a. áki-dóŋ-on</p>			**
<p>b. áki-dóŋ-an</p>		*!	***
<p>c. áki-dóŋ-on</p>	*!		**

3.3 Accounting for other harmony domains

As shown in §1.1, and in Lesley-Neuman (2012: 121–140), Karimojong has three morpho-phonological levels, all which are defined by distinct [ATR] harmony domains. Of the two active domains, the first domain featuring bidirectional root-controlled [ATR] harmony, was successfully captured in OT by the positional faithfulness hierarchy presented in §3.1. The question remains how the presence of the second harmony domain, that of leftward-spreading [+ATR] suffix-controlled harmony, should be represented. As discussed in §1.4, the two domains differ in that the bidirectional root-controlled [ATR] spreading process is a feature-filling operation in which the [ATR] specification of the suffix vowels is supplied by the specification of the root vowel, while the dominant [+ATR] suffix-controlled process is a feature changing operation. These two processes are concurrently present in Karimojong word formation (Lesley-Neuman 2007, 2012). Representation of this in an optimality-theoretic framework builds upon the fact that segments acting as triggers are underlyingly specified. Therefore a highly-ranked constraint with a suffix vowel with an [ATR] value underlyingly specified must be posited to override the dominance of the root vowel specification within the positional faithfulness hierarchy. This constraint appears in (59).

(59) IDENT-IO SPEC SUFFIX [α ATR]: Let β be an output segment in an underlyingly specified suffix and γ the input correspondent of β . If β is [α ATR], then γ must be [α ATR].

Hence, the following constraint hierarchy governs the dominant [+ATR] domain:

(60) Constraint hierarchy for [+ATR] domain

IDENT-IO SPEC SUFFIX [α ATR] >> AGREE [α ATR,PW] >> IDENT-IO ROOT [α ATR] >> IDENT-IO [α ATR]

This hierarchy is featured in (61) for the suffixation of the narrative mood present perfect tense morpheme /tetei/ to [ito-dɔŋ] ‘and you pinch’ forming [ito-dɔŋ-tetei] ‘and you have pinched’. Winning candidate (a) shows exhaustive agreement among all vowels in the word with the underlying specified [+ATR] suffix /tetei/. This entails the alternation of the root vowel [ɔ] of [ito-dɔŋ]. As the head of the prosodic word changed its specification owing to the dominance of the suffix specification, this candidate does not violate the AGREE constraint. It does violate the constraint governing root faithfulness IDENT-IO ROOT [αATR], and incurs two violations of the low-ranking identity constraint governing input-output faithfulness of the [ATR] specifications at the word level. Neither of these violations is crucial. Candidate (b) crucially violates AGREE [αATR,PW] by maintaining the [-ATR] specification of the prefix vowel. Like Candidate (a), it violates the root faithfulness constraint with the alternation of the root vowel to agree with the specification of the suffix. Candidate (c) crucially violates the highest-ranking constraint IDENT-IO SPEC SUFFIX [αATR] by allowing the suffix vowel to alternate and adopt the original [-ATR] specification of the root vowel. This also incurs three violations of IDENT-IO [αATR], with no impact on candidate evaluation.

(61)	[ito-dɔŋ]+ /tetei/ ‘and you have pinched’	IDENT-IO SPEC SUFX [αATR]	AGREE [αATR, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
	<p>a. ito-dɔŋ-tetei</p>			*	**
	<p>b. ito-dɔŋ-tetei</p>		*!	*	*
	<p>c. ito-dɔŋ-tetei</p>	*!			***

3.4 Localized effects and optimality theory

As noted in §1.2, forms such as [áki-dɔŋ-ɔr] and [áki-dɔ-ɔr] represent exceptions to the binary [a/o] alternation frame, and are proposed to stem from the phonologization of a co-articulation effect of [r] on mid and back vowels, also have an apparent relationship with a merger of [ɔ] and [ɑ] in progress. Tongue retraction that occurs in the pronunciation of the rhotic consonant in the suffix [-Ar] produces a [-ATR] feature that overrides the [ATR] specification of the preceding vowel, if it had been originally targeted to be [+ATR]. This anticipatory co-articulation effect has become phonologized to the point where the speech community has generated a bidirectional [-ATR] spreading process with limitations on the domain, a maximum of two syllables to the left and to the right of the affected vowel segment. A [-ATR] feature specification is spread to mid and high vowels to the left, including that of the root of [+ATR] verbs, and selects [-ATR] TMA markers to the right in the cases of Form B conjugations.

Effecting this [-ATR] suffix-controlled spreading process, as shown in (5), requires the action of two constraints. The first is IDENT-IO SPEC SUFFIX [αATR], introduced in (59), which overrides the specification of the root vowel and launches a spreading domain. It should assume the form of constraint mandating faithfulness to the [ATR] specification of the suffix vowel, which in this case is [-ATR] and is ranked more highly than root-specification faithfulness constraint. The second stems

from a constraint family proposed in (62), DOMAIN C[F]. The specific constraint governing this [-ATR] is presented in (63).

- (62) DOMAIN C[F]: The local domain D of a feature [F] from consonant C is N1 syllables to the left of the consonant trigger, and N2 syllables to the right of it, where N1 and N2 are elements of defined sets of natural numbers. The vowels immediately adjacent to the consonant trigger are principally affected by the spread of the feature.
- (63) DOMAIN [r]/[-ATR]: The local domain D of the [-ATR] feature generated by [r] is defined by the set of natural numbers $n \leq 3$, where n is the number of syllables to the left and to the right of the head vowel.

That the spread of a consonant-generated feature is dictated by a natural number representative of the strength of a co-articulatory effect and not any prosodic unit is indicated by forms from two verbs in the Novelli data. These are the verbs [áki-dó] ‘to produce a child’ and [áki-dóŋ] ‘to castrate’. As indicated by the underlined [-ATR] vowels, examples of spreading to a maximum of three syllables appear in (64a,b), although spreading is not required to extend to the limit of this domain.

- (64) Maximum domain for DOMAIN [r]/[-ATR]: (Novelli 1985: 225–226, 411)
- | | |
|--|---|
| <p>a. áki-do ‘to produce a child’</p> <p>ákídó-òr</p> <p>áki-zi-dó-òr</p> <p>ákídó-ón-òr</p> <p>ákizidó-ónòr</p> <p>ákidoíðó-òr</p> <p>ákidoíðó-ón-òr</p> <p>ákidoíðó-ón-òr-játa</p> | <p>b. áki-dóŋ ‘to castrate’</p> <p>ákídóŋ-òr</p> <p>ákizidóŋ-òr</p> <p>ákídóŋ-ón-òr</p> <p>ákizidóŋ-ón-òr</p> <p>ákidoŋíðóŋ-òr</p> <p>ákidoŋíðóŋ-ón-òr</p> <p>ákidoŋíðóŋ-ón-òr-játa</p> |
|--|---|

At this point of the analysis no ranking argument can be formulated that crucially ranks IDENT-IO SPEC SUFFIX [α ATR] over DOMAIN [r]/[-ATR] or vice versa. The two constraints are ranked above AGREE [α ATR, Hd, PW], as shown in this ranking tableau of the derivation of /áki-dóŋ-An-òr/ ‘to castrate frequently (away)’ in (65). For the itive suffix in this derivation the [-ATR] specification from the rhotic is considered the underlying [ATR] specification of the morpheme, and the [-Ar] suffix is therefore represented as [-òr] in the input representation. Candidate (c) respects faithfulness to the specification of the suffix but not the stipulated domain of the consonant-generated feature, therefore showing a root-controlled spreading process with the itive suffix excluded. This also incurs multiple violations of AGREE [α ATR, Hd, PW], as the [-ATR] specification in the dominant PW node does not agree with the specifications of the heads of the root and its immediately adjacent affixes.

The itive suffix has a separate intermediate prosodic word node with its [-ATR] feature, and the remaining affixes share the same intermediate PW node with the root. The dominant PW node receives percolation of the [-ATR] feature. Candidate (b) allows the [+ATR] root specification to percolate to the dominant PW node to override the [-ATR] specification on the itive suffix, crucially violating IDENT-IO SPEC SUFFIX [α ATR] and showing exhaustive spreading through-out the PW. The candidate therefore lacks violations of AGREE, and incurs one violation of IDENT-IO [α ATR] through the feature-filling process carried out in the frequentive suffix [-An]. The winning candidate and attested form respects the [-ATR] specification of the specification of the root, and remains so as it is outside of the domain of the suffix-controlled [-ATR] spreading process. The phonological head of the itive suffix assumes a dominant position in the PW hierarchy, through which the [-ATR] feature percolates.

The C/[+ATR, hi] adjacency effect discussed in §1.2 examples (7) and (8), in which the [+ATR] feature generated by supraglottal cavity enlargement for the prolongation of voicing, is governed by the constraint proposed in (66) as part of the DOMAIN constraint family proposed in (63).

(65) IDENT-IO SPEC SUFFIX [α ATR], DOMAIN [r]/[-ATR] >> AGREE [α ATR,Hd,PW]

/ákI-dəŋ-An-ɔr/ 'to castrate frequently	IDENT-IO SPEC SUFFIX [-ATR]	DOMAIN [r]/[-ATR]	AGREE [α ATR,Hd,PW]	IDENT-IO ROOT[α ATR]	IDENT-IO [α ATR]
<p>a. aki-dəŋ-ɔn-ɔr</p>			*	*	**
<p>b. aki-dəŋ-on-ɔr</p>	*	****			**
<p>c. aki-dəŋ-on-ɔr</p>		**!	***		**

(66) Domain C/[+ATR, hi] adjacency effect

DOMAIN C/[+ATR, hi]: The local domain D of the [+ATR] feature generated on the high vowel by voiced obstruent C is N phonological heads to the right of the consonant, where $N \in \{1,2\}$.

This constraint is crucially ranked over DOMAIN[r]/[-ATR] as it blocks the harmony domain of the rhotic-generated [-ATR] feature. This can be seen in the tableau in (67). Candidate (a) shows alternation of all vowels within the domain of the rhotic-generated [-ATR] spreading show [-ATR] values, which crucially violates DOMAIN C/[+ATR, hi]. Candidate (c) completely overrides the [-ATR] suffix specification and crucially violates IDENT-IO SPEC SUFFIX [α ATR]. Winning Candidate (b), respects the specification of the [-ATR] suffix, but overrides the extension of its domain through the action of the C/[+ATR, hi] adjacency effect. The tableau is shown without the prosodic structure.

The dominance of DOMAIN C/[+ATR, hi] over DOMAIN [r]/[-ATR] explains data showing the [+ATR] causative morpheme blocking the spreading of the rhotic-generated [-ATR] feature. As noted by Novelli (1985) and Lesley-Neuman (2007), the causative morpheme /-tA-/ in [+ATR] verbs appears obligatorily with [+ATR] vowel [-i], which appears as well with voicing and spirantization of the /t/: $t \rightarrow z$. The anterior and markedly higher position of the tongue used to pronounce these segments sharply counters the lowered and retracted position of the rhotic, and of the [-ATR] spreading from the phonologized co-articulation effect. It is proposed that the voicing of the consonant [z] produces an additional [+ATR] feature that blocks the [-ATR] spreading domain that would otherwise remain unaffected.

(67) DOMAIN C/[+ATR, hi] >> DOMAIN [r]/[-ATR]

/akI-bu-or /	IDENT-IO SP SUFX [αATR]	DOMAIN C / [+ATR,hi]	DOMAIN [r]/[-ATR]	AGREE [αATR, Hd,PW]	IDENT-IO ROOT [αATR]	IDENT- IO [αATR]
a. aki-bu-or		*!			*	**
☞ b. aki-bu-or			**	*		*
c. aki-bu-or	*!					*

The tableau supporting this assertion appears in (68). Candidate (c) crucially violates the domain stipulations of the rhotic-generated [-ATR] feature, instead allowing root-controlled harmony. Candidate (b) allows the full expression of the domain of the rhotic-generated [-ATR] feature, which violates the constraint DOMAIN C/[+ATR,hi] by allowing the vowel within the causative morpheme to be [-ATR]. This form is also unattested. Winning candidate (a) accommodates all spreading processes and adjacency effects in their proper domains, and is the attested form.

Another effect to be accounted for in the constraint ranking is that of the dissimilation in [ATR] specification among central and back vowels, which was discussed in §1.2 example (10). Before [-ar] suffixation in [-ATR] verbs, the central and back vowels immediately preceding the suffix verbs raise slightly and assume a [+ATR] feature. Vowel [ɔ] becomes [u], and [ɔ] becomes [o]. The constraint in (69) prohibits a sequence of [-low] central and back vowels with the same [-ATR] specification as the following low vowel.

(68) C/[+ATR, hi] Adjacency effect in [+ATR] causative forms

/akI-zi-doŋ-or/	IDENT-IO SPEC SUFX	DOMAIN C/[+ATR,hi]	DOMAIN [r]/[-ATR]	AGREE [αATR,Hd,PW]
a. aki-zi-doŋ-or			*	**
b. aki-zi-doŋ-or		*!		*
c. aki-zi-doŋ-or			**!	

(69) [+back] [ATR] Dissimilation

* V / ___ + V C : Vowels which are [-low] and [-ATR] at a morpheme boundary followed by a vowel that is [+low] and [-ATR] are prohibited.

(-low)	(+low)
-ATR	-ATR
+back	+back

Ensuring the attested output form in Karimojong involves constraints on the various alternative processes that may be employed by the language in face of the prohibition on the sequence. Among them are metathesis, prohibited by LINEARITY-IO, which is presented in (70), deletion, controlled by MAX-IO in (71), and faithfulness to the feature [low] which is governed by IDENT-IO [low], presented in (72). A demonstration of the combined action of these constraints to generate this conspiracy appears in the tableau in (73), though with the prosodic structure is omitted for ease of exposition.

(70) LINEARITY-IO: The output reflects the precedence structure of the input segments and vice-versa. (Pater 1996)

(71) MAX-IO: Input segments must have output correspondents. (McCarthy & Prince 1995)

(72) IDENT-IO [low]: Any correspondent of an input segment specified as [+low] must also be [+low] in the output. (Pater 1996)

Example (73) shows through the Level 1 the derivation of [adoar] ‘to snap one’s fingers (itive), the constraint conspiracy that permits the attested form. Candidate (d) crucially violates the back vowel [ATR] dissimilation constraint. Candidate (a) avoids this violation through metathesis, but crucially violates LIN-IO. Candidates (b) and (c) likewise avoid the violation through deletions that crucially violate MAX-IO. Candidate (f) avoids the prohibited sequence through a change in the [+ATR] feature of the suffix vowel, incurring a violation of AGREE [α ATR, Hd, PW]. The height specification of the first vowel in the sequence of Candidate (g) complies with this prohibition, but crucially violates the faithfulness constraint IDENT-IO [low]. Winning Candidate (e) effects the vowel change from [ɔ] to [o].

As Karimojong is a language with numerous vowel deletions and processes of coalescence, MAX-IO and LINEARITY-IO must have rankings that are low in comparison with the constraints governing moraic structure, the licensing of voiceless vowels and vowel sequences. Presentation of the relevant constraints governing these processes, and analyses of them are beyond the scope of this discussion.

(73) [+back]/ ATR Dissimilation effect

Level 1 /-dɔ-Ar/	LIN-IO	MAX-IO	IDENT-IO [low]	*V / __ + VC -low +low -ATR -ATR +back +back	AGREE [α ATR, Hd, PW]	IDENT-IO ROOT [α ATR]	IDENT-IO [α ATR]
a. -daɔr	*!						
b. -dor		*!					
c. -dɔr		*!					
d. -dɔar				*!			*
e. -doar					*	*	*
f. -dɔor			*!		*		*
g. -daar			*!				

3.5 Constraint ranking by morpho-phonological level

The overall ranking of constraints governing [ATR] harmony processes and adjacency effects in Karimojong under a stratal model is presented in (74). It can be noted that the number of constraints decreases at the higher levels.

(74) Constraint hierarchy for [ATR] harmony in Karimojong

- Level 1: INTERPRETABILITY [ATR] >> LIN-IO, MAX, IDENT-IO[low]>>

-low	+low
-ATR	-ATR
+back	+back
- IDENT-IO SPEC SUFFIX [ATR], DOMAIN [r]/[-ATR] >> AGREE [α ATR, Hd, PW] >> IDENT-IO ROOT [α ATR] >> IDENT-IO[α ATR]
- Level 2: IDENT-IO SPEC SUFFIX[α ATR] >> AGREE [α ATR, PW]>> IDENT-IO ROOT [α ATR] >> IDENT-IO [α ATR]
- Level 3: IDENT-IO [α ATR]

3.6 The failure parallelism and its alternatives

The failure of the simple parallel optimality-theoretic model can be demonstrated by the inclusion of a variety of affixation processes, including for those affixes transparent to harmony, in a one-step input-output parallel tableau. This is shown in (75). The attested output form, Candidate (a), cannot be a product of the application in parallel of all of the constraints in the hierarchy, as the [+ATR] specification of the vowels in the rightmost morpheme do not agree with the specification on the

dominant PW node, crucially violating the AGREE constraint. Candidate (b) crucially violates INTERPRETABILITY by permitting a vowel segment unspecified for [ATR] to surface. Candidate (c) incurs three violations of faithfulness to suffix vowel input specifications, in the input had [+ATR] specifications and at the output [-ATR] values. Winning candidate (d), in which a [+ATR] feature spreads to every phonological head, is unattested.

This leads to the proposal that the parallel derivation model cannot account for Karimojong output forms, an assertion remaining unproven until alternative parallel strategies have been analyzed and eliminated from consideration. These strategies include Sympathy Theory, Output-Output Correspondence, and Uniform Exponence, all of which are discussed in the sub-sections that follow.

3.6.1 Sympathy Theory

(75) Parallel model tableau

/ákI-tA-dəŋ-eenen/	INTERP [ATR]	IDENT-IO SPEC SUFX [αATR]	AGREE [αATR, Hd, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
<p>a. ákitədəŋ-eenen</p>			***!		**
<p>b. ákitOdəŋ - eenen</p>	*!				
<p>c. ákitədəŋ-eenen</p>		***!			****
<p>d. ákitodəŋ-eenen</p>				*	***

Sympathy Theory is employed where forms analyzed under serial derivation do not appear as winning candidates in parallel tableaux. The idea behind it is that EVAL selects a sympathetic candidate that is the most harmonic in obeying a specified constraint, called a *Selector Constraint*. Rankable constraints require the output form to resemble the sympathetic candidate, which is available to influence the output form. One or more sympathy constraints are used to determine the optimal candidate. Sympathy constraints define faithfulness to a sympathy candidate, (marked with ⊕ in tableaux) which is the optimal candidate obeying the selector constraint (marked with ★). The standard version of this constraint is one that was crucially violated by the attested output form under parallel derivation. This higher ranking constraint allows the previously failed attested output form to be the winning candidate. A Sympathy tableau based on the parallel derivation in (75) appears in (76).

Sympathy Theory raises a series of concerns. The first among them is that when the theory is applied in the present case, a markedness constraint must be employed as the selector. The theory had been restricted to apply only to cases of input-output faithfulness under the principle of Selector Confinement (McCarthy 1999: 339) presented in (77). For Sympathy to apply to the case shown in (76), modifications to the original proposal must be made.

A second concern has been raised by Bermudez-Otero (2003), Green (2004), and Kiparsky (2007) who point out that the generation of a sympathy candidate and its incorporation into the evaluation process constitutes a derivational step, while a Sympathy Theory tableau, within the logic of its formalism, claims otherwise. Bermudez-Otero (2003: 3), using the computation in (78), asserts that a Sympathy process minimally mimics a derivational step.

(76) Sympathy theory approach to parallel derivation

/akI-tA-dəŋ-eenen/	★AGREE [αATR, Hd, PW]	INTERP [ATR]	IDENT-IO SPEC SUFFIX [αATR]	AGREE [αATR, Hd, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
<p>a. ákítədəŋ-eenen</p>				***		**
<p>b. ákítAdəŋeenen</p>		*!		***		*
<p>c. ákítədəŋ-eenen</p>			***!			*****
<p>d. ákítədəŋ-eenen</p>	***!				*	***

(77) ★-Confinement

The selection of a sympathetic candidate must be confined to a subset of candidates that obey an IO-faithfulness constraint F.

(78) Equivalence of parallel sympathy theory and serial derivation

$$I \rightarrow a \rightarrow O \approx I \rightarrow \odot \rightarrow O$$

A third issue is the plausibility of claiming sympathy constraints to be part of the universal constraint set, when they are so specifically tailored to the evaluation of individual forms. The number of cases of opacity across the world's languages is extensive, and no explanation has been made as to how constraints simultaneously so specific, numerous and varied in type could be said to represent universals, which are tendencies verified under careful investigation to occur cross-linguistically. The

ad hoc creation of constraints to explain cases of opacity is minimally a contradiction in methodology. Nevertheless, no studies regarding the implications of Sympathy Theory for universals research have been carried out. Such studies would have to calculate the number of constraints/universals that the theory would potentially generate, a number anticipated to be so unreasonably large that some rationalization or modification to the theory would necessarily be proposed.

This raises a fourth argument involving learnability. Every active input-output faithfulness constraint can potentially be a selector constraint. A learner would contend with a large set of potential \otimes -candidates. Dinnsong and colleagues (2000) invoke the Elsewhere Condition to propose that the most specific candidate be chosen, but no description or algorithm defines how this would occur. Bermudez-Otero (2003) maintains that stratal OT models offer a more coherent view of learnability; this work is to be consulted for more detailed argument.

Given its equivalence to a derivational step, the specificity as opposed to the universality of its constraints, and the need to apply to markedness rather than faithfulness constraints to account for the Karimojong data, Sympathy Theory presents an implausible argument in favor of parallel derivation. As it resembles ad hoc and brute force argumentation, it has been for the most part abandoned as an analytical strategy in defense of parallelism in OT.

3.6.2 Output-output correspondence and uniform exponence

Output-Output Correspondence (OOC) constraints are tightly focused on the relationship between the base and the derived form, strictly adhering to the principle of the base including a smaller number of features than the derived form. The formal definition of a base under OOC follows in (79).

- (79) Base (Kager 1999)
- a. A base is a free-standing output form—a word.
 - b. The base contains a subset of grammatical features of the derived form.

OOC is effected in Karimojong through correspondence between the base and an affixed form, and involves faithfulness to the [ATR] specifications of the base. This requires the faithfulness constraint proposed in (80).

(80) Base-affix identity in [ATR] specification

IDENT B-A [α ATR]: Let γ be a vowel in the base and β a correspondent in the affixed form. If γ is [α ATR], then β is [α ATR].

The case of a verb form containing a pronominal prefix is shown in (81). The form [ε -l \acute{o} -t \acute{e} re] ‘he has been dried in the sun’ contains the neutral pronominal prefix [ε] ‘he’ underlyingly specified as [–ATR]. The input / ε -l \acute{o} -t \acute{e} re/ has the root vowel / ω / underlyingly specified as [–ATR]. Dominant TMA marker [-t \acute{e} re] has a [+ATR] specification. The base in this case, [ε -l \acute{o} - δ] ‘he was drying in the sun’ also has a dominant TMA marker, and shows the pronominal prefix maintaining its underlying [–ATR] specification. Under AGREE and IDENT I-O SPEC SUFFIX, the root vowel assumes the suffix [ATR] value, and, in agreement with the base, the [ATR] specification of the pronominal prefix vowel remains unchanged for optimal Candidate (a). Candidate (b) crucially violates the high-ranking OOC constraint, IDENT-IO B-A [α ATR], by maintaining faithfulness to the root as opposed to faithfulness to the base. Candidate (c) also violates this high-ranking constraint by allowing the pronominal prefix to harmonize with the [+ATR] suffix.

(81) Base-affix identity with neutral pronominal prefix

ε-ló-tére ‘He has been dried in the sun (indicative/subjunctive mood, form A)’

(Novelli 1985: 239)

Input: /ε-ló-térè/ Base: [ε-ló-ò]	IDENT-IO B-A [αATR]	IDENT-IO SPEC SUFX [αATR]	AGREE [αATR, Hd, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
<p>a. ε-ló-tére</p>			*	*	*
<p>b. ε-ló-tére</p>	*!	**			**
<p>c. e-ló-tére</p>	*!			*	**

The example in (81) shows the most common case in Karimojong—that in which the pronominal prefix is neutral to [ATR] harmony processes. If Output-Output Correspondence has validity, the constraint ranking should follow the Principle of Globality—in that it should be valid for forms across the language. However, quite different results are obtained in the Narrative Mood, where the pronominal prefix is underlyingly underspecified. Example (82) shows the case of the Form A Narrative Mood past perfect of [-dɔŋ] ‘pinch’, [*ító-dɔŋ-ítetèi*], in which the dominant [+ATR] TMA marker [-(*i*) *tetèi*] is affixed, but causes the pronominal prefix to alternate. The paradigm base of is the simple past form [*ító-dɔŋ-ì*], in which the [+ATR] TMA marker [-*i*] remains neutral. Under a serial model, the neutral TMA marker would be affixed on a separate level in which [ATR] harmony rules do not apply. An OOC constraint ranking using this as the base form produces pathological results, as Candidate (d), the attested output form with exhaustive [+ATR] spreading, violates the high-ranking OOC constraint by not retaining the [-ATR] values of the root and the prefix vowels of the base. Candidate (a), allowing for suffix dominance and neutrality, and thereby representing the pattern of the great majority of Karimojong forms, violates the OOC constraint. Candidate (c) crucially violates suffix specification faithfulness, and the unattested optimal candidate is that which respects the [ATR] values of the root and prefix of the base form and faithfulness to the suffix [ATR] specification.

Its inconsistency across paradigms and the stipulations of Principle of Globality eliminate OOC as an explanation for affix neutrality. There is also no principled way to determine which of the more than 140 pronominal prefixes serves as the base model for the remaining non-alternating forms. Even if all pronominal prefixes were neutral, an OOC argument at best would be circular: the suffix does not alternate in construction Y because it does not alternate in construction X. It provides no explanation for neutrality in construction X. This also holds for [-*eenen*], and the variable behavior of the TMA marker [-*ere*].

- (82) Base-affix identity with pronominal prefix alternation
 [iɬo-dɔŋ-itetèi] ‘3s/pl had been pinched’ (form A) (Novelli 1985: 272)

Input:/iɬA-dɔŋ- itetèi/ Base: [iɬo-dɔŋ-i]	IDENT-IO B- A [αATR]	IDENT-IO SPEC SUFFIX [αATR]	AGREE [αATR, Hd, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
<p>a. iɬo-dɔŋ-itetèi</p>	*!		*		
<p>b. iɬo-dɔŋ- itetèi</p>			**		
<p>c. iɬo-dɔŋ-itetèi</p>		***!			***
<p>d. iɬo-dɔŋ-itetèi</p>	*!			*	**

Similar to OOC is the analytical framework of Uniform Exponence, which provides a more general format that need not rely on the concept of a base. Its formal statement follows in (83).

- (83) Uniform Exponence a lexical item (stem, affix, word) has the (Kenstowicz 1996)
 (UE): same realization for property P in its various
 contexts of occurrence.

Applying UE to neutral affixes in Karimojong involves a constraint prohibiting the harmonic alternation of pronominal prefixes, the [-eenen] frequentive, and neutral TMA markers. This Uniform Exponence constraint is labeled UE [*Harmony]. The pronominal prefix case is exemplified in (84). Optimal candidate (a) is the attested output form in which the principle of Uniform Exponence mandates the non-alternation of pronominal prefix [ɛ-] under suffix-controlled [+ATR] spreading governed by AGREE [αATR, Hd, PW] and IDENT-IO SPEC SUFFIX [αATR]. Candidate (b) crucially violates the UE constraint by allowing the pronominal prefix to receive a [+ATR] specification. Here the UE constraint assures the neutrality of the pronominal prefix. Under the Principle of Globality, this same constraint set should yield attested output forms throughout the language.

(84) Uniform Exponence and pronominal prefix neutrality

/ε-dɔŋ/ + /jo/	UE [*Harmony]	IDENT-IO SPEC SUFX [αATR]	AGREE [αATR, Hd, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
<p>a. ε- doŋ- jo</p>			*	*	*
<p>b. e- doŋ- jo</p>	*!			*	**

However, its application to a narrative mood derivation produces pathological results. In (85), the attested output form, Candidate (a), which has a pronominal prefix alternating under suffix-controlled [+ATR] harmony, crucially violates UE [*Harmony]. The optimal candidate is an unattested form that maintains pronominal prefix neutrality under an otherwise dominant harmony process.

A similar phenomenon involves the continuous aspect marker [-ere] and dative suffix [-Akin]. In (88), the input form /aka-dɔŋ-Akin-ere/ ‘I am pinching (for the purpose of)’ has a [-ATR] root, a non-alternating pronominal prefix, the dative suffix [-Akin] with an underspecified head, and a [+ATR] TMA marker. Stipulations of the highly-ranked UE constraint keep both the pro-nominal prefix and the TMA marker inactive.

(85) Uniform Exponence with the alternating pronominal prefix

itɔ-dɔŋ+/tetèi/	UE [*Harmony]	IDENT-IO SPEC SUFX [αATR]	AGREE [αATR, Hd, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
<p>a. itɔ- doŋ- tetèi</p>	*!			*	**
<p>b. itɔ- doŋ- tetèi</p>			*	*	*

The attested form, Candidate (a), is optimal. Candidate (c) crucially violates the UE constraint by allowing [+ATR] spreading from the TMA marker and pronominal prefix alternation. Candidate (b) does likewise by respecting UE for the prefix only. Candidate (d) violates UE and suffix specification

faithfulness by altering the [ATR] specification of [-ere]. It incurs three violations of general [ATR] specification faithfulness. Prosodic structure is not shown in the tableau for ease of exposition.

(86) Uniform exponence and the continuous aspect

/aka-dóŋ-Akin-ere/	UE [*Harmony]	IDENT-IO SPEC SUFX [αATR]	AGREE [αATR, Hd, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
☞ a. aka-dóŋ-ákin-ere			**		*
b. aka-dóŋ-ókin-ere	**!		**	*	**
c. aka-dóŋ-ókin-ere	**!			*	**
d. aka-dóŋ-ákin-ere	*!	**			***

It could be argued the dative suffix [-Akin] blocks harmony. As across the language, this suffix has never appeared as opaque in any other instance, so a UE constraint stipulating its opacity would crucially fail. Nonetheless, it is shown in (87) that a UE constraint based on the neutrality of [-ere] also fails, since it appears as an active suffix controlling [+ATR] spreading. The derivation of the input form /aka-dóŋ-eré/ ‘I was pinching’ shows the attested form, Candidate (a), with a [+ATR] feature spreading from the suffix, crucially violating UE. Optimal Candidate (b), allows neither spreading from the TMA marker nor the alternation of the pronominal prefix. Candidate (c), by allowing both dominant [+ATR] spreading and pronominal prefix alternation, crucially violates UE [*Harmony] twice. In all forms, the UE constraint controls the neutrality of [-ere]. Hence, the variability in [ATR] harmony participation of a single TMA marker shows that UE cannot apply to all output forms as the Principle of Globality requires.

(87) Pathological results with UE [*harmony]

/εke-dóŋ-eré/	UE [*Harmony]	IDENT-IO SPEC SUFX [αATR]	AGREE [αATR, Hd, PW]	IDENT-IO ROOT [αATR]	IDENT-IO [αATR]
☹ a. aka-dóŋ-eré	*!		**		*
☞ b. aka-dóŋ-eré			**	*	**
c. aka-dóŋ-eré	***!			*	**

As Output-Output Correspondence and Uniform Exponence cannot work as explanations for affix neutrality, they likewise cannot justify a parallel optimality-theoretic model for Karimojong. Headed Spans was eliminated in §2.1.4 as inviable. It is a stratal model of optimality-theoretic constraints that best explains attested forms and historical change. Future research will determine how a historical lexical phonology model corresponds to lexical storage and the physical processing of language.

5 Conclusions

Eastern Nilotic vowel harmony processes present a number of challenges to the optimality-theoretic framework in phonology. The fact that these languages possess two active harmony domains and a significant presence of neutral morphemes, while experiencing surface irregularities due to local effects, nullifies the bases of most of the constraint systems that have been proposed thus far, in that these stipulate domains of regular and exhaustive spreading. This paper has proposed a constraint system that addresses phonological structure internal to the morpheme, as well as constraints that handle local effects. This system meets the high standards set by the theoretical parameters of OT in that it obeys the Principle of Globality: the constraints work for forms across the language. Further work should focus on constructing a whole language phonological model which more clearly defines the relationships between the constraints and the modules of the grammar and their interfaces that they address.

Abbreviations

A	Verb Form A	L	Left
ACT	Active Voice	LIN	Linearity Constraint
AGREE	Agree Constraint	LO	Low
ALIGN	Alignment	MAX	Maximality Constraint
APPL	Applicative	NARR	Narrative Mood
ATR	Advanced Tongue Root	ODT	Optimal Domains Theory
B	Verb Form B	OOO	Output-Output Correspondence Constraint
BA	Basic Alignment	OT	Optimality Theory
B-A	Base-Affix	PASS	Passive Voice
CAUS	Causative	PFEAT	Prosodic Feature
DOMAIN	Domain Constraint	PRF	Perfect Aspect
EN	Eastern Nilotic	PROG	Progressive
F	Feature	PRS	Present Tense
FREQ	Frequentive	PST	Past Tense
FUT	Future	PW	Prosodic Word
GCAT	Grammatical Category	R	Right
HD	Head	REFL	Reflexive
HI	High	RTR	Retracted Tongue Root
IDENT	Identity Constraint	SPEC	Specified
IMPF	Imperfective Aspect	SUBJ	Subjunctive
IND	Indicative	TMA	Tense, Mood, Aspect
INF	Infinitive	UE	Uniform Exponence Constraint
INTERPRET	Interpretability Constraint	VEN	Ventive
IO	Input-Output	WSA	Wide Scope Alignment
IT	Itive		

References

- Akinlabi, Akinbiyi. 1994. Alignment Constraints in [ATR] Harmony. *Studies in the Linguistic Sciences* 24.1: 1–18.
- Archangeli, Diana B. & Douglas G. Pulleyblank. 1994. *Grounded Phonology*. Cambridge: MIT Press.
- Archangeli, Diana B. & Terry Langendoen. 1997. *Optimality Theory: An Overview*. Oxford: Blackwell.
- Bakovic, Eric. 2000. Harmony, dominance and control. PhD dissertation, State University of New Jersey.
- Bakovic, Eric. 2002. *Vowel Harmony and Stem Identity*. Rutgers Optimality Archive #540-0902.
- Beckman, Jill. 1999. *Positional Faithfulness: An Optimality-Theoretic Treatment Phonological Asymmetries*. New York: Garland Publishing.
- Bermudez-Otero, Ricardo. 1999. Constraint interaction in language change (opacity and globality in language change). PhD dissertation, University of Manchester/Universidad Santiago de Compostela.
- Bermudez-Otero, Ricardo. 2002. "Cyclicality or sympathy? A case study." Handout of paper presented at the 10th Manchester Phonology Meeting May 23–25, 2002.
- Bermudez-Otero, Ricardo. 2003. *The Acquisition of Phonological Opacity*. Rutgers Optimality Archive #593-0403.

- Boersma, Paul. 2011. A programme for bidirectional phonology and phonetics and their acquisition and evolution. In Anton Benz & Jason Mattausch (eds.), pp. 33–72. *Bidirectional Optimality Theory*. Amsterdam/Philadelphia: Benjamins.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.
- Cole, Jennifer S. 2009. Emergent feature structures: harmony systems in exemplar models of phonology. *Language Sciences* 31: 144–160.
- Cole, Jennifer & Charles Kisseberth. 1994. An optimal domains theory of harmony. *Studies in the Linguistic Sciences* 24.2: 1–13.
- Dimmendaal, Gerrit J. 1983. *The Turkana Language*. Dordrecht: Foris Publications.
- Dimmendaal, Gerrit J. 2002. Constraining disharmony in Nilotic: What does an optimal system look like? *Journal of African Language and Linguistics* 23.2: 153–181.
- Dimmendaal, Gerrit J. 2011. Metrical structures: A neglected property of Nilotic (and other African language families). MS, University of Cologne.
- Gordon, Matthew. 1998. The phonetics and phonology of non-modal vowels: A cross-linguistic perspective. *Berkeley Linguistics Society* 24: 93–105.
- Green, Antony D. 2004. Opacity in Tiberian Hebrew: morphology, not phonology. *ZAS Papers in Linguistics* 37: 37–70.
- Guion, Susan G., Mark W. Post & Doris L. Payne. 2004. Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics* 32: 517–542.
- Hall, R.M.R. & Chet A. Creider. 1998. The fates of [+ATR] /a/ in Nilotic. In Ian Maddieson & Thomas J. Hinnebusch (eds.), *Language History and Linguistic Description in Africa*, pp. 45–54. Trenton, N.J./Asmara, Eritrea: Africa World Press.
- Hermans, Ben & Marc van Oostendorp. 1999. *The Derivational Residue in Phonological Optimality Theory*. Amsterdam/Philadelphia: Benjamins.
- Idsardi, William. 1998. Tiberian Hebrew spirantization and phonological derivations. *Linguistic Inquiry* 29.1: 37–73.
- Kager, René W. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kiparsky, Paul. 2007. Paradigm effects and opacity. MS, Stanford University.
- Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291–306.
- Lesley-Neuman, Diane. 2007. [ATR] Harmony in Karimojong: justification for a stratal optimality theory. MA thesis, University of Colorado at Boulder.
- Lesley-Neuman, Diane. 2009. The DOMAIN C constraint family. MS, Michigan State University.
- Lesley-Neuman, Diane. 2012. Morpho-phonological levels and grammaticalization in Karimojong: A review of the evidence. *Studies in African Linguistics* 41.1: 99–169.
- Mantovani, Mario A. 1963. *An Introduction to the Karimojong Language*. Gulu, Uganda: Catholic Press.
- McCarthy, John J. 1999. Sympathy and Phonological Opacity. *Phonology* 16.3: 331–399.
- McCarthy, John J. 2003. Sympathy, cumulativity and the Duke-of-York gambit. In C. Féry and R. van de Vijver (eds.), *The Syllable in Optimality Theory*, pp. 23–76. Cambridge: Cambridge University Press.
- McCarthy, John J. 2004. Headed spans and autosegmental spreading. MS, University of Mass. at Amherst. ROA #685.
- McCarthy, John J. 2008. *Doing Optimality Theory: Applying Theory to Data*. Oxford/Malden, MA/Carleton, Victoria: Blackwell.
- McCarthy, John J. 2009. Harmony in harmonic serialism. MS, University of Massachusetts at Amherst.

- McCarthy, John J. 2010. *An Introduction to Harmonic Serialism*. Available at: http://works.bepress.com/john_j_mccarthy/103.
- McCarthy, John J. & Alan Prince. 1995. Faithfulness and Reduplicative Identity. In Jill Beckman, Laura Walsh-Dickey & Suzanne Urbanczyk (eds.), *University of Massachusetts Occasional Papers in Linguistics* 18: 249–384.
- Noske, Manuela. 1990. Vowel harmony in Turkana. *Studies in the Linguistic Sciences* 20.1: 123–134.
- Noske, Manuela. 1991. Metrical structure and reduplication in Turkana. In M. Lionel Bender (ed.), *Proceedings of the Fourth Nilo-Saharan Linguistics Colloquium*, pp. 245–262. Hamburg: Buske.
- Noske, Manuela. 1996. [ATR] harmony in Turkana. *Studies in African Linguistics* 25.1: 61–98.
- Noske, Manuela. 2000. [ATR] harmony in Turkana: a case of faith suffix >> faith root. *Natural Language and Linguistic Theory* 18: 771–812.
- Novelli, Bruno. 1985. *A Grammar of the Karimojong Language*. Berlin: Reimer.
- Pulleyblank, Douglas. 1997. Optimality theory and features. In Diana Archangeli & Terence Langendoen (eds.), *Optimality Theory: An Overview*, pp. 59–101. London: Blackwell.
- Sasa, Tomomasa. 2009. Treatment of vowel harmony in optimality theory. PhD dissertation, University of Iowa.
- Schröder, Helga & Martin C. Schröder. 1984. The Toposa verb. *Occasional Papers in the Study of Sudanese Languages* 5: 1–47.
- Trigo, Loren. 1991. On pharynx-larynx interactions. *Phonology* 8: 113–136.
- Vannest, Jennifer & Julie E. Boland. 1999. Lexical morphology and lexical access. *Brain and Language* 68: 324–332.

BERTA AND THE EAST JEBEL SUBFAMILY: REINVESTIGATING A NILO-SAHARAN ISOLATE¹

Nate D. Bremer

1 Introduction: Berta in context

Since the revision of Greenberg's (1963) "Chari-Nile" family, scholars have speculated Berta's relationship within the larger Nilo-Saharan phylum. Evans-Pritchard's (1932) work is one of very few to suggest a group of Berta languages, however his classification was later challenged and diffused by Bender (1983, 1989). Bender's reanalysis of Evans-Pritchard (1932), coupled with his own research, led to the creation of the East Jebel (E4) subfamily of Eastern Sudanic. The grouping together of Gaahmg, Aka, Kelo, Molo and later Beni-Sheko, subsequently led to Berta becoming an isolate.

While Bender's classification still has merit, a recent dialect survey (Bremer, forthcoming) done in Ethiopia in 2011, when combined with the historical Berta record reveals notable phonetic, lexical and semantic similarities between Berta and the East Jebel (E4) languages. Research on Berta dialectal variation has been a longstanding deficiency, evidenced by the fact that the majority of known Berta typology has focused on a single variety, Maiyu. Beyond this populous dialect, however, prominent lexical dissimilarities exist. This paper takes some of these features into consideration and revisits the historical record and the case of Berta's genetic affiliation with the East Jebel family. Commonalities emerge between each language's lexicon as well as a high percentage of Berta cognates with the larger proto-Eastern Sudanic isoglosses. These findings suggest that the link between these languages is perhaps stronger than once thought, and thus Berta's inclusion within this family ought to be reconsidered.

Section I examines Berta's historical context, spotlighting existing literature useful for historical reconstruction. Section II focuses on phonological, lexical and semantic shift in Berta. This section begins with an investigation of the phonetic inventories of six different contemporary Berta dialects, leading to a hypothesis about the proto-Berta inventory. Continuing the topic of language shift, lexical and semantic variation are explored and isoglosses are noted. Section III extends beyond Berta and explores some of the commonalities with the East Jebel languages which, among other structures, includes negation, pronouns and interrogatives. Section IV examines the figures generated by lexicostatistics and offers suggestions for further research. Section V concludes the paper.

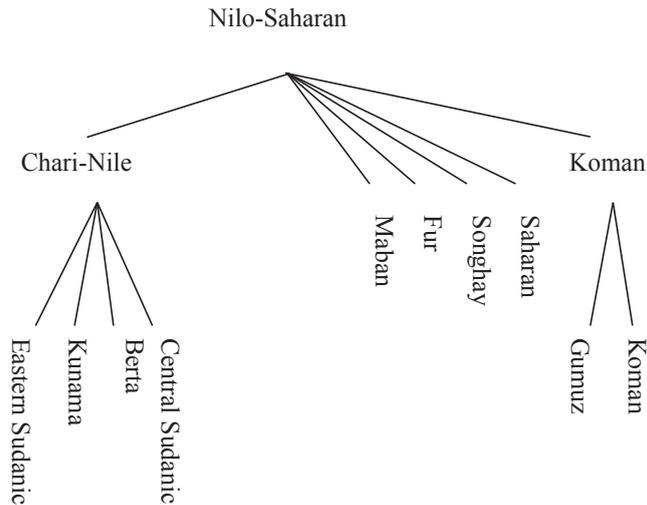
1.1 Placement within the Nilo-Saharan phylum

Placing Berta [wti] in a genetic context is not easy. To quote the late Lionel Bender (1989: 271), "Although great progress has been made in Nilo-Saharan studies since mid-century when Greenberg began his break-through classificatory work, there are still many neglected languages and branches. One of these is Berta." Ever since Greenberg's ground-breaking 1963 analysis in which he first proposed the classification of what has become the four African phyla, Berta's genetic relationship amidst the wider Nilo-Saharan (N-S) phylum has always been puzzling. Although Greenberg's original proposal has largely withstood the test of time, a major amendment to his original proposal has been the elimination of the family called "Chari-Nile", and the subsequent promotion of its members (Central Sudanic, Eastern Sudanic, Berta and Kunama) to families or isolates. Bender proposed this change in 1976, and it is arguably the most widely accepted model today (c.f. Blench 2000). Where previously Berta was an isolate related to the now-dissolved Chari-Nile family, it

¹ Berta is arguably not a true isolate since it typically falls within a classified super-family. For the purposes of this paper, however, since it does not belong to a family, the term isolate is used in a generic sense.

currently stands in isolation, genetically linked to the N-S phylum through what is labeled the “Satellite-Core” group (for a visual depiction, compare Figure 1 with Figure 2 below). Greenberg was not adverse to this new classification, and he “now considers Central and Eastern Sudanic to be independent branches of N-S, with Berta and Kunama ‘closer to East Sudanic’” (Bender 1996: 60). This paper seeks to reexamine the notion of Berta’s proximity to the Eastern Sudanic family and propose that lexical and typological similarities suggest more than mere language contact, but actual genetic relatedness. But first a quick look at the historical development of N-S classification is warranted.

Figure 1: Greenberg’s (1963) Nilo-Saharan structure



Juxtaposing two contemporary Nilo-Saharan family trees proposed by Bender (Figure 2) and Ehret (Figure 3) visibly illustrates that scholars have radically divergent views concerning genetic relatedness within the phylum. Berta’s hierarchical assignment within the tree differs greatly. As mentioned above, Bender (1996) places Berta in the “Satellite Core”, grouped together with Maban, Fur, Central Sudanic and Kunama. Christopher Ehret’s (2001) reconstruction has come under significant scrutiny; however, his placement of Berta more closely reflects Greenberg’s inclination than does Bender’s in that it shares genetic ties to Eastern Sudanic (see Figure 3). The suggestions of this paper go beyond both, placing Berta specifically within the Eastern Sudanic sub-family of East Jebel (E4), and thus returning to Evans-Pritchard’s (1932) original idea of Berta being related to the languages of Aka, Kelo and Molo (AKM).

Figure 2: Bender's (1996) Nilo-Saharan structure

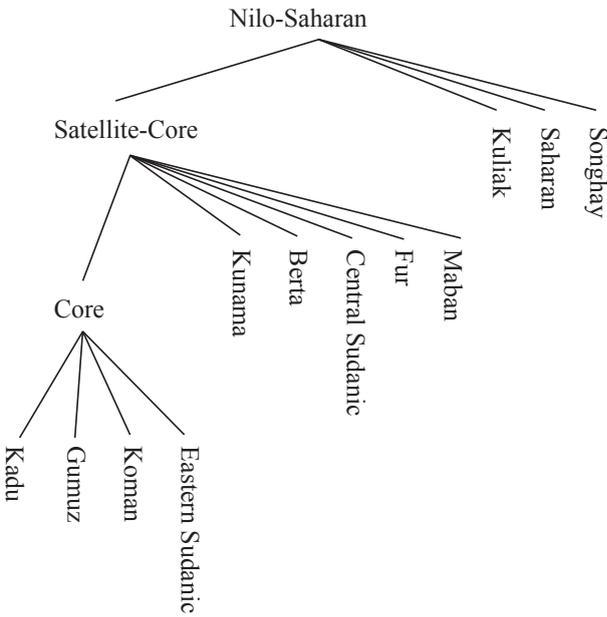
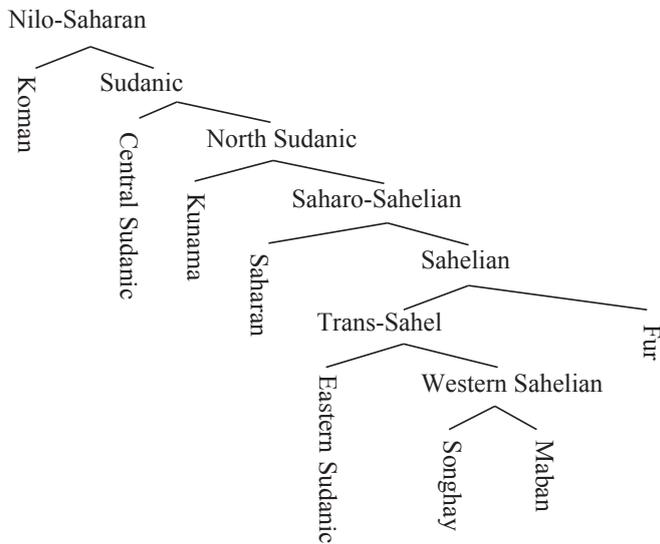


Figure 3: Ehret's (2001) Nilo-Saharan structure



1.2 The historical record

There are four sources from the 19th Century which document Berta varieties. The first record that we have available is from the French explorer, Frederic Cailliaud. His book, published in 1826, is based on a four-year trip in eastern Africa from 1819–1822 in which he traveled along the White Nile and recorded anthropological and linguistic data from that area. Most pertinent to the study at hand is the

more than 200 item word list from a Berta area. The location of its elicitation is uncertain, but based on the lexicostatistics it shows a high percentage of similarity to other Berta sources. A few decades after him, Lorentz Tutschek (1850) collected stories as well as a large word list with over 1000 lexemes in Berta. This data was collected in Europe from a Berta slave, and although it is a valuable resource, the caution about its having “great shortcomings” ought to be noted (Triulzi et al. 1976: 8). In the early 1870’s, Ernst Marno (1874), a German explorer, collected word lists from a number of people groups presently found near the Ethio-Sudan border. His word list shows some interesting lexical disparity between the languages in that region, and due to the age of his sample, the section pertaining to Berta remains an important resource for historical reconstructions. To close off the century, Halevy (1874) provides us with a small word list of just 70 items.

Berta research in the first half of the 20th Century continued at a similar pace to the previous century. Two Italians, Carlo Conti Rossini (1920) and Enrico Cerulli (1947), as well as the British anthropologist Evans-Pritchard, provide the only published data. All three collected word lists, but Cerulli breaks new ground by offering the first available grammar sketch of Berta. Additionally, he demarcates three dialects found amongst the Berta people. Evans-Pritchard, who will be examined more thoroughly later, is also significant since he is the first to suggest the genetic relatedness of other languages with Berta. Notably, his training was as an anthropologist, not a linguist.

The second half of the century and the beginning of the next have shown a significant increase in Berta research. During this period, Fleming documented new data from an isolated group of Berta he terms “Gebeto” living in the Didedsa Valley of Ethiopia, several hundred kilometers east of the majority population.² Sometime after him Bender compiled a lexicon of his personal elicitations. This document included a few other sources, and delineates some of the regional differences amongst the Berta, both in Ethiopia and Sudan. There are many entries where Bender (1989a) shows no dialectal contrast but where Bremer (forthcoming) does. The absence of apparent contrast between Bender’s “Mayu, Undu and Fadasi” is a likely factor for his conclusion that Berta shows “minor regional variation” and is “clearly a single language” (1997: 62).³

Three independent surveys conclude the inventory of available Berta literature. All three are exceptional in their own right. First, Siebert et al. (2002) put forward some intriguing data collected from the Asosa area. This area is predominately inhabited by speakers of the Maiyu dialect, however their word list contains extraordinary entries which are closer to varieties found in Sudan and Ethiopia. This record, if deemed reflective of the linguistic landscape in and around Asosa, suggests a certain amount of trans-regional borrowing and/or contact across dialectal regions. The second survey is notable, since it does not presume to be a Berta survey at all. Krell (2007) surveyed a group located in Yabus al-Ghaba, Sudan. In her report, this group she calls “Baruun be Magtole” were assessed to have only 56% lexical similarity with Berta.⁴ However, as I examined a wider comparative sample it revealed a clear affiliation with other Berta varieties and so Baruun be Magtole’s inclusion in this comparative study is warranted. The last survey, Bremer (forthcoming), offers a comparative look at six different Berta dialects. Several highlights are notable: included are two previously undescribed communities, Wabosh and Metehara. Additionally, significant lexical contrast was discovered between the dialects of Maiyu, Undulu and Fadashi, giving cause to rethink Bender’s (1989a) generous lexical similarity percentages. Finally, Bremer (forthcoming) suggests that low lexical similarity percentages as well as low intelligibility levels between some Berta speech communities may be an indication of multiple Berta languages, not merely dialects.⁵

² Fleming made two trips to the Didedsa area, the first “in the 1960’s” and the other in 1974 (Bender 1989b: 271). The information cited in this paper comes from Bender (1989b), and Bender’s own notation of G1 and G2 corresponds to each of Fleming’s respective trips.

³ Bender’s Mayu, Undu and Fadasi correspond to my Maiyu, Undulu and Fadashi respectively. For more information on names and their variations, see Appendix C.

⁴ Krell notes that this group is not listed as a Sudanese language in the Ethnologue, and so no ISO code is used to identify it.

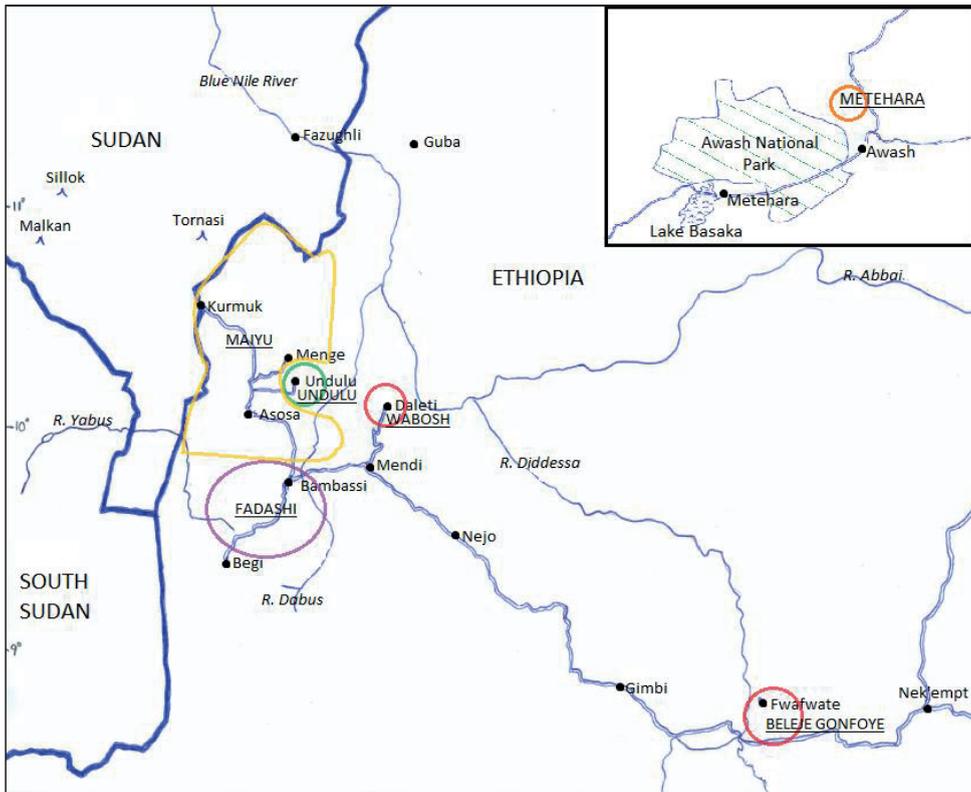
⁵ The categorization of languages and dialects is controversial and will not be focal in this paper. In this paper I use the term “dialect” as a precaution which is more or less synonymous with “speech variety”.

2 Evidence of language shift

2.1 Berta's contemporary phonemic inventory: examining six Ethiopian varieties

For an understanding of the basic phonological behavior as well as the phonemic structure of Berta, Andersen (1993: 55–67) and Bender (1997: 191–92) offer suitable introductions. Andersen's treatment is much more thorough and arguably more reliable than Bender's but due to the phonetic differences (of the phonemic inventory) found in the Sudanese variety he surveyed, Bender's Ethiopian-Berta analysis will be the primary point of comparison in this paper. My conclusions about the phonemic inventory differ somewhat from Bender's; each divergence will be treated independently as it arises. Since the current research touches on half a dozen different dialects, noting the similarities and differences across Berta speech communities offers a good starting point for later comparison.

Figure 4: Approximate location for 6 Berta dialects in Ethiopia



The six dialects of Maiyu, Metehara, Undulu, Fadashi, Wabosh and Beleje Gonfoye all have contemporary phonemes rooted in the following 17 proto-consonants: *b, *b̥, *m, *f, *d, *d̥, *n, *r, *l, *z, *ts, *s, *k, *g, *j, *h and *ʔ.⁶ The Maiyu dialect has many observable nuances. For example, the voiceless egressive stops [p], [t] and [k] are in free variation with their voiced counterparts [b], [d], and [g]. Additionally, the voiced bilabial stop /b/ may often be pronounced as a bilabial fricative, [β], but never the labio-dental fricative [v]. The velar stop /g/ is often realized as [ʒ] or [dʒ] intervocally or preceding a front vowel. The implosive /b/ is most often realized as an ejective [pʰ]

⁶ Different dialects' phonemic inventories have unique features and often have variant phonetic realizations of the proto-consonant. Some contemporary varieties have more phonemes than the posited proto-inventory. Namely, *t has merged with *l in Wabosh and BG and is therefore not amongst this list.

in Maiyu, however frequency and preference depends on the speaker. The voiceless sibilant [s] is in free variation with the voiced counterpart [z], although personal observation suggests that the voiced variety seems to be preferred in Maiyu-speaking regions outside of Asosa. The ejective [kʰ] is in complementary distribution with [tʰ]; it, like other phones in Berta, is palatalized in environments where it precedes a front vowel. Similarly, the velar [ŋ] and the palatal [ɲ] are allophones; the velar nasal is realized as a palatal in environments preceding a front vowel. In addition to the list of 17 proto-consonants above, Maiyu has a number of additional phonemes which are not part of the phonemic inventories of all other Berta dialects. These are the interdental fricative /θ/, the glide /j/ and the bilabial glide /w/, bringing the total number of Maiyu consonant phonemes to 20.

Lexically and phonetically, Metehara most closely reflects the Maiyu variety.⁷ Phonetic differences can be observed in the voiced sibilant /z/ which is most often realized as voiced, not voiceless. Furthermore, in cognate forms the interdental fricative /θ/ found in Maiyu is consistently realized as [t] in Metehara to the total absence of [θ]. In this regard, Metehara is the only variety surveyed which shows no phonetic deviation from *t. Metehara has maintained phonemic contrast by disallowing free variation between [t ~ d] as in Maiyu. In sum, the Metehara [t] corresponds to /θ/ in Maiyu, and [d] corresponds to [t ~ d]. Interestingly, both the voiced and voiceless bilabial and velar stops are in free variation in Metehara. Metehara does not seem to have a phonetic ejective [pʰ] like Maiyu, but rather the corresponding phoneme /b/ only surfaces as implosive [ɓ]. Like Maiyu, Metehara has a phonemic /w/. The velar stop /g/, the ejective /kʰ/, as well as the velar /ŋ/ all undergo the previously described allophonic palatalization under the same phonological conditioning as Maiyu.

Undulu speakers prefer the voiced sibilant [z] over the voiceless [s] in most cases, although records suggest that there is free variation between [s ~ z]. Bremer (forthcoming) offers no record of phonetic alternation between ejective [pʰ] and implosive [ɓ] in Undulu, but as in Metehara, the phoneme /b/ is consistently realized by the reflex [ɓ]. Importantly, this is not a neutralization of contrast, since no known Berta dialects contrast [pʰ] and [ɓ], rather what is notable here is that not all varieties actually have the phone [pʰ] in their inventory, and those that do only have it as an allophone of [ɓ]. As in Maiyu, /θ/ is a phoneme and is contrastive to /t ~ d/, however the phonemic nature of [w] needs to be researched further to determine whether or not it is a separate phoneme or in free variation with [f] as is the case with other Berta dialects. The allophonic behavior of /g/, /kʰ/, and /ŋ/ are uniform in both Maiyu and Metehara.

The consonant system of Fadāshi has little variance from the three previously described dialects; however, there is one glaring exception: the interdental fricative [θ] is in free variation with the lateral [l]. This alternation is intriguing, since the voicing and manner of articulation are different. Even more surprising, /l/ is a phoneme in all Berta dialects, and so free variation between [θ ~ l] fails to preserve phonemic contrast that exists in other Berta varieties. Due to the regularity of this sound correspondence across dialects, this is almost certainly an example of merging in Fadāshi rather than splitting elsewhere. The limited data make it difficult to predict the directionality of the [θ ~ l] sound change, but speaker intuition suggests movement towards the interdental fricative. Like the varieties examined thus far, it is clear that Fadāshi has free variation between voiced and voiceless sibilants and stops, however Bremer (forthcoming) does not record any examples of a /p ~ b/ alternation; only [b] is attested. More research is needed to determine whether or not free variation exists in this bilabial stop. As with the other varieties treated thus far, the allophonic consonants /g/, /kʰ/ and /ŋ/ behave equally, and /w/ is phonemic.

⁷ There are an estimated 1000 Berta living north of the town of Awash (close to the town of Metehara) about 300 kilometers east of Addis Ababa, and nearly 1000 kilometers from Asosa. Oral accounts suggest that these Berta were resettled there around the turn of the 20th Century, where they lived in isolation from “western” Berta for about 100 years. In 2005, a governmental initiative to return the Metehara Berta to their indigenous land around Asosa was launched, however due to differences in climate and culture, many chose to return to Metehara. Most Maiyu-speaking Berta living in Asosa are aware of the “Metehara” Berta, and most of them have a favorable response to their speech variety.

Table 1: Phonemic chart of six Berta dialects⁸

Proto Consonants	Phonetic Realizations					
	Maiyu	Metehara	Undulu	Fadashi	Wabosh	Beleje Gonfoye
*b	p ~ b ~ β	p ~ b ~ β	p ~ b ~ β	b ~ β	b ~ β	b ~ β
*ḅ	p' ~ ḅ	ḅ	ḅ	ḅ	p' ~ ḅ	p' ~ ḅ
*m	m	m	m	m	m	m
*f	f	f	f	f	f ~ w	f ~ w
*	w	w	w	w	-	-
*t	θ	t	θ	θ ~ l	l	l
*l	l	l	l	l		
*d	t ~ d	d	t ~ d	t ~ d	t ~ d	t ~ d
*d'	d'	d'	d'	d'	d'	d'
*n	n	n	n	n	n	n
*r	r	r	r	r	r	r
*z	s ~ z	s ~ z	s ~ z	s ~ z	dʒ	dʒ
*ts	s'	s'	s'	s'	tʃ'	tʃ'
*s	ʃ	ʃ	ʃ	ʃ	s ~ ʃ	s ~ ʃ
*	j	j	j	-	-	-
*k	[k', [tʃ']]	[k', [tʃ']]	[k', [tʃ']]	[k', [tʃ']]	k'	k'
*g	[k ~ g], [dʒ ~ ʒ]	[k ~ g], [dʒ ~ ʒ]	[k ~ g], [dʒ ~ ʒ]	[k ~ g], [dʒ ~ ʒ]	k ~ g	k ~ g
*ŋ	[ŋ], [ŋ]	[ŋ], [ŋ]	[ŋ], [ŋ]	[ŋ], [ŋ]	[ŋ], [ŋ]	[ŋ], [ŋ]
*h	h	h	h	h	h	h
*ʔ	ʔ	ʔ	ʔ	ʔ	ʔ	ʔ

Due to significant similarity in their phonemic inventories and their phonetic realizations, the Wabosh and Beleje Gonfoye varieties will be treated together.⁹ Grammatically and phonologically, these two dialects are the most unique of all the known Berta varieties of Ethiopia.¹⁰ Like Fadashi, there is no evidence that the voiced and voiceless bilabial stops, [b] and [p] are in free variation. In all elicitations thus far, only [b] surfaces. Comparatively, although there is a record of free variation between both [t] and [d] as well as [k] and [g], the voiced varieties are much more common, surfacing in the majority of samples. There is notably no /w/ phoneme found in these geographically peripheral dialects, but uniquely [f] and [w] are in free variation, again with the voiced variant being most common. Additionally, there is no [θ] in either Wabosh or Beleje Gonfoye. In cognates, the /θ/ of Maiyu and Undulu corresponds to [l]. Additionally, /l/ in Maiyu corresponds with /l/ in Wabosh and Beleje Gonfoye, which means that the eastern varieties do not preserve phonemic contrast found in

⁸ Allophones are indicated with brackets [], and free variation with a tilde ~. In the case of *g, there is an allophonic pair in which there is also free variation amongst the allophones. In most dialects surveyed, allophones are typically conditioned by the successive vowel, i.e. they are the result of palatalization rules.

⁹ The Wabosh live in and around the small town of Daleti, located 42 kilometers north of the town of Mendi, situated on the eastern periphery of the Berta language area. There are an estimated 3000 – 5000 Wabosh speakers in that area, and a recent resettlement has brought them together from outer-lying areas. In recent years the Wabosh have had increasing contact with Maiyu speakers; the town of Daleti belongs to a predominately Maiyu-speaking *woreda* 'county', Oda, in the Benishangul-Gumuz region. Most Wabosh have heard of the Beleje Gonfoye Berta living in the Didessa Valley, and many of them regard them as "Wabosh". The historical record is uncertain, but the linguistic record confirms a historical connection between these groups, suggesting that the Beleje Gonfoye originated from the Daleti area. Slave raiding was an unfortunate reality in that area, and this is believed to be the cause behind how the Beleje Gonfoye became displaced from the larger Berta community. The Beleje Gonfoye Berta have had little to no contact with other Berta, living hundreds of kilometers removed from the Berta speech area, about 30 kilometers northwest of the town of Arjo, nearby the Didessa River. The Beleje Gonfoye surveyed in this study were from the small town of Fwafwate.

¹⁰ This statement is based on original research and personal observation and is currently not attested in any literature.

other Berta dialects. The phoneme /z/ in Maiyu, Undulu and Fadžashi is realized as an affricate, [dʒ], and it is important to realize that this phoneme contrasts with another phoneme /ʃ/ which is realized as [s ~ ʃ] in the eastern varieties. Furthermore, the allophonic relationship between [dʒ] and [g] found in the other varieties does not exist in Wabosh and Beleje Gonfoye. The ejective [sʰ] found in the majority of Berta dialects is realized as a [tʃʰ]. Similar to the behavior of the allophones [g] and [dʒ], [tʃʰ] does not share the same allophonic relationship with [kʰ] that other varieties have. Unlike other varieties, the phonemic ejective /kʰ/ does not have any allophones regardless of the successive vowel, and is always realized as [kʰ]. At this point in my research, the palatal [ɲ] and the velar [ŋ] are the only observable allophones in Wabosh and Beleje Gonfoye, and these pattern identically to the other Berta varieties. In sum, there are just 17 consonant phonemes in Wabosh and Beleje Gonfoye, making it the most condensed system amongst the Ethiopian Berta dialects. For an overview of the phonemic inventories, Table 1 below offers a comparison of the six dialects mentioned above and also includes a hypothesis of proto-consonants which are treated more fully in section 2.2.

Comparison with Andersen and Bender

Based on research conducted on a Sudanese variety in Abeegu, Andersen (1993: 56) lists 13 phonemic consonants which are “unproblematic”. The majority of these phonemes are stable in Ethiopian Berta varieties as well. Metehara’s phonological inventory most closely reflects Andersen’s “unproblematic” list. Indeed, of those consonants listed, only the /b/ has multiple surface forms in Metehara, being the result of free variation. Of the allophonic pairs listed, only [ɲ ɳ] and [j ɡ] have been observed in Ethiopia. The uvular stop [q] and the palatal ejective [cʰ], as well as the fricatives [h] and [χ] have not been recorded east of the Sudanese border. Andersen (1993: 59) notes that his own data seems to indicate a dialectal correspondence between the Sudanese [q cʰ] allophones and the Ethiopian /kʰ/, and the Ethiopian dialects surveyed thus far attest to the credibility of his assessment. Likewise, Andersen’s description of the glides [j] and [w] shows identical patterning amongst all the varieties surveyed. The voiceless pharyngealized dental stop [tʰ] and the [v] described by Andersen remain unattested in Ethiopian research.

Bender (1997: 191) lists 26 consonantal phonemes, however this number is presumably inflated. Bender does not note free variation between [s ~ z] or the allophones [ɲ ɳ] and [g dʒ], but rather lists each of these surface forms as separate phonemes. Additionally, differences in analysis have led to varying interpretations of prenasalized stops as phonemes, namely Bender’s /mb/, /nd/ and /ŋg/.¹¹ He lists these as phonemes, thus parsing them as units, however Andersen’s (1993: 55) consonantal inventory as well as the phonological analysis employed in the standardized Berta orthography suggests that they can be understood sequentially.¹² As we will see later, nasal-stop sequences are common throughout all Ethiopian Berta varieties but are rare in the East Jebel languages. This regional and genetic distribution of nasal-stop sequences provides some important clues about the historical development of Berta and the East Jebel subfamily.

2.2 Historical and contemporary lexical variation

Appendix A offers an extensive comparative list of both regular and irregular sound correspondences observed in contemporary and historical Berta word lists. Much of the observable phenomena found are further described in the sections which follow. Recognizing the range of sound behavior within the Berta language offers a baseline for extra-Berta comparisons. Only when the entire corpus is considered can appropriate generalizations about phonetic patterns and language shift be made. The next sections are structured according to current phonological inventories in most Berta speech varieties. Within each of these sections (Stops, Fricatives, Nasals, Glottals, Liquids and Glides) historical phonal distribution will be examined, and from this data I offer a hypothesis of Berta’s

¹¹ Andersen (1993: 44) says “It is not clear, however, whether the initial nasal constitutes a syllable N by itself, or whether we are dealing with a NCV syllable, which contains an initial consonant cluster. On the one hand the nasal seems to carry its own tone, as does any vowel, but on the other hand the articulatory energy of the nasal seems to be rather weak.”

¹² Cross-reference to A. and S. Neudorf’s *Bertha-English-Amharic Dictionary* (2007).

proto-consonantal inventory. Each section offers a table which juxtaposes sound correspondence, with a complete proto-Berta phonological inventory offered at the end in Table 8.

2.2.1 Stops

Amongst cognates, stops have a wide range of alternation across Berta speech varieties which is often unpredictable. Assuming that free variation between voiced and voiceless phones of the same place of articulation (POA) is an old feature of Berta (at least 200 years), a relatively high frequency of word-initial and word-medial alternation exists between *b, *d and *g.¹³ Looking beyond the distribution of these three egressive stops, we see that both ejectives and implosives have infrequent alternation, but importantly such examples are irregular. Both ejectives and implosives will be treated separately in section 2.2.4.

The glottal stop is a curious phoneme in Berta, and its evolution is largely indeterminable. Complicating analysis, the oldest records seldom record word-initial or word-final glottals, if at all. Cailliaud (1826), whose transcriptions rely heavily upon the French orthography, is the only source who uses [h] in places where a phonetic glottal [ʔ] would be expected. Examples include *ohon* ‘meat’, *horgodit* ‘clothes’ and *hier* ‘milk’. In places where a glottal stop was likely present, Tutschek (1850) often records successive vowels, omitting any kind of further indication of a present glottal stop. Because of diphthongs and other phonetic variation, these too can be difficult to interpret, however the following cases are presumed to have had intervocalic glottal stops: *oong* ‘meat’ and *oo* ‘grandmother’. In contemporary Berta varieties, the glottal stop is rarely seen in nouns, however it is common as a verbal morpheme in cases like 1.SG.SUB ʔáʔ- as well as the non-past tense/aspect marker -ʔi.

Table 2: Berta stops and their historical distribution

Proto consonants	Cailliaud	Tutschek, Marno, Halevy	EP	Maiyu, Undulu, Fadashi, BG, Wabosh	CerK, CerR	CerDS	S & W, BBM
*b	p ~ b	b	b	p ~ b ~ β	b	p ~ b	b
*d	t ~ d	t ~ d	d (t) ¹⁴	t ~ d	d	t ~ d	t ~ d
*g	k ~ g	k ~ g	k ~ g	k ~ g	k ~ g	k ~ g	k ~ g
*ʔ	h	-	ʔ	ʔ	-	-	ʔ

2.2.2 Fricatives

Historical behavior amongst Berta fricatives displays extreme dissimilarity. As observed above, most contemporary dialects have contrastive phonemes /f/ and /w/, but the eastern-most varieties do not preserve this contrast, showing free variation between /f ~ w/ (with a preference towards a surface [w]). Examining the oldest Berta records offers some possible clues regarding the catalyst behind this development. Cailliaud does not record any instances of [w], however some of his multi-vowel sequences suggest that [w] is a phonologically conditioned semi-vowel.¹⁵ In his data, the consonant /f/, on the other hand, appears word-initially, -medially and -finally, before any vowel. Marno’s (1874)

¹³ The standardized Berta orthography (based on Maiyu) represents [p ~ b], [t ~ d] and [k ~ g] by the voiced variants, /b/, /d/ and /g/.

¹⁴ It seems as though [t] surfaces only in the word-final position in Evans-Pritchard’s data. Data in this regard is limited, but a phonological rule of this type has not been observed elsewhere amongst Berta dialects.

¹⁵ Cailliaud, noticeably more than others, transcribes Berta in a manner which closely reflects his language-of-origin, French. Examples of consecutive vowels include but are not limited to: *areguio* “afternoon”, *gouacheyé* “sing”, *ouatichy* “sneeze” and *ouara* “tendon”. Notably, in the case of [w], neither the French nor the German orthographies treat this letter as a labio-velar semi-vowel, so it is also plausible that their respective phonemic biases could have altered their recordings of this sound.

data matches Cailliaud's, having no instances of [w] but rather [f] occurring in all environments. Tutschek differs markedly from these two. He is the first to record examples of /w/, found in just three words: word-initially in *wao* 'abandoned, widow' and *wala* 'not', as well as word-medially in *zawuta* 'shoot'. The origin of *wao* is uncertain, but the other two lexemes are Arabic loan words, suggesting that /w/ is a borrowed phone, probably adopted into most Berta dialects' phonemic inventories about 200 years ago.

An interesting twist to the variety that Tutschek surveyed is that the fricative /f/ never appears before a back vowel. This phonological restriction cannot be found elsewhere in other Berta varieties, historical or contemporary. Comparative analysis reveals an unexpected dialectal sound correspondence between [f] and Tutschek's [h]. In many contemporary cognate examples where /f/ precedes a back vowel, this /f/ is transcribed as [h] by Tutschek. This correspondence may suggest an allophonic relationship between [f] and [h] in Tutschek's data, which would be unique. Further observation shows that Tutschek's [f] and [h] both regularly appear before the low vowel [a]. Second, [h] seems to be in complementary distribution with another allophone in Tutschek's speech variety. He has multiple entries in which he transcribes a rare sound with the digraph "gh". These words are *ghinga* 'eat', *ghiba* 'fear', *ghio* 'inside', *ghima* 'squeeze' and *gighe* 'yesterday'. One can only speculate about the precise phonetic quality of this sound, but [ɣ] or [χ] are plausible deductions. In the five examples found in Tutschek, this phone occurs only before a front vowel, and also fits nicely as an allophone of [h] which often surfaces before a back or low vowel.¹⁶

Why then is there a regular sound correspondence between [f] and [h], which can be observed primarily in Tutschek but also in other varieties?¹⁷ There is strong likelihood that these cases involving [h] are an example of overlapping sound correspondence sets. Furthermore, apart from Tutschek's data there is little justification for not including both *f and *h in a proto-Berta inventory. Tutschek's examples above suggest a possible split between these two phones which may eventually lead one to discount the incorporation of both *f and *h in the proto-inventory, but until more evidence is established, the inclusion of both seems warranted.

Further complicating matters, by the time of Tutschek, another proto phoneme, *t had also undergone a shift. In the variety that Tutschek surveyed, *t, like /f/ in certain environments, had nearly dissolved into /h/. Elsewhere in other Berta dialects, this *t maintained its phonemic nature but went on to have a very exotic development.¹⁸ The proto-stop *t surfaces as an interdental fricative [θ] in Maiyu, and is the most variant phone in Berta's sound system. In contemporary dialects, this surfaces predictably as [t], [ð], [θ] and [l], but older records regularly transcribe it as [t], [h] or [ɣ].¹⁹ Cailliaud, Tutschek and Marno all lack the phone [θ] that is present in contemporary Maiyu and Undulu varieties. Presumably using the French orthography as a base, Cailliaud transcribes this sound as [r]. In the vast majority of Tutschek's cognates, this interdental fricative surfaces as [h], however it also corresponds at times with "gh". Maiyu words *θija* 'eat' and *θa...iyu* 'inside' correspond nicely to Tutschek's *ghinga* and *ghio*. Complicating matters slightly, *ghiba* and *gighe* have contemporary correspondences of *hiba* and *gihe*, corresponding with [h] rather than the [θ] seen elsewhere.²⁰ Marno, on the other hand, regularly records *t as [t]. Just like in Metehara, this adjustment has not resulted in a lack of contrast, since [t] and [d] are not in free variation. Therefore, the [t] of Marno and Metehara corresponds to *t, and [d] corresponds to *d, showing no digression from the proto-consonants. Some of the varieties surveyed have merged *t with another phoneme, no longer maintaining proto-phonemic contrast. Cases where phonemic contrast are no longer maintained are represented by shading in Table 3.

¹⁶ One exception to this is *hiri* "under".

¹⁷ Contemporary elicitations from Fadashi also indicate alternation between [f] and [h] (as seen in words *fadi/hodi* "white" as well as *fu:da/ho:da* "gold").

¹⁸ Andersen (1993: 60) notes that a voiceless uvular fricative [χ] can still be found in certain Sudanese speech varieties and is an allophone of h. This [χ] corresponds to [k'] in most Ethiopian dialects and is therefore representative of a different proto-consonant altogether and must not be confused with *t. Andersen's (1993, 1995) data, like Metehara, has a correspondent [t] to the Maiyu/Undulu [θ].

¹⁹ None of the older documents use the icon [ɣ], however this is my interpretation of Tutschek's [gh] and Cailliaud's [r].

²⁰ The historical record offers no cognates with *ghima*

In contrast to the fricatives observed so far, the sibilants *z and *s show regular correspondence in Berta. Both phonemes undergo free variation in certain varieties although not in all dialects. The voiced *z has a relatively consistent geographic distribution, wherein those varieties north and west of Asosa tend to use the voiced [z], but those regions in Asosa and east of Asosa tend towards a surface [s]. The *s is realized as [ʃ] in all varieties except Wabosh and Beleje Gonfoye where [s ~ ʃ] is normative. This proto-consonant, like other sibilants and stops in Berta, has undergone palatalization. Free variation between [s ~ ʃ] in Wabosh and Beleje Gonfoye is made possible by a phonetic shift in *z to [dʒ]. A more detailed description about this historical process is provided in Section IV.

Table 3: Berta fricatives and their historical distribution²¹

Proto consonants	Cailliaud	Tutschek	Marno	Halevy	EP	Maiyu	Undulu, CerK, CerR,	CerDs	Fadashi	Wabosh and BG, Rossini	BBM	S & W
*f	f	f	f	f	f	f	f	f	f	f ~ w	p ~ f	f
*z	s ~ z	z	s	s	z	s ~ z	z	z	s ~ z	dʒ	z	s ~ z
*s	“ch”	sh	“sch”	s	sh	ʃ	ʃ	ʃ	ʃ	s ~ ʃ	ʃ	ʃ
*t	“r”	gh, h	t, d	t, d	t	θ	θ	θ, ð	θ ~ l	l	t	θ, t, ð, h, z, t, d
*h	h	h	h	h	h	h	h	h	h	h	h	h
*	-	w	-	w	-	w	w	-	w	f ~ w	w	β, w
*l	l	l	l	l	l	l	l	l	l	l	l	l
*d ²²	d	d	d	t (?)	d	t ~ d	t ~ d	d	t ~ d	t ~ d	t ~ d	d, l

2.2.3 Nasals

Nasals are the most stable phones in the Berta sound system. Historically there is very little observable alternation, and all varieties show phonemic evidence of the three nasals /m/, /n/, and /ŋ/. The precise nature of the palatal [ɲ] is debatable, and it may be different amongst different speech varieties. As described in Section 2.1 (and illustrated in Table 1), the palatal [ɲ] is an allophone of the velar /ŋ/ in most varieties, however some contemporary word lists hint at a possible shift occurring. Siebert et al. (2002) record *bɛŋeni* ‘red’, *s’ur(i)ɲi* ‘long, far’ and *di:ɲi* ‘louse’. These examples unexpectedly show the velar nasal [ŋ] occurring before a front vowel where a palatal [ɲ] would be expected. Elsewhere, ‘flower’ is transcribed *du:du bɛŋeni* (lit. ‘red leaf’) using the palatal [ɲ]. This together with their transcription *nɪŋe* and *nɪje* ‘woman’, seems to discount any type of phonological predictability and may indicate free variation between [ɲ ~ ŋ].

Krell (2007) also presents evidence which disrupts the theory of a trans-Berta [ɲ ŋ] allophonic relationship. Examples of an unexpected velar nasal include *duɲi* ‘big’, *tusɲi* ‘burn’, *soriɲi* ‘long, far’ and *busiɲe* ‘teach’, however this distribution does not eliminate the presence of a phonetic palatal [ɲ] which appears elsewhere in words like *ɲara* ‘grass’, *ɲinsi* ‘heavy’ and *di:ɲi* ‘louse’. Assuming the reliability of Krell’s transcriptions, at least two possible explanations exist. First, in Baruun be Magtole, the allophonic relationship between [ɲ ŋ] may have dissolved (or more likely, *be dissolving*), leading to two separate phonemes, /ɲ/ and /ŋ/. If this is what is happening, it is still in the early stages, since her data shows no examples of a palatal [ɲ] appearing before a back vowel. There is also a conceivable second option: while [ɲ] and [ŋ] evidence a previous allophonic relationship, they may be

²¹ Shading indicates lack of phonemic contrast.

²² Cailliaud sometimes records a [t], but this doesn’t surface in words that have a cognate [d]. Marno’s [t] most often surfaces in cognates where there is a Maiyu [θ], but “elephant tusk” is an exception to this pattern. Evans-Pritchard frequently records [d] for /d/ except when it is in the word-final position, it tends towards [t]. This may be coincidental or it may be due to a phonological pattern of word-final devoicing.

shifting towards free variation (similar to what can be observed in Siebert et al. (2002)). In this regard, Krell's own data offers a possible clue. Her word *duŋi* 'many, much' corresponds semantically with *duŋi* 'big' found in the Ethiopian varieties. In fact, this is the same (historical) lexeme, but is transcribed differently in each elicitation. This example could be used to support either option: In the first regard semantic disambiguation would be the catalyst towards the creation of a new phoneme /ɲ/, in the second it still is the same lexeme with two different surface forms, the result of free variation. Cailliaud and Tutschek as well as my data from Wabosh all offer historical evidence of the semantic correspondence between 'big' and 'many'.

Nasal-stop sequences are a common characteristic that all Berta varieties share. In such cases, the nasal matches the POA of the subsequent stop. As we will see later, the East Jebel languages do not have nasal-stop sequences as regularly as Berta, but in cognate forms the E4 languages typically preserve either the nasal or the stop. This development suggests that the nasal-stop combination is old and that Berta has preserved it despite its erosion elsewhere in the East Jebel family.

Table 4: Berta nasals and their historical distribution

Proto consonants	Cailliaud	Tutschek	Marno	Evans Pr	CerR	CerDs, CerK	Maiyu, Undulu, Fadashi, BG, Wabosh, Ros.	Baruun BM	S & W
*m	m	m	m	m	m	m	m	m	m
*n	n	n	n	n	n	n	n	n	n
*ɲ ²³	[gn ng]	[ñ ng]	[nj ng]	[ni ng]	ny, ɲ	ɲ	[ɲ ɲ]	$\begin{matrix} \text{ɲ} \\ \text{ɲ} \end{matrix}$	ɲ ~ ɲ

2.2.4 Glottals: ejectives and implosives

Some of the older records attest the presence of both implosive and ejective stops, however the phonetic representations used in these transcriptions vary. Those instances which lack differentiation are likely due to the fact that early researchers did not have linguistic training, rather than assuming that ejective and implosive stops are recent innovations in Berta. Cailliaud appears to use the digraph [ts] to record [s'], as in *mitsa* 'chicken', and [qu] for [k'] as in *quedit* 'laugh' and *asisique* 'sand'. His transcriptions do not differentiate between ejective and implosive stops. Conversely, Tutschek uses the apostrophe ['] to denote implosive consonants [b] and [d], but does not record any ejectives. Notably, Tutschek's [dʒ] regularly corresponds to the [k'] found in contemporary dialects. Neither Marno nor Halevy offer any evidence of ejective or implosive stops, however, their data samples are much smaller than either Cailliaud's or Tutschek's.

All contemporary word lists sampled in Ethiopia attest both implosives and ejectives. Krell's (2007) Sudanese survey on Baruun be Magtole does not record any ejective stops, but does have a [d] in *doŋ* 'axe' and *duŋi* 'big'. Bender's (1989) Al Azharia, also sampled in Sudan, has an ejective [k'] as in *k'aβa* 'blood' and *mank'ura* 'pot' but no other implosives or ejectives are evidenced in these two varieties. Bender (1997: 192) is probably correct when he writes, "given that glottals are found in Berta but generally not in the related Eastern Sudanic family, one may wonder whether Berta gained them from neighboring Koman or Afrasian languages (notably Oromo and Amharic)." In Ethiopia, both Amharic and Oromo are languages of wider communication (LWC) in most Berta-speaking areas. The LWC's impact on Berta glottals is plausible, however it is not possible to determine at what point in history this shift would have occurred. Based on anthropological estimates on Berta migration, a proto-Berta inventory without glottals would necessitate a reconstruction which

²³ Marno also seems to use "gn" for [ɲ], however its use is unpredictable and may not be significant. The only example of "ni" in Evans-Pritchard is for the word *neri* 'grass', but this may be a way of transcribing the palatal [ɲ]. Although not attested elsewhere, it is possible that the allophonic palatal [ɲ] in contemporary Berta dialects may have arisen from a phonological assimilation of the vowel into the nasal in order to avoid vowel hiatus. For reasons cited in section 2.2.3 above, both Krell's (2007) data on Baruun be Magtole and Siebert et al. (2002) present data which disrupts the widespread distribution of an allophonic pair [ɲ ɲ].

predates 1600 AD. Triulzi (1981) suggests that the Berta people immigrated to the Ethiopian plateau approximately 400 years ago, which would be a likely estimate for the time when Berta acquired glottalics (assuming they didn't have them before then).

Table 5: Berta glottalized consonants and their historical distribution

Proto consonants	Cailliaud	Tutschek	Mamo, Halevy	EP	Maiyu, Jebel Ura	CerDs, CerR	CerK	Rossini	Undulu, Fadashi, Metehara	Wabosh and BG	BBM	S & W
*ḃ	-	bʼ	-	-	pʼ ~ ḃ	-	-	-	ḃ	pʼ ~ ḃ	-	pʼ
*ɗ	d	dʼ	d	-	ɗ	d, dʰ ²⁴	-	-	ɗ	ɗ	ɗ	ɗ
*ts	s, ts	-	-	s	sʼ	tsʼ	cʼ	sʼ	sʼ	tʃʼ	s	sʼ
*k	qu	dʒ	-	q	[kʼ tʃʼ]	cʼ, kʼ	kʼ	-	[kʼ tʃʼ]	kʼ	-	kʼ, tʃʼ

2.2.5 Liquids

Liquids in Berta seldom show variation; both the historical and contemporary records provide persuasive evidence that both *l and *r were part of proto-Berta's inventory. Words containing /l/ like *mili* 'black' and *juli* 'house' exhibit little to no variation, similar to words with /r/ like *are* 'eye' and *ʃinir* 'donkey'. Rare alternation between /l/ and /r/ can be found in a few examples like *agulgulu* and *adurdur* 'whirlwind' (Metehara and Undulu respectively) and also Tutschek's double entry *bele* and *bar* 'stone'. All varieties have contrastive /l/ and /r/ and these often correspond regularly in cognates. Notably, Cailliaud uses the same symbol "r" to record both *t and *r.

Table 6: Berta liquids and their historical distribution

Proto consonants	Cailliaud	Tutschek	Mamo	Halevy	EP	Maiyu, Undulu, CerK, CerR	CerDs	Fadashi	Wabosh and BG, Rossini	BBM	S & W
*l	l	l	l	l	l	l	l	l	l	l	l
*r	r	r	r	r	r	r	r	r	r	r	r

2.2.6 Glides

Glides in Berta are difficult to reconstruct. Indeed, even in contemporary dialects their behavior can be difficult to analyze. Andersen (1993: 60) puts it succinctly, saying "the phonemic status of the glides [j] and [w] is somewhat problematic." My analysis suggests that glides in Berta which are phonologically conditioned in vowel sequences (see example 57 in Andersen 1993) should not be interpreted phonemically, but those examples which lack phonological conditioning are phonemic. Phenomena such as compounding and borrowing have begun to alter predictable phonological behavior and are likely leading to splits in the phonemic inventories of some Berta varieties.²⁵ In current Berta speech varieties, both [j] and [w] appear word-initially in unmotivated cases, but their presence in proto-Berta is doubtful.

²⁴ Cerulli is the only author to record a retroflex [d] but interestingly also uses an implosive [d] for his phonetic transcriptions as in *hos'adaa* "today", *ingodee* "tree", *budu* "bull".

²⁵ A good example of the effect nominal and verbal compounding have on glides can be seen in Andersen (1993: 60) example 58. The first item, *tàayú*, is a compound of *ta* + *iyu*, meaning "at stomach", or "at inside". This compound has become lexicalized and is now used only within the semantic range of "at home". The phonologically predictable glide of *iyu* has been preserved, despite the elision of the preceding vowel. Shifts such as this do indeed make the phonemic status of glides "problematic".

Please refer to Section 2.2.2 above for treatment on the glide [w]. The glide [j], like [w], is difficult to reconstruct since it rarely appears in phonologically unmotivated cases. Word-initially, there are only a handful of cases where a [j] can be found. The glide in Wabosh *jeho* ‘salt’ corresponds to [h] in Bender’s (1989) *he’he(y)o* and Fleming’s *hehe, heehə*. Furthermore, the two examples, *yagut* ‘fish’ and *yamut* ‘mosquito’ listed in Andersen (1993: 60) are no less problematic, since they also find both historical and contemporary correspondence in *igu*²⁶ and *umuθ* respectively. There are therefore very few examples of either glides [j] or [w] appearing in phonologically unmotivated environments. The glides appear to be the result of phonemic splits, wherein most contemporary speech varieties are adopting both glides into their phonemic inventories.

Table 7: Berta glides and their historical distribution

Proto consonants	Phonemes	Cailliaud, Marno	Tutschek, Halevy	Evans Pr	Maiyu, Undulu, Metehara	CerK, CerR, CerDs	Fadashi	Wabosh, BG	BBM	S & W
*	/w/	-	w	-	w	-	w	f ~ w	w	β, w
*	/j/	-	-	-	j	-	? ²⁷	j (?)	-	j

2.2.7 Summary

Berta’s consonantal inventory looks very different depending upon the variety surveyed, but regular sound correspondences provide a window into sound change. Berta speech communities, contemporary as well as historical, all have between seventeen and twenty phonemes. I suggest that the glides that are clearly not the result of phonological conditioning are borrowed phones and were not part of proto-Berta’s inventory. Both are phonetically present in Berta’s sound inventory as phonologically conditioned semi-vowels, however due to their rarity and also the lack of predictable sound correspondence, neither are presumed to have belonged to proto-Berta. Table 8 below lists the 18 phonemes of proto-Berta:

Table 8: Proto-Berta phonological inventory

	Labial	Alveolar	Velar	Laryngeal
Stops:		t, ts	k	
	b	d	g	?
	β	dʰ		
Fricatives	f	s		h
		z		
Nasals	m	n	ŋ	
Liquids		l, r		

2.3 Lexical and semantic shift: Exploring polysemy and diachronic semantics

In this section and the next, certain Berta semantic categories will be investigated to show isoglossic trends and threads of relatedness within Berta and beyond. In the first case, Berta numerical systems will be treated, and observations on an east-west isoglossic distribution will follow. The section concludes by broadening the scope of the survey to include the East Jebel languages, noting evidence of historical semantic shift.

Not many of the older Berta sources have left behind traceable information, allowing for the precise identification of survey locations. Thus, one can only deduce approximate regions for the Cailliaud, Marno and Rossini surveys. Furthermore, due to political instability and multiple wars that have affected this region in the last half century, resettlement and geographic movements are expected. The table in Appendix C provides a list of known information about each source’s origins, and this

²⁶ There are many variants to this word, including *nəgu* and *d’āgul*.

²⁷ My records do not indicate the presence of a glide [j] in Fadashi, however Bender (1989) suggests it.

information has been plotted on Figure 5 below, as best as possible.²⁸ Generally speaking, the distribution of Berta lexemes reveals a conservative periphery with a more innovative core, but there are exceptions which will be noted.

Plotting the geographical distribution of cognate lexemes reveals some interesting patterns. Certain words in the *Berta lexicon* are highly stable throughout recorded history, retaining homogenous regional distribution. Body part lexemes as well as high frequency verbs for example, show very little alteration throughout time. Other words are surprisingly variable. A number of animals have an east-west isoglossic split, with Fadashi, Wabosh and Beleje Gonfoye often sharing cognate forms, and Maiyu and Metehara forming a separate group. These are illustrated on the map in Figure 6.

2.3.1 Numerals

There are multiple numeral systems in Berta which, like other lexemes, follow an east-west distribution. Actually, there are not two, but three systems currently in use within the Berta speech area, when the borrowed Arabic system is taken into account. Neither Gaahmg nor the AKM languages share any of these systems, nor do they share each other's. After the number 'two', Maiyu and Metehara have borrowed their numerals from Arabic. Lexemes for the numbers 'one' and 'two' are closer to the western Berta variety than to the eastern, a probable indication of a closer historical relationship rather than contact.

The number 'one', in Maiyu and Metehara is *dúk'unu*. This is phonetically very close to Cailliaud, Evans-Pritchard and Siebert et al., as well as the word lists collected from Baruun be Magtole and Jebel Ura. Most eastern Berta varieties use *məneŋku* or some close variant for 'one'.²⁹ Interestingly, Tutschek's and Rossini's word lists offer items which strike middle ground between the eastern and western numerals. Their words for 'one', *mod'ogono* and *midiku* respectively, preserve the copula prefix found in the eastern varieties, but also have the post-copula alveolar stop which characterizes the western varieties, as opposed to the nasal of the eastern varieties. It is possible that these elicitations are the missing links between Berta's eastern and western numerical systems.

Turning to the number 'two', similar patterning can be observed. The eastern varieties have some form resembling *məhole*, and true of every number in the eastern dialects, has incorporated the word-initial *ma*. This is contrastive to the western variety, *hələŋəniŋ*, which, like 'one', does not have the *ma*-prefix. Here again Tutschek offers what may be a missing piece to the puzzle: *madzâling*. Once again *ma* is present in Tutschek's phonetic transcription. As in the western varieties, his grapheme concludes with a word-final nasal. From 'three' to 'ten', the western dialects' numeral records (Jebel Ura and Baruun be Magtole) all end either with a word-final nasal N or a word-final syllable NV. This may merely be coincidental, or it may be an indication of some morphological component which is not found in the eastern varieties.

Table 9: Representative numerical systems for five isoglossic regions³⁰

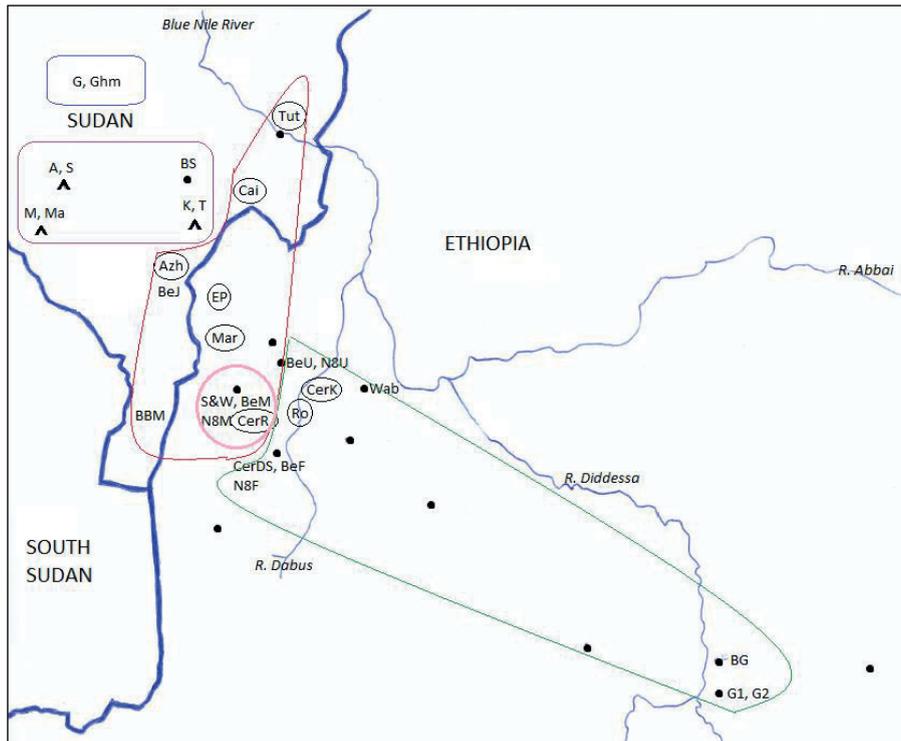
	Gaahmg	AKM	Western Berta	Asosa	Eastern Berta
one	taman	ligidi	duguni	dúk'unuŋ	məneŋku
two	dhaag	waasi	koleni	holoŋoniŋ	məhole
three	ɔɔ	eede	zitigini	talata	məte
four	yaahsah	laala	sogon	arba	man:amu
five	asaman	kamuuje	asing	hamsa	ma:koŋu

²⁸ The abbreviations of each researcher correspond to the abbreviations found in Appendices A and B. Those researchers with a circle around them are approximate, whereas the un-circled names are believed to be reliably positioned. Halevy has been omitted from the map due to an insufficient amount of comparative data.

²⁹ From a philological perspective, *məneŋku* was likely a compound, consisting of the copula *ma*, the third person subject agreement marker *-ne* and the first person possessive marker *-ŋko*. In other words, the number one has become lexicalized from "it is mine!". Still today, the first person possessive, "mine", in Maiyu is *maŋk'o*.

³⁰ The sources for each of the five isoglosses are: Gaahmg (Madal et al. (2004), the number three Bender et al. (1980)), AKM (Aka, Bender (1997)), Western Berta (Dul, Evans-Pritchard (1932)), Asosa (Maiyu, Bremer (forthcoming)), Eastern Berta (Wabosh, Bremer (forthcoming)).

Figure 5: Numerical Systems of Berta and East Jebel languages



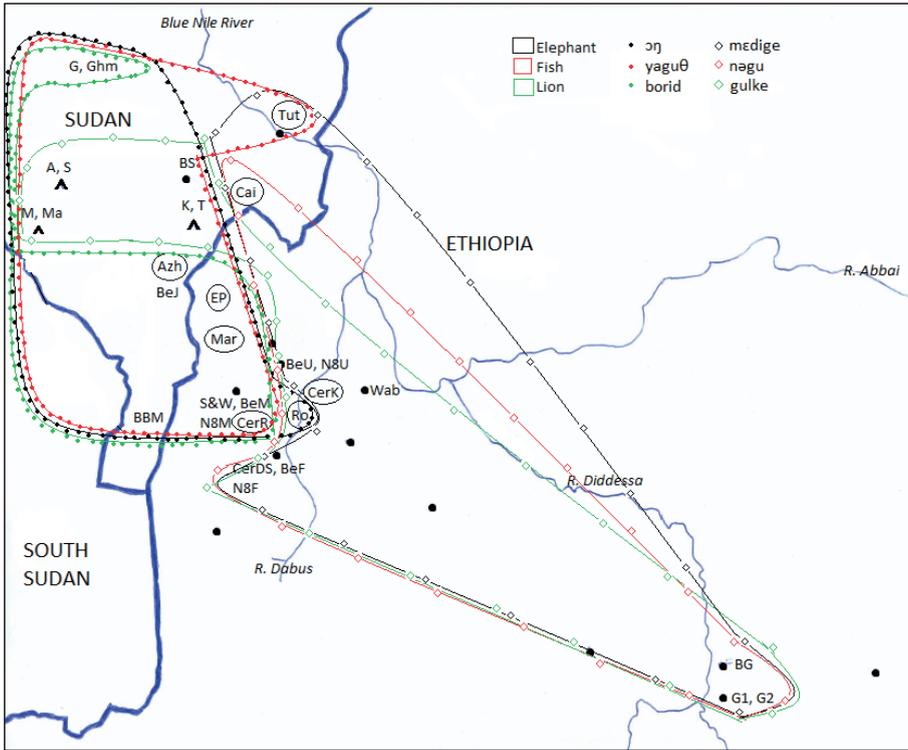
2.3.2 *Naturalia*

Most animals, especially the domesticated ones, show no lexical alteration in Berta. For instance, the words for animals such as ‘buffalo’, ‘chicken’, ‘cow’, ‘dog’, ‘donkey’, ‘goat’, ‘horse’, ‘pig’, ‘porcupine’, ‘rabbit’, ‘ratel’, ‘scorpion’, ‘snail’, ‘spider’, ‘termite’, ‘tick’ and ‘turtle’ are homogenous throughout Berta dialects, but these examples show no predictable regular sound correspondence when compared with the East Jebel languages.³¹ Lexical items referring to other animals, however, like ‘elephant’, ‘fish’ and ‘lion’ exhibit trans-Berta variation.³² None of these three lexemes show identical geographic distribution, but a general east-west trend is observable. Comparison with the East Jebel languages makes these lexemes especially intriguing, since sometimes an eastern Berta lexeme is cognate with Gaahmg and other times it is not. Their distribution can be seen on the map below.

³¹ The lexemes for ‘buffalo’, ‘pig’, ‘scorpion’, ‘termite’ and ‘tick’ are not cognate with Gaahmg and are unattested in AKM. ‘Dog’ and ‘donkey’ are cognate with AKM, but not Gaahmg. ‘Chicken’ and less likely ‘rabbit’ are cognate with both AKM and Gaahmg. ‘Goat’ is cognate with Gaahmg, but not AKM. ‘Horse’, ‘porcupine’ and ‘spider’ are cognate with Gaahmg but unattested in AKM. ‘Cow’ and ‘turtle’ are not cognate between Berta, AKM and Gaahmg. ‘Ratel’ and ‘snail’ are unattested outside of Berta.

³² The lexeme ‘elephant’ shows a nearly perfect east-west divide, but there is discrepancy between Evans-Pritchard’s (1932) Tornasi/Kelo elicitation and Bender’s (1997b). Otherwise, western Berta, AKM and Gaahmg are cognate, and the eastern Berta varieties are cognate. ‘Fish’ is difficult to pin down definitively, but there is an east-west tendency. The northern varieties, Gaahmg and Tutschek are tough to place, but they lack the word-initial nasal that characterizes the eastern dialects, and a case for similarity can be made with the western varieties. ‘Lion’ is interesting, since western Berta and Gaahmg are cognate, and eastern Berta and AKM are cognate.

Figure 6: Elephant, fish and lion



‘Tree’ is an interesting word in Berta, since it typifies a whole class of words which have undergone semantic shift due to polysemy. Amongst the Ethiopian dialects, elicitation of the English word ‘tree’ will often yield a response similar in form to either *s’is’iva* or *ingole*, depending on the isoglossic region. This is a clear example of semantic shift, and Marno provides a good clue of how it likely happened. Marno’s ‘tree’ is glossed *angole*, but ‘wood’ is *sisia*. Presently, in the Maiyu region, *s’is’iva* has come to define any generic tree, but *ingole* is a specific type of tree.

Analogy is another catalyst behind other similar examples of metonymy like ‘wealth’ and ‘chair’. The lexeme *fimbil* has come to mean ‘wealth’ in the western dialects, but is only used for ‘cattle’ in the far eastern varieties. From a pastoralist perspective, it is easy to see how these two categories are related. Similarly, ‘saddle’ and ‘chair’ are undifferentiated in eastern dialects, whereas western varieties have borrowed ‘chair’ from Arabic and differentiate it from a saddle. The same is true of ‘cup’ and ‘gourd’, where exposure to Arabic categorization has prompted Maiyu and other dialects nearby to borrow such words.

The verb ‘eat’ also shows some evidence of shift, but not in the same fashion as those examples above. In Maiyu there are four different verbs used for eating (*θija*, *k’ɔla*, *k’a*: and *s’ek’ela*), depending upon the physical nature of the object being eaten. Contrastively, most non-Maiyu dialects have just two or three of these in their lexicons: *θija*, *s’ekela* and *k’ɔla*. *K’ɔla* is almost certainly derived from the verb *k’ula* ‘to break’, and makes sense why it is reserved for harder objects like meat (often eaten together with the bones). Since the eastern varieties do not have a cognate for *θija*, it seems plausible that *s’ekela* could be the oldest word for in Berta’s inventory for ‘eat’. Cailliaud’s unique elicitation, *nangaly*, complicates this analysis but adds a further wrinkle to the enigma. *Nangaly* is significant since it provides a hint that Berta may once had a verb ‘to eat’ which is similar to the East Jebel verb *naam*. It is impossible to say for certain, but probably Cailliaud’s lexeme has a

tense/aspect suffix *-li*, affixed to the verb root, signifying a completed action by a first person actor.³³ Reconstructing this leaves a verb root **naŋ(a)*, which is quite similar to the East Jebel *naam*.

2.3.3 Polysemy and semantic shift between Berta and East Jebel

Looking beyond the low lexical similarity percentages, other comparable points are available which suggest a genetic relationship between Berta and the East Jebel languages. When the geographical scope is extended to include the East Jebel languages, certain diachronic processes are easier to see. Table 10 below juxtaposes a number of comparable lexemes between Aka, Kelo, Molo and Beni-Sheko with selections from the Berta dialects.

Table 10: Semantic shift between AKM, B-S and Berta

AKM	Gloss	Berta	Gloss
pəŋɔɾ, bəŋɔɾɔ	shield	bəŋɔɾ	soldier/boy
kəkɔne	skin (n)	kʰəkɔna	scratch (oneself) (v)
bi:di, bid:i	sun	bɛdi, bɛdi	hot
podur	star	bɔɾuru	planet/Venus
aberi, abɛra	girl	be:re	child
pɔɾɔra, fɔɾɔra	weak (ADJ)	NEG.padeʔe	not strong
imili, amilu	person (SG)	mili	black
-gɔre	MASC.SG	gure	penis
-miki	FEM.PL	mihi	women/wives

2.4 Summary

The six Berta varieties found within the current political borders of Ethiopia show significant phonetic, lexical and semantic variation. The phonological inventories of each dialect are unique (except for Wabosh and Beleje Gonfoye), and their lexicons also reflect historical drift. Nevertheless, each community included in this survey ascribes to a “Berta” ethnic identity and the historical records also point towards a common lineage. Furthermore, the historical linguistic record suggests an east-west isoglossic divide in Berta, with Maiyu and Metehara patterning closer to those Berta varieties found in Sudan, but Undulu, Wabosh and Beleje Gonfoye showing more similarity to Fadaši. This geographical delineation is not clearly visible via lexicostatistics alone, but rather in the synthesis of linguistic and socio-linguistic records. The east-west division is probably reflective of conservative versus innovative tendencies, with those varieties in the east being more conservative than those in the west.

3 Beyond Berta: Cross-lingual comparison with the East Jebel languages

This paper suggests that Berta and the East Jebel languages share a deeper genetic affiliation than originally thought. The data presented herein has been selected in order to pique the interest of those Nilo-Saharanists interested in historical reconstruction, but further research (i.e. a syntactic and grammatical comparison of Berta and the East Jebel languages, shared irregularities within paradigms, etc.) ought to be conducted to either support or diffuse this proposed merger of Berta and the East Jebel subfamily.

The topics of negation, pronouns and interrogatives have been chosen specifically because Bender (1989: 160ff.) highlights them to illustrate the disparity between Berta and East Jebel. Reopening this case is not in any way meant to defame or discredit the enormous legacy that Bender has left to Nilo-Saharan linguistics, but it is rather meant to supplement a linguistic gap with more data and fresh stimulus. Bender himself saw that Berta’s need for dialect research was desperate, saying “[i]n view of its wide geographical extent and the difference among Andersen’s (1993), Wedekind et al.’s (1993), and my (1989a) descriptions, it appears that Berta is a language badly in need of a dialect survey” (Bender 1997: 191). Based on Bender’s regular juxtaposition of Berta and East Jebel in his writings, it seems as though he was unwilling to completely discount Berta’s potential inclusion within the

³³ In Maiyu, ‘I have eaten’ is expressed *θijali*.

subfamily. In an article describing the phonology of East Jebel languages, he includes a description saying that, “Berta is a neighboring Nilo-Saharan language which is *close geographically to E.J. [East Jebel] and genetically to East Sudanic.*” (italics mine, Bender 1997b: 189). Taken alone, this is little more than a restatement of Greenberg’s own conviction. However, Bender goes a step further by recurrently juxtaposing Berta with his East Jebel comparisons.³⁴

Use of computer programs like WordSurv 7 has greatly enhanced our ability to compare and contrast lexemes from different speech varieties. For this paper I compiled eighteen historical Berta sources and nine East Jebel sources together with my own data on six Berta dialects (for more information, consult Appendices A and B). WordSurv 7 allows for easy comparison of data which represents a historical span of nearly 200 years. While my own data has added to the Berta dialect corpus, the record remains incomplete and even inadequate in some places. Word lists from Undulu and Fadashi need to be expanded and more importantly, a survey focusing on the grammatical and syntactic structures of each speech variety ought to be conducted in both Ethiopia and Sudan. Furthermore, the geographical distribution of Berta needs to be researched, specifically the northern regions of the Berta language area (in Ethiopia) stretching from Kurmuk to Guba, as well as the Fadashi variety spoken around Begi and Mendi.³⁵ Despite the gaps in contemporary records, the existing corpus is sufficient enough for a preliminary comparative analysis.

3.1 Negation

Negation varies considerably within Berta, and shows only partial commonality with East Jebel negation. Sudanese Arabic has left an indelible mark on both Berta and the East Jebel’s negation structures, but what remains unaffected shows some intriguing resemblances.³⁶ The eastern varieties of Wabosh and Beleje Gonfoye exhibit a unique negation form, namely an *sV(r)* affix. This is most likely a grammaticalized affix from the negative verb of existence, *far-i*. No other known Berta varieties share this negative affix, and it is a defining innovation of the Wabosh and Beleje Gonfoye dialects.

The *sV(r)* affix of Wabosh and Beleje Gonfoye may be related to Gaahmg’s postponed *wa*. In order for a relation to be established, then a process involving the elision of [s] and the subsequent addition of a word-initial glide [w] would need to have occurred. Correspondence between the innovative Berta affix *sV(r)* and Aka’s *iisen* is also possible. Even more reliable correspondence exists in the negated imperative between the Gaahmg *ar-* prefix and Undulu’s *ari*, in which case Gaahmg’s affix would be the more innovative form. It is also possible that Fadashi’s negated interrogative *ala* shares a common source to Undulu’s *ari*. If cog-nate, then their variant grammatical functions (imperative vs. interrogative) would be the result of diachronic grammatical shift. Adding to the puzzle, the origin of the negative imperative marker *baka* is unknown, although it may be related to the Amharic *bek’a* ‘enough’. Regardless, it does not appear to have a cognate in any of the East Jebel languages, although notably the data for the East Jebel negation table is incomplete.

Table 11: Negation in Berta and East Jebel

	Declarative	Imperative	Interrogative	Existence
Gaahmg	Postposed “wa”	“ar-” prefix +postposed “wa”	?	?
Aka	suffixes -tu, -tə, -ðo, etc.	wəl	?	?
Kelo	aa	iisen	?	wəla
Molo	(i)n-	wəl	?	iito (?)
Maiyu	wela	baka	wela	ʃari
Fadashi	wela	bəka	ala	ʃari
Undulu	wela	ari	wel:a	?
BG/Wabosh	-sa	-se	si(r)-, sa(r) -	sari, ʃari

³⁴ Other publications from 1983, 1989, 1998 all juxtapose Berta with the East Jebel languages.

³⁵ I have personally never been in this northern area, but I have recently heard reports of a unique Berta dialect spoken by those Berta residents living in Guba.

³⁶ Specifically, the negation lexeme *wela* is borrowed from Arabic (c.f. Bender 1989a: 169)

3.2 Pronouns

Comparing Berta and the East Jebel subject pronouns shows potential for genetic relatedness in most cases. Alternatively, Bender came to the opposite conclusion, firmly stating that “examination of these pronouns would surely have modified Evans-Pritchard’s thinking about the classification of the Eastern Jebel languages” (Bender 1989: 159). Bender’s comparative selection of “Berta” pronouns includes only the innovative Maiyu variety, and is therefore an incomplete comparison. Table 12 below lists all known subject pronouns found in historical and contemporary varieties. It often reveals not only vocalic variation, but even consonantal variation in most environments.

Table 12: Subject pronouns in Berta and East Jebel³⁷

	Bender (1989: 160)					Comparative Word list	
	Gaahmg	Aka	Kelo	Molo	Bender’s Berta	Beni-Sheko	Berta Variation
1SG	aan	ee	əŋ	ɔŋ	ali	iŋgi	ali, elli, ali ide, ?eli?
2SG	oon	in, ni	iŋ	ɪŋ	ŋgo	mi	iho, ŋk’o, ŋgo, əŋgo?, rho, mko
3SG	een	ənə, ɛɛne	ɛɛnɛ	een	ŋine	?	in:e, ŋine, ine, ɲine, in:e, hine
1PL	a’gan	ɛge, ɛɛgi	əy	ɔy	haθaŋ	iije	həni, hado, ado, haðaan, ngani, haθaŋ, ŋaʔ, in:ka
2PL	o’gon	ɛgu, ɛɛgu	uu	uu	haθu	uusuu	halu, dora, haθu, hau, haðu, hallu, hallo, hatu, hazu
3PL	egen	ɛge	igeegə (?)	?	mere	?	mare, mere, meeree, mæree, mæɛ?

The 1.SG subject pronoun has an intervocalic [l] in all Berta dialects. There are no East Jebel cases which have a lateral, but rather a nasal in four out of five cases. Looking at Appendix A, there are multiple examples exhibiting alternation between a nasal [N] and a lateral [l] in Berta, and further examples exist between East Jebel and Berta.³⁸

The 2.SG pronoun shows considerable variation in Berta. In most cases there is a nasal, either velar or palatal, but not in all cases. Bender (1989: 160) notes that “Berta” is unique since it includes a velar -g in the 2.SG. First, this observation does not apply to all cases of Berta, although I would concede that the proto-East Jebel pronoun expectedly had a velar stop. Assuming a form like *igo or *iko in proto-East Jebel, then a plausible sequential development could be that the nasal seen in the East Jebel languages was incipient, and the stop subsequently was elided. As is common in Aka, Kelo and Molo, the final syllable was elided. The 3.SG subject pronoun is the easiest of all to see the correlation between the E4 pronouns and Berta. Like the E4 pronouns, most Berta cases also lack the word-initial nasal found in Maiyu. Indeed, the majority of Berta 3.SG pronouns look just like what is found in Aka and Kelo, and are not far from Gaahmg or Molo either.

Berta plural pronouns are harder than the singular pronouns to reconcile with E4, however some important observations can be made. First, the plural pronouns in both Gaahmg and Aka all have an

³⁷ This table expands upon Bender’s (1989: 160) table which juxtaposed the East Jebel languages of Gaahmg, Aka, Kelo and Molo with “Berta”. The right section of this table adds Bender’s data on Beni-Sheko (the fifth East Jebel language, not originally appearing in his (1989: 160) table) as well as known Berta variants, both historical and contemporary.

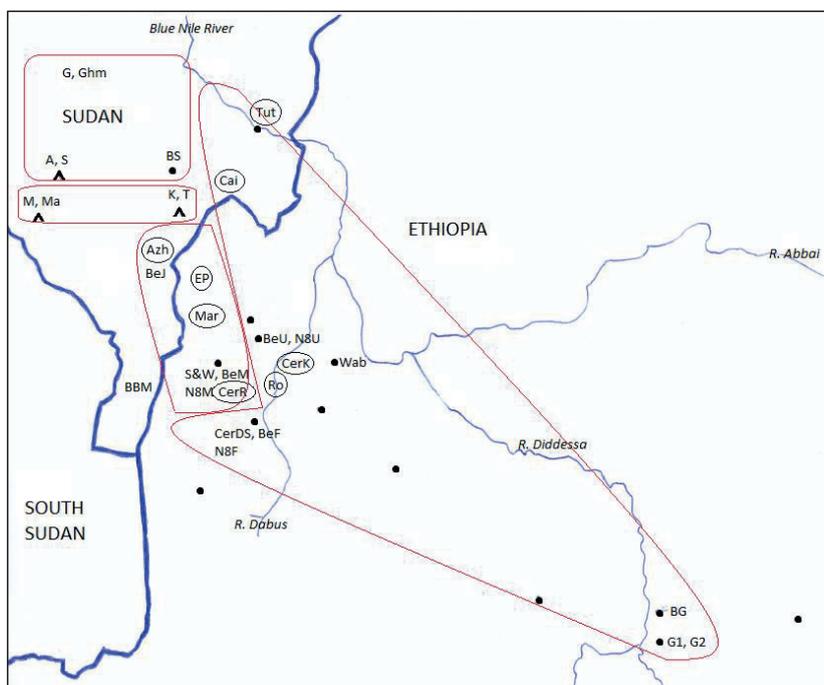
³⁸ Examples include but are not limited to “blow” (v) *fiina*, *fiino*, *piij* (A, M, K) and *fola* (B), “grain” *aajjene*, *addni*, *aajjene*, *ajjene* (A, M, K, BS) and *dzele*, *zili* (B), “hunger” *kona*, *kəna*, *kun* (A, M, G) and *hulaʔ*, *hulaŋ*, *huluy* (B), “push” (v) *ganna*, *genə* (A, M) and *dəlaʔ*, *ɛ:laʔ* (B), “say” *jəgənnə* (A) and *k’al:a* (B), “who?” *ŋa*, *inu*, *na.ŋai*, *ŋai* (A, M, K, BS) and *alo*, *dala* (B), and “yawn” *haamən*, *tahaamne*, *təʔamudən* (A, K, M) and *hamula* (B).

intervocalic [g], whereas Kelo and Molo do not. In Gaahmg, “nearly all plural marking involves the segment *gg*,” which is likely the origin of the [g] in these plural pronouns (Stirtz 2011: 99). Notably, the *gg* segment nearly always takes the slot of a suffix in contemporary Gaahmg, but in the case of the pronouns, it was likely a prefix that underwent vowel metathesis to become intervocalic. This plural marker is thought to correspond to the [y] in Aka 1.PL, and the [h] of the Berta 1.PL and 2.PL. Additionally, most of the Berta 1.PL and 2.PL pronouns still have what is likely a corresponding plural prefix **hat-* which matches the placement of Gaahmg’s plural pronoun prefix.³⁹ The **hat-* prefix is clearly seen in the Berta 2.PL, with nearly all varieties maintaining a correspondent form of **t*. The Berta 1.PL shows the same geographical isoglossic distribution seen before, having a clear east-west divide.⁴⁰

The 2.PL is less difficult to reconstruct than the 1.PL, since it shows regular correspondence in Berta. The E4 languages once again show extreme reductionism in AKM, B-S. For Berta, a proto-2.PL which incorporates a **hat-* plural prefix with the 2.SG pronoun fits all varieties (except Cerulli’s Wadashi, *dora*) nicely. All varieties reduced the first syllable of the 2.SG pronoun but preserved the vowel [u].

Turning to the 3.PL pronoun, there is little that can be offered to reconcile Berta with E4. The 3.PL pronoun lacks the **hat-* prefix found in 1.PL and 2.PL, and may be borrowed. In neighboring Gumuz, there is a *ma-* plural prefix that affixes to the 3.PL pronominal. Furthermore, *marra* meaning ‘many’ in Gumuz may offer a clue of contact or distant relatedness to the Berta 3.PL pronoun.

Figure 7: 1.PL subject pronoun isoglosses⁴¹



³⁹ Not treated elsewhere in this paper, Berta plurals can be marked with a suffix *-gu* which is thought to correspond with Gaahmg’s *gg* segment. Interestingly, in Wabosh, the plural suffix is *-gule*, which like the plural pronoun prefix **hat*, maintains a correspondent form of **t*, thus offering further evidence of a **hVt/*gVt* plural affix in Berta.

⁴⁰ Just as in the case of ‘fish’, the eastern varieties contain a nasal, [n], but the western varieties do not.

⁴¹ These distributions are perhaps controversial, but they are based on syllable structures and phonological make-up. Gaahmg (*aga*), Aka (*ε:gi*) and Beni-Sheko (*i:dʒe*) all have VCV, where the C is either a [g] or a [dʒ]. Molo (*ɔy*) and Kelo (*əy*) are *Vy*, the western Berta varieties mostly resemble *haθaŋ* (*hado*, *haðaŋ*, *a:ðaŋ*), and the eastern varieties show the most variation, but all have a word-initial or word-medial nasal. Most examples are close to a (CV)NV structure (*ŋani*, *ŋa*, *həni*), but Undulu looks different: (*in:ka*).

3.3 Interrogatives

Interrogatives, like pronouns and negation, show significant dissimilarity in Berta. Unfortunately, few interrogatives were collected for my dialect word list; my own data is incomplete for Undulu, Faḏaŝi, Metehara and Beleje Gonfoye. Apart from ‘who?’, Bender (1989) shows no variation between Maiyu, Faḏaŝi and Undulu. While Bender’s data shows little or no variation among dialects, my data (where available) exhibits significant variation. Furthermore, samples which I have taken from Maiyu and Wabosh suggest a similar east-west isogloss to what has been observed before. Admittedly, more research is needed to determine the range of their distribution within the Berta community. The table below lists five interrogatives and a question particle, and juxtaposes the four E4 languages of Gaahmg, Aka, Kelo and Molo with Berta. As before, Bender’s comparative sample is juxtaposed with known variants (both historical and contemporary).

Table 13: Interrogatives in Berta and East Jebel

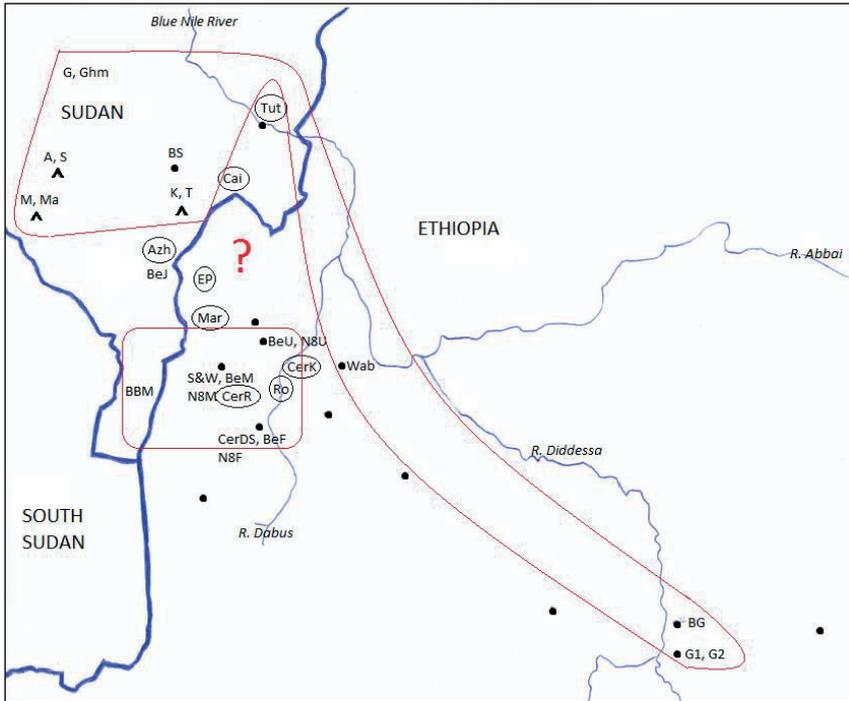
	Bender (1989: 170)					Comparative Word list
	Gaahmg	Aka	Kelo	Molo	Bender’s Berta	Berta Variation
who?	nai	ŋa/ ŋaaga	(na)ŋai/ (ka)ŋai	ɪnu/ iku	ndolo	alo, ambala, dala, mbolo, nda, ndɛɔ?
what?	ɲiin/ ɲiigin	ɲiɲi/ ɲige	ɲi/ ɲani	ni/ sini (?)	naano	ɛn, mbola, naŋ
when?	dɔi	ɔru	wɛyu	ɔru	awune	alaro, alal, amune, me:ne
where?	da	-tu	inda	inde	waane	nda, ndasin, ŋwo, ŋo, ŋwane
whence?	?	bidi, widi	?	inde	?	min wa:ne
Q (Particle)	-bɛ	-ra	ʔi (?)	--	-(y)a	-ja/- ^w a, tone

The table above reveals a significant amount of variation not only amongst the Berta varieties, but also between the E4 languages. Bender (1989: 169) writes that “[i]nterrogatives are more uniform in the Eastern Jebel languages than negatives. Still, a fair amount of variation exists.” Generally, ‘who?’, ‘where?’ and ‘what?’ are more homogenous in form in E4 than other interrogatives. Looking at these three first, we see that despite E4 homogeneity, there is considerable variation in Berta, some of which more closely resemble E4 constructions than others.

The interrogative ‘where?’ adds a further mystery to the lexical development of ‘who?’. Wabosh and Beleje Gonfoye have respective lexemes *nda* and *ndasin* for ‘where?’, which are possibly similar to the E4 ‘where?’. Whether or not ‘where?’ and ‘who?’ became semantically crossed in Berta is a valid question, but the likelihood of the Eastern Berta varieties’ ‘where?’ and the E4 languages’ ‘where?’ being cognate is better interpreted as further evidence that the conservative Berta varieties lie in the outer peripheries. Unfortunately there are too many gaps in both the historical and the contemporary records to pin down definitive isoglosses, but the figure 8 below gives a rough projection of what is currently known.

Finally, the interrogative ‘what?’ shows regular sound correspondence between Berta and the East Jebel languages. In nearly all cases, some remnant of a NVNV lexeme exists. Fleming’s G1 and G2 are exceptions and are most likely correspondent to *neŋ mba-lo* in Maiyu. This common question in Maiyu combines a question word with a relative pronoun and a demonstrative to mean “What (is) it that that (is)?” It is possible that this high frequency phrase has become grammaticalized in Beleje Gonfoye, morphing into merely *mbola* as Fleming suggests, but more research is needed to determine this.

Figure 8: Distribution of the interrogative ‘where?’⁴²



4 Statistical analysis

Lexicostatistics is a controversial tool used to hypothesize possible relatedness of languages and language families. Bender (1983: 53) uses lexicostatical percentages to fortify his claim that Berta is not related to the East Jebel languages. His statistics are juxtaposed with my own in the table below, and the sections which follow elucidate some of the possible ways in which the different statistical figures were generated. Largely speaking, my own percentages are very close to Bender’s for the AKM languages as well as AKM against Gaahmg, however AKM against Berta and Gaahmg against Berta show considerable variation. These topics and others are discussed more thoroughly in the sections which follow.

⁴² Records for the western Berta varieties show some discrepancies. Bender (1989) lists Maiyu, Undulu and Fadashi as identical, uniformly using *waane*, *wa-thaane*, whereas Cerulli (1947) presents alternatives *no* and *no* for similar regions. I can confirm Bender’s Maiyu as *waane*, but I have no records from Undulu or Fadashi. Cerulli’s word-initial nasal can also be seen in Baruun be Magtole’s *no*, and could be another isogloss used to further delineate the western Berta varieties.

Table 14: Comparative lexicostatistics⁴³

	Bender's (1983: 53) lexical similarity percentages	my lexical similarity percentages (with 50+ items)
AKM among themselves	63%	Bender avg: 63.7%, Evans-Pritchard (EP) avg: 65%
AKM against Gaahmg	34%	Bender avg: 30.5%, EP avg: 31.7%
AKM against Berta	17%	Bender avg: 31.1%, Berta WL 31.9%
Gaahmg against Berta	14%	Bender avg: 24.5%, Berta WL 24.6%

4.1 Berta lexicostatistics

Lexical similarity levels amongst the Berta speech varieties are surprisingly low and may therefore generate more questions than answers. My lexicostatistics between Berta and the E4 languages vary considerably from Bender's, so before delving into the figures themselves a brief explanation is warranted. The primary difference between Bender's results and my own can be observed at a local level amongst the three primary Berta dialects: Maiyu, Undulu and Fadashi. Bender's (1989) lexicon suggests a very high percentage of lexical similarity (in most cases these three varieties have identical lexemes in *Berta lexicon*), yielding figures that are 96% and above. The comparative percentages for Bender's Jebel Ura shows a level of variation which is more consistent with what I have observed between Berta speech communities.

It is uncertain, but one reason for Bender's uniform data may be due to the linguistic conditioning of his informants. In the case of both Undulu and Fadashi, it does not appear as though these dialects were elicited in each respective area. Rather, these were likely speakers who were found around Asosa and who had therefore had significant exposure to Maiyu. If this is so, then it is not unlikely that their responses were often Maiyu lexemes. Peer pressure can be a significant social force which dramatically affects word list elicitations, and if Bender's data collection occurred in a Maiyu area, there would plausibly be Maiyu speakers present during field work who may have impacted elicitations. Another theory about why Bender's word list is so uniform and my own so disparate, may be due to a time-tested technique commonly referred to as "cognate fishing". His 'guts, intestines' example suggests this, since his Maiyu, Undulu and Fadashi records are all cognate forms of *afɔːnʃɔros* 'lung' and not the expected *ɪu k'a:si* 'guts, intestines'. That three speakers would independently offer the same inaccurate reply is unlikely.

⁴³ The statistics found in the "My Lexical Similarity Percentages" column are my calculations for the data I used for my lexicostatistical analysis. Most of these figures reflect my scoring of other people's data for cognates between different language groups. The figures given for "Bender avg" were calculated by taking the data he published (Bender 1997: 204–214) and then comparing those AKM icons which have comparable points and averaging them against one another. I have a significant amount of data: 250 Aka icons, 230 for Molo and 242 for Kelo. Averaging the lexical similarity percentages yields a figure of 63.7% average lexical similarity within AKM, which is nearly identical to Bender's own assessment of his own data. The same procedure was also done with Evans-Pritchard's data for Sillok, Malkan and Tornasi, which also yielded a very similar figure (65%) to the number generated from Bender's data. The same AKM data was then compared to Gaahmg (the average of both Bender (1980) and Madal et al. (2004) was taken), and in this case my "Bender data" yielded a figure of 30.5% whereas Evans-Pritchard's data was slightly higher, registering a number of 31.7%. Both samples are slightly lower than Bender's own figure of 34% similarity, but neither difference is statistically significant. The next two rows show sharp contrast between my figures for Bender's data and his own calculations. In the case of the AKM and Berta comparison as well as the Gaahmg and Berta comparisons, my statistics with Bender's data is very close to the percentages generated from the entire Berta historical word list (which is a compiled average of more than a dozen different sources). Theories about how and why my calculations are so different than Bender's are explored in Section 4.1.

Due to the small comparative samples that I collected in some cases, my own lexical statistics are arguably lower than what may be generated from a larger sample. However, there is reason to believe that my statistics are credible within 5%.⁴⁴ Of further note, my informants for the Undulu and Fadashi word lists were elders (each over 60 years old), and their elicitations are not necessarily representative of contemporary speech, but rather that of previous generations. Indeed, the Undulu of my word list is thought to be endangered, and more research should be conducted soon before a more extensive record of this dying variety is lost entirely. Maiyu and the eastern-most varieties of Wabosh and Bejeje Gonfoye show the most dissimilarity, but the higher percentages of these communities with the western varieties suggests a more conservative periphery (east and west) and an innovative core.

Table 15: Berta lexical similarity percentages⁴⁵

	Fleming G1	Fleming G2	Bender Mayu	Bender Undu	Bender Fadasi	Bender Jebel Ura	Siebert and Wedekind	Baruun be Magtole	Bremer Maiyu	Bremer Undulu	Bremer Fadashi	Metehara	Bejeje Gonfoye	Wabosh
G1	100													
G2	94	100												
BeM	62	76	100											
BeU	71	80	98	100										
BeF	70	82	96	98	100									
BeJ		60	90	84	79	100								
S&W		76	86	86	86	82	100							
BBM		77	77	77	78	74	69	100						
BrM	61	73	82	80	79	82	83	73	100					
BrU			72	80	78		77	73	68	100				
BrF			88	90	91		91	87	81	77	100			
Met			91	89	87				85	68	82	100		
BG			78	82	82		83	75	62	64	68	64	100	
Wab	80	83	70	69	70	63	68	73	63	62	69	62	76	100

4.2 East Jebel lexicostatistics

This section offers a cursory treatment of the East Jebel languages, since my own figures closely mirror Bender’s. For a thorough reference, Bender (1997, 1998) offers a convincing discussion of lexical similarity between Aka, Kelo, Molo, Beni-Sheko, and Gaahmg. Additionally, Evans-Pritchard’s data is more-or-less lexically consistent with Bender’s, but due to minor variation as well as additional lexemic input, his data is included for a wider comparative analysis. Evans-Pritchard’s Aka sample has complete lexical similarity with Bender’s, Molo agrees 9 out of 10 times, but Kelo shows the most variation with just over 80% lexical similarity. Both Bender and Evans-Pritchard have an average right around 64% lexical similarity within the AKM languages.

Just two comparative sources were used for Gaahmg, and these show noteworthy variation. In this study, these two sources had a total of 236 common lexemes which have a comparative 79% lexical similarity. This number is surprisingly lower than expected, however literature on Gaahmg does make reference to multiple dialects which may accommodate for the lower figures. It is uncertain whether or

⁴⁴ My lexicostatistics between Maiyu and Wabosh changed only 3% when I increased the sample size from 200 words to 663 words. These same 200 words were sampled for Undulu, Fadashi and Bejeje Gonfoye, but Metehara has 353 words.

⁴⁵ For this and the tables which follow, cells which have less than 50 comparable items are left blank.

not the sources selected for this research reflect dialectal variation, or whether they rather manifest semantic variation between the source language (presumably English or Arabic) and Gaahmg. The table below lists the lexical similarity percentages of the five East Jebel languages, with two sources each for Gaahmg, Aka, Kelo and Molo, and one source for Beni-Sheko.

Table 16: E4 lexical similarity percentages

	Gaahmg	Gaam 1979	Pritchard Sillok (Aka)	Pritchard Malkan (Molo)	Pritchard Tornasi (Kelo)	Aka	Molo	Kelo	Beni-Sheko
Gaahmg	100								
Gaam 1979	79	100							
Pritchard Sillok (Aka)	35	39	100						
Pritchard Malkan (Molo)	33	33	74	100					
Pritchard Tornasi (Kelo)	27	30	59	63	100				
Aka	33	33	100	81	72	100			
Molo	29	29		90	70	67	100		
Kelo	31	28	81	80	81	64	61	100	
Beni-Sheko	32	34				67	64	85	100

4.3 Berta and East Jebel comparative lexicostatistics

Broadening our scope to compare known Berta speech varieties and the East Jebel languages, we see that although lexical similarity percentages between Berta and E4 are low, the percentage is higher than once presumed and this increase is statistically relevant. Although the genetic affiliation between these languages is limited, my comparative statistics approach a figure similar to Bender’s AKM and Gaahmg (34%), which was deemed high enough to warrant these five languages being lumped together to form the East Jebel sub-family of Eastern Sudanic. Furthermore, while my own lexical similarity percentages are higher between Berta and E4, my analysis for AKM and Gaahmg is actually slightly lower than Bender’s. According to my statistics, Berta shares a higher similarity percentage to AKM than Gaahmg does (although not statistically relevant), and Berta and Gaahmg display the most dissimilarity of all.

The reason for these disparate figures can only be speculated since no published records exist which show Bender’s comparative samples. While I have used Bender’s own data to compute my percentages, it must be clearly stated that I did not use the same list that Bender used, and this is almost certainly the primary reason for our different figures. Contrastively, I have drawn on numerous different sources available from both the E4 languages and Berta, and have compiled these into a single database. This approach generated an output in which some comparisons had many hundred comparable points, and others only 20 or 30. For this reason, discretion is required when examining the percentages of similarity, since the ratios underlying the percentages are different in all cases. In an attempt to weed out extraneous data, I have limited the tables found in Appendix D to only those comparisons which have a minimum of 50 and 100 comparative points. Unfortunately, this criteria discounts most of the older word lists which have less than 50 comparable items, and therefore generate an erroneous figure which must be discarded. Table 17 below depicts a condensed diagram of Berta and the E4 languages, showing some of the general cross-linguistic trends.

Table 17: Berta and E4 lexical similarity percentages (with 50+ items)

	Gaam 1979	Aka	Molo	Kelo	Beni-Sheko	Pritchard Berta (Dul)	Siebert and Wedekind	Baruun be Magtole	Bremer Maiyu	Bremer Undulu	Bremer Fadashi	Metehara	Beleje Gonfoye	Wabosh
Gaam 1979	100													
Aka	33	100												
Molo	29	67	100											
Kelo	28	64	61	100										
Beni-Sheko	34	67	64	85	100									
Pritchard Berta (Dul)	20	33		37		100								
Siebert and Wedekind	26	34	30	34	34		100							
Baruun be Magtole	28	35	29	36	39		69	100						
Bremer Maiyu	22	32	27	32	39	86	83	73	100					
Bremer Undulu	17	23	22	28			77	73	68	100				
Bremer Fadashi	19	34	25	31			91	87	81	77	100			
Metehara	17	29	26	30	35				85	68	82	100		
Beleje Gonfoye	17	30	22	30			83	75	62	64	68	64	100	
Wabosh	19	30	27	31	41	66	68	73	63	62	69	62	76	100

Lexical statistics cannot inform decisions about genetic relatedness and must therefore be held lightly. To glean a more accurate understanding of the potential for historical relatedness between Berta and the E4 languages, the comparative method, balanced with dialectology, offer solid comparable points. The historical story takes shape somewhere in the middle of these two poles. Furthermore, as I was scouring the many lexemes for cognates, my interest was piqued by the high frequency of isolated cognate examples. That is to say, there are many times that just one of the four languages of Aka, Kelo, Molo and Beni-Sheko have a cognate with either Berta or Gaahmg. Appendix E lists these isolated examples, and Table 18 presents my findings. Both offer a fresh vantage point about the possible shared histories of these languages, which arguably extends beyond borrowing. When AKM, B-S is compared with Berta, a surprisingly high percentage (just over half) of cases have at least one AKM lexemes which is similar with Berta. If the semantic shift examples found in Table 9 are combined with the statistics below, a staggering figure of 56% of the glosses surveyed have comparative similarity (either lexical or semantic) with Berta. Again, this number should not taken in isolation, but concurrently with other data.

Table 18: Combined AKM, B-S statistics

	any selection from AKM, B-S			
	not similar with Berta or Gaahmg	similar to Berta	similar to Berta & Gaahmg	similar to Gaahmg
Total	99	71	70	31
Percentage	99/271 36.5%	141/271 52.0%	70/271 25.8%	101/271 37.3%

5 Conclusion

Berta's placement in the Nilo-Saharan phylum has baffled scholars for decades. However an extensive linguistic record spanning nearly 200 years of history is not entirely void of clues useful for reconstruction. Evans-Pritchard, the first to suggest related groups (he referred to them as "Group A" and "Group B") of Berta languages, correctly identified the presence of Berta words in some of the East Jebel languages' lexicons. Years later, Bender expounded upon that theory, agreeing that such words had Berta origins, but re-labeling them as "loanwords" and deciding against Berta's relationship with the East Jebel family. Bender's postulation remains normative still today, and the general dearth of linguistic data has been the primary hindrance against placing Berta in any genetic context. While this encumbrance still continues, more information about Berta dialects is now available and allows for a fresh examination into the origins of these speech communities.

In my opinion, this data, coupled with a reexamination of the historical record, suggests a genetic relationship between Berta and the East Jebel languages. The frequency of lexical similarities between Berta and E4 hints of a shared history that extends beyond mere contact. More research is needed before an authoritative claim can be made, however based on the current information available, my hypothesis about these languages is that they should be categorized together.

Abbreviations

A	Bender's "Aka" language data
AKM	The 3 Jebel languages of Aka, Kelo and Molo
AKM,B-S	The 4 Jebel languages of Aka, Kelo, Molo and Beni-Sheko
Avg	Average
Azh	Al Azharia's Berta language data
BBM	Krell's "Baruun be Magtole" language data
BeM	Bender's "Mayu" language data
BeU	Bender's "Undu" language data
BeF	Bender's "Fadasi" language data
BeJ	Bender's "Jebel Ura" language data
BG	Bremer's "Beleje Gonfoye" language data
BrF	Bremer's "Fadashi" language data
BrM	Bremer's "Maiyu" language data
BrU	Bremer's "Undulu" language data
BS	Bender's "Beni-Sheko" language data
Cai	Caillaiud's language data
CerK	Cerulli's "Wa-Kosho" language data
CerDs	Cerulli's "Wa-Dashi" language data
CerR	Cerulli's "Rikabiyyah" language data
E4	East Jebel Language Family, including the five languages of Aka, Kelo, Molo, Beni-Sheko and Gaahmg
EP	Evans-Pritchard's "Dul" language data
G	Lexemes collected from the Gaam 1979 Dictionary
Ghm	Lexemes collected from the Gaahmg 2004 Dictionary
G1	Fleming's "Gebeto 1" language data
G2	Fleming's "Gebeto 2" language data
Ha	Halevy's language data
K	Bender's "Kelo" language data
LWC	Language of Wider Communication
M	Bender's "Molo" language data
Ma	Evans-Pritchard's "Malkan" language data

Mar	Marno's language data
Met	Bremer's "Metehara" language data
N-S	Nilo-Saharan
POA	Place of Articulation
Ro	Rossini's language data
S	Evans-Pritchard's "Sillok" language data
S&W	Siebert et al.'s language data
T	Evans-Pritchard's "Tornasi" language data
Tut	Tutschek's language data
Wab	Bremer's "Wabosh" language data

References

- Andersen, Torben. 1993. Aspects of Berta Phonology. *Afrika und Übersee* 76: 41–80.
- Bender, M. Lionel. 1983. The Eastern Jebel languages. In Ivan R. Dihoff (ed.), *Current Approaches to African Linguistics*, 1), pp. 49–57. Dordrecht: Foris.
- Bender, M. Lionel. 1989a. The Eastern Jebel languages. In M. Lionel Bender (ed.), *Topics in Nilo-Saharan Linguistics*, pp. 151–180. Hamburg: Buske.
- Bender, M. Lionel. 1989b. Berta lexicon. In M. Lionel Bender (ed.), *Topics in Nilo-Saharan Linguistics*, pp. 271–304. Hamburg: Buske.
- Bender, M. Lionel. 1996. *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom.
- Bender, M. Lionel. 1997a. *The Nilo-Saharan Languages: A Comparative Essay*, 2nd edn. Munich: Lincom.
- Bender, M. Lionel. 1997b. The Eastern Jebel languages of Sudan I: Phonology. *Afrika und Übersee* 80: 189–215.
- Bender, M. Lionel. 1998. East Jebel languages of Sudan II: Comparative lexicon. *Afrika und Übersee* 81: 39–64.
- Bender, M. Lionel. 2005. *The East Sudanic Languages: Lexicon and Phonology*. Carbondale, IL: Southern Illinois University.
- Bender, M. Lionel & Malik Agar Ayre. 1980. *Preliminary Gaam-English-Gaam Dictionary*. Carbondale, IL: Southern Illinois University.
- Blench, Roger. 2000. The classification of Nilo-Saharan. *Afrika und Übersee* 83: 293–307.
- Bremer, Nate. (forthcoming). Survey of six Ethiopian Berta dialects.
- Cailliaud, Frédéric. 1826. *Voyage à Meroé, au fleuve blanc, au delà de Fâzoql, dans le midi du Royaume de Sennar, à Syouah et dans cinq autres oasis; fait dans les années 1819, 1820, 1821, et 1822*, 1, pp. 421–25. Paris: Imprimerie de Rignoux.
- Cerulli, Enrico. 1947. Three Berta dialects in Western Ethiopia. *Africa* 17.3: 157–169.
- Conti Rossini, Carlo. 1920. Popoli dell'Etiopia occidentale. II. La tribù ed il linguaggio dei Gamilà sul fiume Dabus. *Rendiconti della Accademia Nazionale dei Lincei, Serie Quinta* 28: 319–325.
- Ehret, Christopher. 2001. *A Historical-Comparative Reconstruction of Nilo-Saharan*. Cologne: Köppe.
- Evans-Pritchard, Edward E. 1932. Ethnological observations in Dar-Fung. *Sudan Notes and Records* 15.1: 1–61.
- Greenberg, Joseph H. 1963. *The Languages of Africa*. Bloomington: Indiana University.
- Krell, Amy. 2011. A sociolinguistic survey of Ganza, Komo and "Baruun be Magtole". <http://www.sil.org/resources/publications/entry/41657> (13.05.2013).
- Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the World*, 16th edn. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com/> (25.09.2014).

- Madal, Hashim O. A., Alawia O. Yousif, Saf A.H. Ateeb & Annaim K. F. Yasim. 2004. *Gaahmg – English Dictionary, Trial edition*. Khartoum: Sudan Workshop Program.
- Marno, Ernst. 1874. Reisen im Gebiete des Blauen und Weissen Nil. In Carl Gerold (ed.), *Egyptischen Sudan und den angrenzenden Negerländern, in den Jahren 1869 bis 1873*. Vienna.
- Neudorf, Andreas & Susanne Neudorf. 2007. *Bertha-English-Amharic Dictionary*. 1st edn. Addis Ababa: SIL Ethiopia.
- Siebert, Ralph, Kati Siebert & Klaus Wedekind. Sociolinguistic survey report on languages of the Asosa-Begi-Komosha area. *Part I* (2002). <http://www.sil.org/resources/archives/9058>. (last access 13.05.2013).
- Stirtz, Timothy M. 2011. *A Grammar of Gaahmg, a Nilo-Saharan Language of Sudan*. Leiden: Leiden University dissertation.
- Triulzi, Alessandro, Atieb A. Dafallah & M. Lionel Bender. 1976. Some notes on the Ethiopian Berta and their language. *Annali dell'Istituto Orientale di Napoli* 36.1: 1–23.
- Triulzi, Alessandro. 1981. *Salt, Gold and Legitimacy: Prelude to the History of No-Man's Land, Belū Shangul, ii al-laggā, Ethiopia (ca.1800–1898)*. Naples: Istituto Universitario Orientale, Seminario di Studi Africani.
- Tutschek, Lorentz. 1850. A vocabulary of the Fazoglo language. *Proceedings of the Philological Society 1848–1850, IV*, pp. 139–153. London.

Appendix A: Historical and contemporary Berta sound correspondences

Sources Cited			
Source	Abr. Form	Meaning	Elicitation Date
Bender, Lionel 1989a	Azh	Al Azharia	1974–1979
	BeM	Bender, Mayu	
	BeU	Bender, Undu	
	BeF	Bender, Fadasi	
	BeJ	Bender, Jebel Ura	
	G1	Fleming Gebeto 1	1960's
	G2	Fleming Gebeto 2	1974
Bremer, Nate forthcoming	BG	Beleje Gonfoye	2011
	Met	Metehara	
	BrM	Bremer, Maiyu	
	BrU	Bremer, Undulu	
	BrF	Bremer, Fadashi	2013
Wab	Wabosh	1819–1822	
Cailliaud, Frédéric 1826 (French)	Cai	Cailliaud	1819–1822
Cerulli, Enrico 1947	CerK	Cerulli Wa-Kosho	1928
	CerDs	Cerulli, Wa-Dashi	
	CerR	Cerulli, Rikabiyyah	
	Ha	Halevy	1874(?)
Krell, Amy 2011	BBM	Baruun be Magtole	2007
Marno, Ernst 1874 (German)	Mar	Marno	1869–1873
Evans-Pritchard 1932	EP	Evans-Pritchard Dul	1926
Rossini, Carlo Conti 1920 (Italian)	Ro	Rossini	1920 (?)
Siebert, Siebert and Wedekind 2002	S&W	Sieberts and Wedekind	1993
Tutschek, Lorentz 1850	Tut	Tutschek	1844

The following examples aim to illustrate both the predictable and unpredictable variation which exists within Berta speech varieties. The examples below reflect both regular sound correspondences as well as less-frequent alternation between cognate forms.

The examples are clustered together to show alteration, while also noting known sound patterns in contemporary varieties. In the emboldened section, commas (,) signify free variation, slashes (/) denote phonemes in most Berta varieties, and the tilde (~) represents alteration between these phonetic sounds. Tone has been omitted.

p,b ~ d/d' ~ g,k/k'

p/g	1. tie loosely	BeMUF lip(i)s'a	Tut lag ^o za	
b/d/g	2. dirty	BBM badi	BeMUF, BRM didi	S&W gid
	3. morning	Wab nabusi	BG ma:bus	BrF mendɛʃi
b/k'/k	4. spit	BeMUF bus'a	Wab k'utʃ'a olo	BBM koze
b/g,k	5. frog	BrMF boŋoʃ	Tut gānggāss	BG koŋosi
	6. send	Wab hebina	BG ʔebene	BeUF aagena
	7. stir	Tut bera	Wab gora	
	8. false banana	BeF baθe	Wab k'əla	BrM ga:de
d/g	9. know	BeUF mo:da	BBM ma:dine	Tut māaga
	10. dish	BrM daʔhe	Wab gi:he	BeMU goi BeF gihe?

d ~ mb	1. many, much	G2 dugura, mbogura				
mb ~ nd	1. which?	BrM mba	Wab, BeU nde	BeM ande	BeJ andi	
	2. who (rel)	Wab mba	BeU mbə, mba	BrM nda		
p ~ g ~ w	1. false	BrM pus'aŋ	Tut gudzaŋ	Wab wuŋfaŋ		
β/b ~ f ~ h ~ w						
β/f/w	1. blood	BeMUF kaβa	CerR k'aafa	BrMUF kafa	G1, G2 kawa	
b/f/w	2. kill	G2 biyeetəŋo	BeMUF fiya	Wab wii doŋo		
	3. water	S&W badi	Tut feri, ferio	BeMUF firi	G2 fire, were	Wab weri
b/f	4. sweep	Wab beka alo	Tut fea	BBM fejetaŋ	S&W fɛ:ða(na)	
b/h	5. blow	Wab bəla	Tut hula			
	6. full	Tut b̄aralo	Wab h̄orolo	BeMUF h̄ori		
	7. pour	Tut b'oda; hod'a				
f/h	8. bark of tree	BeM fahaβu	G2 hohaasa			
	9. dance, play	BBM f'wazi	Wab h̄oh̄ozi	Tut h̄assa		
	10. dig	BeMUF fud̄a	Tut, BeJ huda			
f/h/w	11. money	BrM fu:da	Tut, BrF hoda	Wab wuda		
	12. rabbit	BrMU fu:fur	Tut hodzorr	BeF h̄oh̄or	BG awori	Wab awowuri
	13. white	BeMU fuudi	BrM fudi	BBM fo:di	Tut hoti	Wab wodi
h/w	14. stick	BeMUFJ hodiya	Mar hodio	S&W wodi:ʔa		
	15. tendon	Tut hoara	Wab wa:raʔ			
	16. twist (v)	BeMUF h̄orhos'a	Wab worhosa			
d,t,t̄ ~ ð,θ,l/h ~ j ~ r						
t/l	1. bow	G1 tia, lia				
d/l	2. straight	S&W gadi, go:li				
d̄,d/l	3. saddle	BrM amadi	BeMUF amodiʔ	BG: a:mele		
d,t,t̄/θ,l/h/r	4. hand	Ha deba	CerDs tabaa	EP tuba	S&W ʔaba	CerK,CerR θabaa
		Wab, G1, G2 ləba	Tut habba	Cai raba		
d,t/ð,θ/r	5. three	G1 mouidiʔ	Wab m̄o:te	Tut mote	G2 m̄o:te	CerDS muuði
		CerK mooθe	BeU muuθe	BeF m̄o:θe	Cai more	
d,t/θ,l/h	6. cow	Mar dagn	Ha taŋ	EP tung	CerK, CerDs θaaŋ	
		Wab, BG, G1,2, Ro laŋi		Tut haŋ		
θ,l	7. cut	BeMUF k'iiθa	Wab, BG k'ili			
	8. mosquito	BrF umuθ	BrMU jamuθ	Wab umuli		
l/h	9. under, below	Wab liri	Tut hiri			
ð,θ,l/h	10. finger	CerDs ðabaloo	CerK,CerR θab olo	Wab ləbəlɔ	Tut habbalɔ	
θ,l/h	11. home	BeMUF θaayu	Wab leʔoʔw	Tut haoai		
ð,θ,l/β	12. rise, stand up	CerDs að	BeMUF a'haθa	Wab ahala	Tut haya	
h/r	13. antelope	Met bohosʔ	BeMUF b̄orosʔ			
θ,l ~ dʒ/s,z/tʃ ~ k'/k,g						
θ/dʒ/s,z	1. rat	S&W amaθiŋ	Wab amadzɪ	BeUF amasiŋ	BrF amaziŋ	BrU amazi
l/k',g	2. drop	Tut d'ala	BeMUF daak'a	BBM dage	Wab d̄ak'a	

d/tʃ/k/kʻ	3. all	Tut dill	S&W tʃʻil:iŋ	BBM kiliŋ	BeMUF kʻilliŋ		
θ ~ gh	1. eat	BeMUF θiŋa	Tut ghinga				
	2. squeeze	BeMUF θɔɔra	Tut ghima				
g ~ s,z	1. rainbow	Wab amogaoʷ BeM mozolo	BrU mogol BrF mozol	BG amuʷgaʔo	Tut mássaʻll	BrM amuzalo	
k,g ~ kʻ ~ dʒ,dz							
g/kʻ	1. cut	Tut geda,gya	BeMUF kʻiira	G2 kʻila	BBM kere		
g/kʻ	2. wash/bathe	Wab gitʃʻa	Tut gidza	BeUFJ gisʻa	BrM gisʻa, kʻisʻa		
k,g/kʻ	3. elbow	BeJ kʷɔnkʷɔlɔŋ CerR kʻoŋyolo	Tut gongdʒáŋleng	EP kongkolong	BeMUF kʻoŋkʻoloŋ		
g/kʻ/dʒ	4. bone	BBM gara	BeMUF, G2, BrM kʻaara		Tut dʒare		
g/dʒ	5. thing	BeUF,G1 gin	BeM jin				
k,g/dʒ	6. yesterday	CerDS kiye S&W gi:di	Wab gi:he BrM, BBM dʒi:di	Tut gighe	BrF gi:je	BrU grfe	
g/kʻ/dz	7. peel	Wab kʻenalo	BeMUF kʻeena	Tut dzira			
	8. bitter	BrMF, Met kʻasi	Tut dʒassi				
	9. blood	Mar gafa	BG kʻawa	BrM kʻafa	Tut dʒaua		
t ~ s ~ dz	1. whistle	Wab fɛn:tina	BG wenteje	BrF, Met fensiŋa		Tut fedzinga	
d ~ k,g ~ dʒ ~ l ~ h							
d/g/dʒ/β/h	1. knife	Met handir Wab haʔhiri	BG haŋgiri BrU haŋ:hir	CerDs, R, BeMUF hanjir		G1 haiyiri	
g/β	2. snail	Tut mashgo	Wab maʃio				
g/h	3. snore	BrU gorodí	BrF harodí				
k/h	4. brain	Met afutek	BrM afuθah	Wab aʷolehi			
	5. flower	BBM kundunduluŋ		Met hudundulu			
k/h/l	6. dwarf	Met faterŋkɪfu	BrM faθiŋhiʃo	Wab wəlilisu			
d/d/dʒ,dz/ʒ ~ s,z/ʃ ~ sʻ,tʃʻ/tsʻ,tsʻ/cʻ							
dz/s,z/sʻ,tʃʻ/tsʻ/cʻ	1. earth	Tut dzadzɔ	BeM zakʻa, sʻakʻa	BBM saka	Wab tʃʻekʻa	CerDs tsʻaakʻaa	
		G1 tsagaʔ	G2 cəkʻa; cʻyəkʻa	CerK cʻaakʻaa			
dz/tʃʻ,sʻ	2. guest	Tut dzidze	Wab atʃʻitʃʻe	BrMUF asʻisʻi			
	3. rainy season	Tut adzadzá	Wab atʃʻikɛle	BeMUF asʻekʻale			
dʒ/s,z	4. moon/month	Wab, BG, G1 dʒigi	BeUF sigi	Tut, BeM, BrMU, Met zigi			
ʒ/ʃ	5. dry	S&W hoʒɛdi, hoʃɛdi					
s,z/ʃ	6. die	Wab, G2 muse	Tut muze	BBM muʃe	S&W muʃe		
	7. bad	G2 dɔgɔsi	Tut dagoazi	BeMUF dɔgɔʃi			

s/f/z	8. claw, nail	Wab maʃifo; masiso		BrU mənʃo	Tut mazo	
z/ʃ	9. knee	Tut guzuŋ	Wab guʃu	BeMUFJ guʃoŋ		
dz/z/s',tʃ'	10. dew	Tut gadzia	S&W kəzi:jaʔ	BrMUF gas'ia	Wab getʃ'ia	
dz/s/s'	11. earth	Tut dzadzʒa	Mar saga	BeMUF, BrMUF	s'ak'a	
	12. body	Tut budzege	BrU bɪsk'e	BeMUF	bus'(u)k'e	
dʒ/tʃ'/s,z	13. village	Wab he:dʒa	BG he:tʃ'a	BeU heesa	BeM heeza	
dʒ/ts'/s,z	14. grain	Wab, Ro dʒele	CerDs ts'eelee	CerK zeelee	BeUF sili	BeM zili
dʒ/s,z	15. lazy	Wab, BG dʒɔdʒɔ	BrF so:soʔ	Tut zāzā		
dʔts'/s/s'/tʃ'	16. heavy	Tut ŋedi	S&W m'ets'i	BBM ɲinsi	BeUF ɲees'i	Wab netʃi
dʔdʒ/ʃ/c	17. sit	BrF, Met dɔŋɔfa	BBM dāwafɛ	Wab dʒɔho:la	BrU ʃɔŋo:θa	G2 cohola
d/d,t,θ ~ ɲ/n/ŋ ~ g ~ h ~ ʃ						
dʔnʃ	1. fish	Tut d'āgul	Wab, BrF, BG ŋegu		BeMUJ	yu'guθ
d,t,θ/ɲ	2. sleep (v)	BeJ dirʃa	BBM tarʃine	BeUF θirʃa	Wab ɲirʃa	
d/n	3. count (v)	BrM idda	BeMUF iina			
hʃ	4. hyena	CerR aliliha	CerDs liliya			
	5. salt	BeMUF he'heʔ(y)o	G2 hehe, heeho	Wab jeho		
n/h/ʃ/θ	6. send	Tut, BrF ne'bena	Wab hebina	BrU he:bena	BG yebene	BrM agena
h/ʃ/θ	7. star	Wab hi:tʃ'o	BrM hi:s'u	BBM ji:so	BeMUF is'u	
ŋ/h	8. door	BeMUF handuŋ	Wab handuh			
	9. 2.SG	CerDs, CerR, BeU, BG ŋgo		CerK ŋk'o	Wab iho	
	10. meet (v)	Wab ŋolu woleliʒa	BeJ huluθaŋ	BrM hulu		
ŋ/h/θ	11. 3.SG	CerR, BeUJ ŋine	BrFM ɲine	BBM hine	Wab in:e	
ɲ/n/g/h	12. snore	BG ɲardɔdɔ	Wab ɲardɔdɔ	BrU gorodɔ	BrF harodɔ	
ŋ/g	13. throat	Tut ŋgallo	BrM ŋalu	Mar gualu		
ʃ/p' ~ m	1. keep away	Wab ʃadene	BrM p'a:rane	BrFU mɔriene		
b ~ l	1. darken	BrU am:buwe	Wab ami:loʒe	BG aŋ:a miloʒe		
	2. python	BeF habali	BrM halali			
l ~ ɲ ~ n ~ ŋ						
l/ɲ/ŋ	1. donkey	Tut ʃilerr	BrMUF ʃɲir	BeMUF ʃɲir		
l/n	2. hear	Wab hale liio	BeMUFJ haala	BBM ha:nio		
	3. four	Ro malamumu	CerK, CerDs mannamu	BeU mənɔmu	Tut manamo	
	4. blow	Wab bo:la	BeMUF, BrM fo:la	Tut hula, hāna	Bar fune	
l/n/ŋ	5. horn	Tut balulo	BeJ bulunʒ	BeMUF bo:loŋ		
ɲ/n/ŋ	6. be pregnant	BrFM olgɲi	Wab wolgine	BrU woldʒinʒi		
	7. red	BeMU beŋeni	Tut bene, beni	BrF bini	Wab beŋaʒi	S&W beŋeniʒ
l ~ m	1. lamb	BeMU k'elme	BeF k'ombe			
m ~ ɲ ~ n ~ ŋ						
m/n	1. morning	BrU mendʒɛf	BG ma:bus	BrF mendɛʒi	Wab nabusi	

m/ŋ	2. return	BeMUF mooða	BrF moliaʔi	BrM meθa	Wab ŋoloʔe	Tut ŋoa
m/ŋ/ø	3. suckle	Wab tʃʷumbina	BeMUF sʷuŋθiŋa	Tut dzabita		
m/n/ŋ	4. vomit	S&W gogoma	Wab guguna	Tut guinga	BeMUF	guuʷguuŋa
ɟ/n/ŋ	5. whistle	BG wentepe	Wab fɛn:tina	Tut fedzinga	BrF, Met	fɛnsiŋa
m ~ n/ŋ ~ w						
m/n/w	1. when?	BrM amune	BBM me:ne	S&W awuɛʔ	BeMUFJ	awine
ŋ/w	2. sit	Met dɔŋoʃa	BeMUF dɔŋɔʃa	BBM dɔwɛʃe		
d ~ r						
	1. cut	Tut geda	BeMUF kʷiira	BBM kere		
	2. keep away	Wab ɓadene	BrM pʷa:rane	BrFU mæriene		
	3. sweat (n)	Wab bædɛŋga	S&W bæruwɛŋ			
	4. water	S&W badɪ	Wab wɛri	Tut feri, ferio	BeMUF firi	G2 fire, were
	5. eight	CerDs θaboode	CerK θaboore	BeUF θaboore, θaboothe		
	6. pull	Tut fida, fira				
b ~ r						
	1. night	Wab habole	Tut habien	BeM habiθɔŋ	Mar	naharale
s,ʃ ~ r						
	1. testicle	Tut dosi, dori	BrM θoriʷ	EP turi		
	2. Venus	BeF busura	Wab bɔrɔrɔ			
l,θ,t ~ r						
l,θ,t/r	1. smoke	Wab, BG dɪla	BeMU dɪuθa	BeF dɪuθa, dɪla	BBM du:ta	S&W u:raʔ
l/r						
	1. stone	Tut bele, bar				
	2. thunder	Wab rɔbale	BeMUF buluŋ	roo Tut bare	BBM	rober
	3. when?	BeMUF alal	Wab alaro			
	4. whirlwind	Met agulgulu	BrU adurdur			
n ~ r						
n/r/ø	1. excrement	BeU ŋgun	BeM ŋgu	BeF ŋgu		
n/r	2. peel (v)	BeMUF kʷeena	Tut dzira			
	3. bad, ugly	Tut zuni	BrM, S&W ʃuni	Mar schurai		
h ~ s						
	1. cough (v)	Wab hohona	BeJ hohoŋa	Tut hosiŋa		
h ~ k/kʷ						
h/kʷ	1. liver	Wab a:nəhe	Tut nehe	BBM nanaheʔ	BeMUF	nekʷei
h/k	2. tongue	Wab, BeF, G1 hale	BeJ kala	EP kula		
h ~ r ~ x						
	1. exit	BeMUF huha	Wab hɔra	BeJ xooxʷa		
ʔ ~ d						
	1. louse	Wab aʔi:ni	G2 aini olo	Tut dʃiŋi	BeMUF	dʃiŋi
ʔ ~ ʃ						
	1. no	Wab iʔi	BeMUF eʔee, eeʷyo, eʔeeyo		Mar	ischu

Appendix B: Sound correspondences within AKM, B-S and Gaahmg

Sources Cited		
Source	Abbreviated Form	Meaning
Bender and Malik 1979	G	Gaam 1979 Dictionary
Bender, Lionel 1997b	A	Aka
	M	Molo
	K	Kelo
	BS	Beni-Sheko
Evans-Pritchard 1932	S	Sillok
	Ma	Malkan
	T	Tornasi
Madal, Yousif, Ateeb and Yasin 2004	Ghm	Gaahmg 2004 Dictionary

b ~ ɸ	1. sun	A, M biidi	K, BS ɸiidi			
b ~ d ~ g	1. ashes	AMK tɔdɔ	G, Ghm duug			
	2. bark of tree	A bəɸba	G baga	M gɔga		
	3. ear	BS si	K siidi	A sigii, sigiide	M sigi	
	4. field	M əddɔfa	K aagewi	G gɔfa	Ghm gaffa	A kɔfa
	5. fish	M uudo	A ɔɔgu			
	6. friend	K mərde	A mərgəna			
	7. scratch	A abbəbaano	K dʰedʰene	M g ^(ɸ) eegero		
	8. shoulder	M bɔreda	A pɔrga			
	9. smooth	A darbəde	M dərgəje			
t ~ d/d	1. drink (v)	A mætu	M mootu	K mədeə	BS mɪdi	
	2. fall (v)	A bətese	K bɔdia, podia			
t ~ k/g	1. full	M totuudu	A tɔkɔjaw			
	2. leaf	K kɔca siiti	A kɔc.a.sigi	M kɔs.a sigi		
	3. lion	A tɔk ^w e	K tɔgu, dɔku	BS tuuke	M kɔdu	
	4. liver	BS tee	K teete	A tɔgi		
	5. push (v)	K tɔgenɔ	A gɔganna			
	6. want	M tɔli	K dɔɔy	A ɔgole		
p/b ~ f ~ w	1. alive	K beebeə	M feefere			
	2. blood	K g ^w eeba	A, BS g ^w ɛfa	M g ^w eewa		
	3. clean	G ber	M waari			
	4. field	A kɔfa	M əddɔfa	K aagewi		
	5. mountain	A peela	K bela	BS beela, bela	M feła	
	6. blow (v)	K pii ^ɸ	A fiina	M fiino		
	7. donkey	A pildi	M filiir			
	8. feather	A pɔta	K pɔtɔ	BS pɔɔdo	M fəta	G, Ghm fiid
	9. smoke (n)	A pɔru	K pɔyu	BS boru	M fɪr ^ɪ	
k/g ~ w ~ ʔ	1. black	A ɔɔɔga	M, BS ɔɔ(ɔ)wa			

	2. fear (v)	A lægɛdi	K ləwdi		
	3. goat	A lago	M ləw	K laawo	
	4. good	A. kərrə	M ware	K wəədi	BS waar
	5. long, far	A fuluuge	M fəluugo	BS uuluwe	K əluʔo
	6. lose, hide	A dugi	M d(u)woo		
	7. night	A ɔrga	K ɔrrə, ɔyyɔwa	BS ɔɔrə	
f ~ w	1. chief, king	A fenedi	Ma fonodi	G wɛɛda(n)	
b ~ s	1. bark of tree	A bəəba	K səba		
d ~ s,z	1. horn	M kɔdul	A kəsuli.ge	K kəsəl.tə	BS gosuuli
	2. kill, punish	M fəodu	BS bəəsi	K pəəsi, bəəsi	
	3. knee	Ma kudu	M kɔdu	S kusu	A, K, BS kosu G, Ghm kusumi
	4. knife	M kidəm	A kisəm	K kesa	
	5. two	M, K waadi	BS wadi	A waasi	
	6. sand	M ɔdəəla	A assessel		
d ~ dʒ ~ s ~ k,ç	1. fly (n)	K ɛɛje	BS geje	M gu'soi	A geece
	2. be green	K jɛriyə	M see'ro	BS eekeera	
	3. hand	M soro	BS cəro, sərə	A carga	K cəyə
	4. heavy	K lijə	M liiso	A liice	
	5. how?	A jaani	M sɛɛn	K kye ⁹	
	6. star	M mideero	K maaju w(ey)a	BS maceera	
	7. tooth	K niidi	BS niidi	M niisu	A niice
	8. tree	K kəja	M kəsa	BS kəca	A kəca
d ~ dʒ	1. crop, grain	M addmi	A, K, BS aajjene		
	2. full	M totuudu	A təkəjaw		
	3. rub (v)	M ebeedo	A efeje	K abbeje	
	4. smooth	A darbəde	M dərgəje		
	5. urine	K (d)jəw	A jaa	M jəw	
	6. star	A podur	G, Ghm pəjə		
k/g ~ dʒ ~ ʔ,q	1. louse	A gene	M giŋə	G, Ghm jiŋd	
	2. meat	K kəkəkə	A ɔʔə	BS ɔʔə	
	3. moon	A, M aguwa	K aʔowa	BS aqa	
	4. thorn	A kiigi	M jiiji	K kiɪʔi	
	5. rainy season	A məkase	M mukəm	K məʔase	
	6. stab, pierce	A kaageno	G kən	M jeeno	
d ~ l	1. one	M deedi	A, S ligidi	Ma ledi	
f ~ s	1. swell (v)	A affai	K essəi		

s ~ z	1. all	M tıbbısi	A táb'bizı			
	2. sit down	A abbese	K εbbε'sı	BS aabbızı		
s,z ~ l	1. brain	Ghm laaza	G lała			
	2. full	Ghm parras	G pərł			
	3. grandmother	Ghm taaz	G taal			
	4. shadow	Ghm kuuz	G kuul			
s ~ c ~ g	1. manure	M sɔɔ	A cɔɔgo	K cɔɔdi	G gɔɔd	
t/d ~ r ~ j	1. donkey	A pıldı	M fıliir			
	2. down, low	G, Ghm taad	A taarga	M taara	K taaya	
	3. eye	K εedi	G εd	Ghm ed	A ere	M eri
	4. good	K wɔɔdi	G wɛɛda(n)	A kɔrrɔ	M ware	BS waar
	5. hard	A pırgıdi	M dımgıri			
	6. rabbit	M bɔta	A cɔbɔrɔ	K cɔbɔyɔ		
	7. spear	K wɛɛda	A weere	M wɔr ⁹		
r ~ l ~ j	1. back, spine	A arra	K ɔɔyɔ			
	2. bad, ugly	A wara	K wɔya			
	3. down, low	A taarga	M taara	K taaya		
	4. earth, land	A kɔrɔ	M kala	K kɔyɔ		
	5. be green	BS eekeero	M see ⁹ ro	K jeriyɔ		
	6. hand	A carga	M sɔro	BS cɔro, sɔrɔ	K cɔyɔ	
	7. honey	A keerdi	M kɔrrdı	K keydi, keydimo		
	8. know	BS kɔɔrmi	K k ⁹ ɔyɔ			
	9. new	A araaya	M araaya	BS aayara	K ayaya	
	10. ten	A caraare	K k ⁹ eeraali			
	11. turtle	A pɔrɔŋa	K bɔyɔɔŋa			
l ~ n, ŋ	1. big, large	BS yaamɔdɛl ⁱ	Ghm dhomol	K amadenɔ	G mɔdɔ(n)	
	2. blow (v)	G bel	A fiina	M fiino	K pii ⁹	
	3. stick (n)	Ma loro	M ŋɔra			
l ~ m	1. person	A ele.gora	M ulisɛkɔlo	K amɔlɔra	BS amoolorɔŋ	
m ~ b	1. crooked	M kamnu	A kɔɔbino			
	2. drop	M mijiro	A bijinna			
ŋ ~ d	1. cry, mourn	A ɔŋa	M aado			
ŋ ~ k/g	1. come	K laaŋ	A, M laako	BS laago		
	2. hair	BS j ⁹ ɔŋge	A jeke			
	3. near	K leeŋ ⁹	A leeŋi	M loogi		
ŋ ~ ŋ ~ n	1. food	T nyuma	Ma ngam	G, Ghm names		

ŋ ~ s	1. enough!	M məŋane	K məs'səne	
m ~ s	2. rainy season	M mukom	K məʔəse	A məkase
n ~ r	1. scratch	A abbobaano	K dʰedʰene	M g ⁶ eegero
	2. drop	A bijinna	M mijiro	
	3. run	A daan	M dɔɔni, dɔɔŋi	K ɓeeri, ɓeria
ŋ ~ ʃ	1. break	K bɔ'gɔŋgo	A bəgɔyo	
h ~ g	1. tomorrow	M mər(u)hudɔ	A margude	
h ~ ʔ	1. yawn	A haamɔn	M tahaamne	K tɔʔamudən
r ~ ʔ	1. breathe	A saar	M saari	K sɔɔʔi
ʃ ~ ʔ	1. basket	K gɔdaiyo	A gɔtaʔa	

Appendix C: Locations and dates of sources

Publication	Language (variety) name and abbreviation	Informant info, locations and notes
Bender 1997b	Aka, A	This is the same as Evans-Pritchard's Sillok "found around several Jebels (hills) in eastern Sudan south of the Ingessana Hills in Blue Nile Province" (Bender 1997: 189). Specifically, "Aka is spoken around Jebel Silak, south of the Ingessana Hills, west of Keli (11.10 N 33.45E)" (Bender 1997: 190).
	Molo, M	E-P's Malkan "Molo is spoken at Jebel Malkan, south of Jebel Silak (10.50 N, 33.45 E)." (Bender 1997: 190)
	Kelo, K	E-P's Tornasi "spoken in and around the village of Keli, on the main road south of the Ingessana Hills in east-central SudanC (10.50N, 34.20 E) (Bender 1997: 190)
	Beni-Sheko, BS	This language is located on Jebel Beni Sheko, "a striking 'Inselberg' on the main road north of Keli" (Bender 1997: 190).
Cerulli 1947	Wa-Kosho, CerK	Cerulli describes the linguistic area in general terms, saying that the Berta "...inhabit the region south of the Blue Nile and east of the former Sudan province of Dar Fung. Their eastern limit in Ethiopian territory is usually indicated as the River Dabus (or Yabus); but some scattered groups of Negro tribes, probably akin to the Beni Shangul, live east of the Dabus as far as the confluence of the River Didessa with the Blue Nile." (Cerulli 1947: 157). Although he provides a map (158), he does not give details of where the three dialects of Wa-Kosho, Wa-Dashi and Rikabiyyah are located, only that they are "spoken by different tribes in the Beni Shangul area" (Cerulli 1947: 157).
	Wa-Dashi, CerDs	
	Rikabiyyah, CerR	
Tutschek 1850	Fazoglo, Tut	Tutschek's word list "was collected from a boy born at Hobila, in the south of the Fazoglo country, purchased out of slavery at Alexandria by the Duke Maximilian." (Tutschek 1850: 139). Unfortunately I have not been able to indentify Hobila on any

		maps, and so I placed Tutschek's source not far from the town of Fazughli (based on the language name).
Bender 1989b	Mayu, BeM	"The Arabized Berta of town-dwellers around Asosa." (Bender 1989b: 271)
	Undu, BeU	"The 'pure' rural Berta... found in the area north of Asosa, western Wellegga Province, Ethiopia" (Bender 1989b: 271)
	Fadasi, BeF	"The Berta of the Bambeshi (Fadasi) area, southeast of Asosa." (Bender 1989b: 271)
	Sudan (Jebel Ura), BeJ	"Collected in Khartoum from a speaker of the Jebel Ura area, near Kurmuk. Numerals collected at Jurut, near Kurmuk" (Bender 1989b: 271)
Al Azharia Jahn 1970, in Bender 1989b	J: Sudan, Azh	Collected in Sudan by "Samia Al Azharia Jahn in the 1970's; her main concern was with ritual and ethnomusic." (Bender 1989b: 271)
Fleming 1960s, in Bender 1989b	G1 Gebeto	"Gebeto of the Diddesa bridge area, western Wellegga" (Bender 1989b: 271)
Fleming 1974, in Bender 1989b	G2 Gebeto	"Gebeto of the Diddesa bridge area, Dimtu village, forty kms. west of Gimbi" (Bender 1989b: 271). ⁴⁶
Bremer (forthcoming)	Fadashi, BrF	A Fadashi village, 13 kilometers north of the town of Bambassi.
	Undulu, BrU	The town of Undulu, northeast of Asosa.
	Metehara, Met	The displaced Berta living nearby the towns of Metehara and Awash, approximately 250 kilometers east of Addis Ababa.
	Beleje Gonfoye, BG	Speakers located in the small town of Fwafwate, about 30 kilometers north of the town of Arjo and likely about as far from where Fleming did his surveys.
	Wabash, Wab	Those Berta living in and around the town of Daleti, 42 kilometers north on the road taken from the town of Mendi.
	Maiyu, BrM	Those Berta living in and around the main towns of Asosa, Menge and Kurmuk.
Krell, 2011	Baruun be Magtole, BBM	This survey was collected in a village called Yabus al-Ghaba, located in the Blue Nile Region of Sudan, not far from the town of Yabus.
Siebert, Siebert and Wedekind 2002	Berta (Asosa), S&W	This word list "was provided by Asherif Ahmed (21 yrs.), who is a citizen of Asosa; he also speaks Arabic, Amharic, and English. The word list was transcribed and recorded by K. Wedekind." (Siebert, Siebert and Wedekind 2002: 17).
Halevy 1874, in Cerulli 1947	?, Ha	?
Marno 1874 (481–495), 132 words	Bertat, Mar	Marno's first journey took him from Cairo, Egypt to Bambasi, Ethiopia. His journey, recorded in German, notes his travels through "Dar Bertat", which began from Fassoql, went through Famaka, into "Beni Schanqol" and ending at Bãmbãschì. In his description, he distinguishes between what seems to be a general "Bertat", and specific Berta peoples "Fossoql" and "Bãmbãschì". Also, his map seems to suggest that Beni Schanqol is a synonym for modern-day Asosa. The language described in his word list is "Bertat", which could easily have been elicited anywhere between Fassoql and Bambasi, and could reflect any of the dialects found therein: Maiyu, Fadashi or Undulu.

⁴⁶ This is almost certainly a mistake; it should read "east" of Gimbi. To my knowledge, there are no Berta living 40 kilometers west of Gimbi, and the Didessa Bridge is indeed east of Gimbi.

Rossini 1920	Gamila, Dabus River, Ro	This word list was elicited by Gimiri Arar, a captured slave. Rossini's document suggests that Gimiri believed his language to be different from "the Fazoklo and Barta" (Rossini 1920: 319).
Evans-Pritchard 1932	"Dul", EP	Evans-Pritchard, like Marno, travelled extensively amidst multiple people groups living nearby the present Sudan-Ethiopia border. He refers to this area as "Dar-Fung" and to the Berta as "Dul". Although the town of Dul was not included on my map, it is in close proximity to Kurmuk, so 15 miles southeast, but it's not certain that this was the location for Evans-Pritchard's wordlist.
Cailliaud	?, Cai	?

Appendix E: Rare Berta lexemes that have similarities with some E4 lexemes

Gloss	Berta		East Jebel	
	Source	Lexeme	Source	Lexeme
basket	Tut	<i>ngânde</i>	Kelo	<i>gə'daiyə</i>
break (v)		<i>bezoa</i>	Aka	<i>begɔɔ</i>
centipede		<i>dagalgazang</i>	Gaahmg	<i>dongol</i>
fat, thick		<i>dundulung</i>	Aka, Kelo	<i>dɔŋɔl</i>
near, close		<i>nidze</i>	Aka, Molo	<i>leegi, loogi</i>
forget		<i>d'odzoinga</i>	Kelo	<i>yiziŋa</i>
frog		<i>gānggāss</i>	Aka	<i>gaage</i>
give		<i>dia</i>	Beni-Sheko	<i>bi</i>
thin		<i>d'afet</i>	Beni-Sheko	<i>deede</i>
touch		<i>bāaza, bādza</i>	Kelo	<i>bə'basəna</i>
good		<i>dzab</i>	Gaahmg	<i>jerrad</i>
pour (v)	Tut, BrM	<i>b'oda, p'u:da</i>	Aka	<i>peegɔ</i>
rat	Mar	<i>doroa</i>	E4	<i>dii, tii, tia</i>
there		<i>ati</i>	Aka, Molo, Kelo	<i>tiiti, tiide, tii</i>
belly, abdomen	Ros	<i>ile</i>	E4	<i>e:lg, elle, illi, ele</i>
sister		<i>mbune</i>	Aka, Molo, Kelo	<i>moguna, muuna, mɔɔne</i>
elephant	Mar, Ros	<i>un, oŋi</i>	E4	<i>ange, rrɔŋɔ, krɔŋɔ, ɔɔŋɔ</i>
girl	G1	<i>here, beri</i>	Aka, Kelo	<i>aberi, abera</i>
hot		<i>hulu?</i>	E4	<i>luhluh, luuge, luu(wa), lua</i>
new	G2	<i>hara</i>	E4	<i>yaa, araya, ayaya, aayara</i>
seed		<i>beceri</i>	Aka, Molo	<i>wiico, wɔɔso</i>
lick	BeJ, BrM	<i>raaθa, raθa</i>	Aka	<i>rɛɛde</i>
down, low	BrM	<i>θiri</i>	E4	<i>taad, taarga, taara, taaya</i>
field, pasture		<i>gafa</i>	E4	<i>gaffa, kafa, əddafa, aagewi</i>
want, hunt		<i>k'aba</i>	Kelo	<i>kəbəne?</i>
beard		<i>adigiŋ</i>	EPT	<i>andiginbuna</i>
hyena		<i>bɔŋu</i>	EPT	<i>bungu</i>
cry, mourn		<i>hoona</i>	Aka	<i>ɔɔŋa</i>
mountain	BrM, BBM	<i>bele, bele</i>	Aka, Molo, Kelo, BS	<i>peela, fela, bela, beela</i>
deaf	BrF	<i>ma:la</i>	Molo	<i>miji</i>
jump	BrU, BrF	<i>ho:re, gu:dɛ</i>	Aka	<i>peede</i>
few	Wab	<i>neŋ:gel:e</i>	Kelo	<i>kedɛnele</i>
love, like		<i>gɔdi l:are</i>	Molo	<i>taaridiade</i>
why		<i>malu^wa</i>	Aka, Molo	<i>aal, aali</i>
monkey		<i>ori</i>	Aka, Kelo	<i>teeri, toi</i>
where	Wab, G1	<i>nda, ndasin</i>	E4	<i>daa, tuu, mde, intə, inda</i>
star	BG	<i>bəruru</i>	Aka	<i>podur</i>
throwing stick	Azh	<i>hankol</i>	Aka, Kelo	<i>kooda</i>
taste (v)	BeM	<i>geena</i>	Aka	<i>duugəna</i>
friend	BeMUF	<i>amedel</i>	Aka, Kelo	<i>mərgəna, mərde</i>
person, man	Cai	<i>meloko</i>	Kelo, BS	<i>amɔlɔra, amoolorɔŋ</i>

Appendix F: Rare E4 lexemes that have similarities to Berta lexemes

Gloss	East Jebel		Berta
	Source	Lexeme	Lexeme
bite (v)	Aka	<i>gorri</i>	<i>gora, k'ora</i>
dirty		<i>doogade</i>	<i>didī, donḡon, gidanḡa, mili dokochi</i>
gourd		<i>ææbo, abo</i>	<i>abube, ago</i>
say, answer		<i>jegesi, jəḡanna</i>	<i>dʒalla, dzedzarna, k'ala</i>
see		<i>gəllako</i>	<i>gaayela</i>
shadow		<i>korre</i>	<i>agijjikeri, angijir, a:foren, asori</i>
short		<i>ledor</i>	<i>lihidi, duguri, k'idere</i>
rub (v)		Aka, Gaahmg	<i>ɔɔḡ, guur</i>
sand	<i>assessel, sassa</i>		<i>ase:sege, fe:k'e</i>
wait	Aka, Kelo	<i>tiḡano, diḡtasi</i>	<i>diḡa</i>
smooth, soft	Aka, Molo	<i>darbade, dəḡəje</i>	<i>daadas'i, de:de:si</i>
beer	Aka, Molo, Gaahmg	<i>baago, bawodi, saa</i>	<i>ba:ɸ'a, baaca, ba:s'a</i>
bone	Kelo	<i>ora</i>	<i>k'ara, gara</i>
cloud		<i>aaguiyu</i>	<i>angoru, a'goru</i>
come		<i>ɔya</i>	<i>adɔ</i>
granary		<i>luuba</i>	<i>lu:be, luuba</i>
river		<i>teeli</i>	<i>dale, ʰele, deli</i>
sky		<i>ɣombom</i>	<i>rump'uḡ, rumbun</i>
stick		<i>kodia</i>	<i>hodiya, hodio, wodi:ʒa</i>
strong		<i>bəəde, bɔəde</i>	<i>badiya, beɸiɸi</i>
yesterday		<i>d^v</i>	<i>gi:he, gife, gi:di</i>
big, large		Kelo, Beni-Sheko, Gaahmg	<i>amadenə, yaamədel, dhomol</i>
clean (adj)	Molo	<i>waari</i>	<i>s'ari, s'ara</i>
skin	Beni-Sheko	<i>səəba</i>	<i>fəwa, fafa, fəfa, zafa</i>
crocodile ₁	Gaahmg	<i>dhagar</i>	<i>dabro, daburu,</i>
crocodile ₂		<i>naan</i>	<i>nagu</i>
grain		<i>jer</i>	<i>dzele, sili, zili</i>
lion		<i>bahrd, bərd</i>	<i>bɔərid, purid</i>
hyena	EPT	<i>bungu</i>	<i>bumu?, bo:ḡu</i>

THE DIACHRONIC DEVELOPMENTS OF *KI* CONSTRUCTIONS IN THE LUO AND KOMAN FAMILIES

Richard Griscom

1 Diachronic syntax: The comparative method, grammaticalization, and construction grammar

The field of diachronic syntax has its roots in the comparative method, a means of determining language relationship and change through the comparison of individual features across different languages. The comparative method has been in use in various ways since antiquity, but was first refined and standardized during the 19th century by Rasmus Rask, Karl Verner, Jacob Grimm and the neogrammarians following Sir William Jones' observations of the similarities between Sanskrit, Latin, and Greek in 1786. Schlechter (1874) was the first to posit reconstructed forms for a proto-language based on sound correspondences between Indo-European languages. Despite recent critique of its limitations, the comparative method stands out as the most capable approach of determining genetic relationships between languages (Cambell & Poser 2008).

Until the 20th century, the field of historical linguistics was primarily focused on the comparison and internal reconstruction of phonological features and cognate morphological and lexical elements. The reconstruction of syntactic structures was not seen as a worthwhile endeavor by the neogrammarians, and it was not until 1912 when Meillet first proposed grammaticalization as a process through which lexical items become integral to the grammatical system (DeLancey 2004). Since the 1970's, grammaticalization has been extensively researched both cross-linguistically and within individual languages (Heine & Reh 1984; Traugott & Heine 1991; Hopper & Traugott 1993).

More recently, the development of theories of construction grammar has had a significant impact on approaches to diachronic syntax. Authors including Goldberg (1995) and Croft (2001) have argued that the symbolic units of language must necessarily include compositional schemas, rather than merely atomic units, and that these schemas must be meaning-bearing units in their own right. The relationship between diachronic construction grammar and grammaticalization remains to be clarified (Noël 2007), but within the past decade diachronic construction grammar has won increasing recognition as a viable means of reconstructing the syntactic structures of proto-languages and the stages of development that lead to synchronic syntactic patterns. The present study takes such an approach in order to further understand the relationships between grammatical constructions in two language families: Luo and Koman.

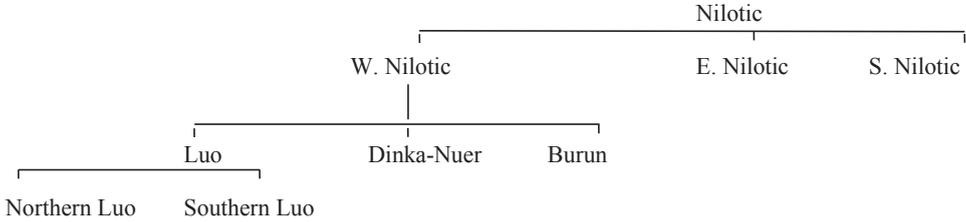
2 Language background

The Luo family is one of the three primary families of Western Nilotic (Dinka-Nuer, Burun, Luo), and itself is divided into two groups: Northern Luo, and Southern Luo (Storch 2005). The Northern Luo languages are spoken primarily in South Sudan and on the Sudan-Ethiopia border and include Shilluk, Belanda Bor, Thuri, Luwo, Anywa, and Pāri. The Southern Luo languages are spoken mostly in Northern Uganda and Western Kenya, and include Adhola, Kumam, Dholuo, Alur, Lango, Acholi, and Labwor. The categorization of the Northern Luo and Southern Luo languages within the Nilotic family is presented in Figure 1.

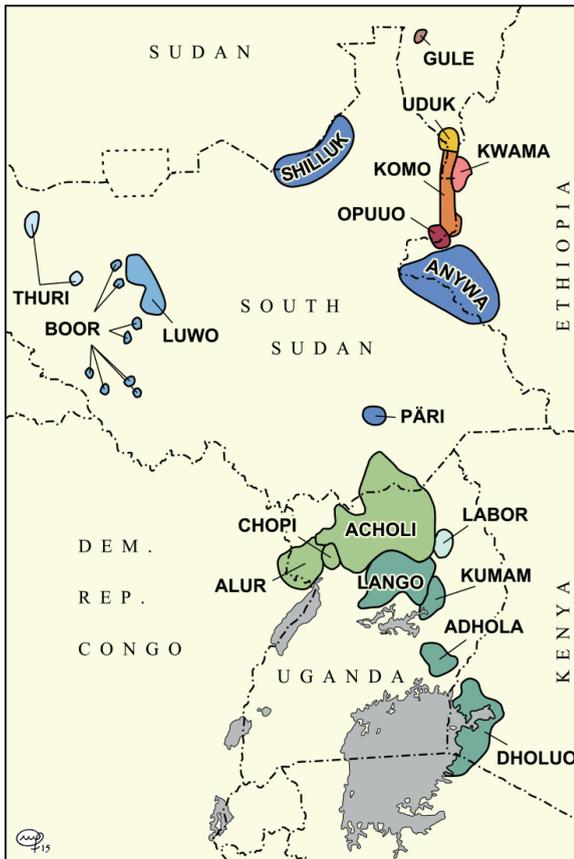
The Koman family consists of 5 languages (Uduk, Kwama, Komo, Opuuo, and Gule), spoken near the border of South Sudan and Ethiopia. The Koman family has been both included (Greenberg 1963; Bender 2000; Ehret 2001; Blench 2012) and excluded (Dimmendaal 2008) from the Nilo-Saharan phylum by various scholars due to the lack of regular correspondences and limited data. When it has been included in the phylum, it has most often been described as an early offshoot from the other Nilo-Saharan languages (Greenberg 1963; Ehret 2001; Blench 2013). The relationship between the Koman languages and the Luo languages is thus distant at a minimum. The Koman languages are, however, spoken in an area that is adjacent to the areas in which the Western Nilotic languages are spoken. Map

1 shows the areas inhabited by the speakers of each of the Western Nilotic languages with the area inhabited by the speakers of the Koman languages indicated.

Figure 1: The classification of the Luo languages with reference to the Nilotic family.



Map 1: The Western Nilotic languages (Storch 2005), with the area of the Koman languages.



3 KI Constructions

Within both the Luo and Koman families, there are numerous constructions that include a preposition *kV* (most commonly /ki/, /ka/, /gi/, /gi/, or /ke/), which in some languages reduces to a cliticized form *k=* before constituents that begin with a vowel segment. The three constructions that will be discussed in this paper include:

- Basic KI Construction
- Coordinate NP KI Construction
- Verbal complementation KI Construction

The Basic KI Construction is the most pervasive of these constructions and most often appears post-verbally. It contains a noun phrase that codes numerous semantic roles and functions, including Instrument, Manner, Time, and the Theme of antipassive constructions. This construction is represented using the structural schema in Figure 2.

Figure 2: The Basic KI Construction.



Although the Basic KI Construction is present in all of the languages described in this study, some of the semantic properties mentioned above are restricted to certain language groups. In the examples below, the various functions of the Basic KI Construction are discussed with reference to both the Luo and Koman families. The labels S, A, O, and KI are used for the single participant of an intransitive clause (including antipassive constructions), the most active argument of transitive clauses, the least active argument of transitive clauses, and the preposition taking the form *kV*, respectively (Dixon 1994; Comrie 1981). R and T refer to the Recipient and Theme of ditransitive clauses. It should be noted that these labels are not intended to reflect crosslinguistic categories, as argued by Dryer (1997).

Instrument¹

In these examples, the Basic KI Construction is used in both intransitive and transitive clauses to code the Instrument with which an action is completed.

Anywa (Northern Luo)

- | | | | | | |
|-----|---|----------|--------------|-----|-----------------|
| | A | O | V | | [KI NP] |
| (1) | ōbúrré | riŋó | ā-cám-gí | kī | lwēt-gí |
| | children | meat | PST-eat-3.PL | OBL | finger:PL:mN-3P |
| | ‘The children ate the meat with their fingers.’ (Reh 1996: 320) | | | | |

Päri (Northern Luo)

- | | | | | | |
|-----|---|----------|----------|------|----------------|
| | O | V | A | | [KI NP] |
| (2) | thiir | á-tiic | cícò-ê | kí | pál`-à [sic] |
| | spear.shaft | C-make | man-ERG | PREP | knife |
| | ‘The man made the spear shaft with a knife.’ (Andersen 1988: 303) | | | | |

Dholuo (Southern Luo)

- | | | | | | |
|-----|--|----------|----------|------|----------------|
| | A | V | O | | [KI NP] |
| (3) | Oloo | kunyo | bur | gi | beti |
| | Oloo | dig-IMP | hole | with | machete |
| | ‘Oloo is digging a hole with a machete.’ (Okombo 1997: 52) | | | | |

Acholi (Southern Luo)

- | | | | | | |
|-----|--|------|-------|--|----------------|
| | V | | | | [KI NP] |
| (4) | aŋolo | kí | pala | | |
| | I.cut | with | knife | | |
| | ‘I cut with the knife.’ (Malandra 1955: 151) | | | | |

¹ Some glosses have been slightly altered for ease of comprehension. A full list of abbreviations is included at the end. Original glosses for KI prepositions is retained.

Kwama (Koman)

- S V [KI NP]
 (5) mà jě kwèy jáf-mì gi gáájǎ mà-bábá
 1.PL.EX.SBJ also hoe eat-1.PL.EX.SBJ.PFV with hoe PL-father
 ‘We also hoe in order to eat [lit. we also hoe and eat] with a hoe (like our) fathers.’ (Kievit & Robertson 2011: 39)

Uduk (Koman)

- []_A V O [KI [NP]]
 (6) tādā-m pīnī thíp^h é kī tō pābā-m pīnī
 mother-LNK POSS.2.SG care.for 2.SG KI thing father-LNK POS.2.SG
 ‘Your mother rears you with your father’s things.’ (Killian, p.c.)

Komo (Koman)

- V-s [KI NP]
 (7) kīs’-í-r-í gi wat’i
 light-ITV-3.SG.M-3.SG.N INST fire
 ‘He lit it with fire.’ (Otero, p.c.)

Opuuo (Koman)

- A V O [KI-NP]
 (8) uteni kaay- yep-a buwaa-chaá go-wuy
 he PERV- cut-PST PL-tree INSTR-axe
 ‘He cut down trees with an axe.’ (Kebebw 2010: 20)

Manner

In these examples, the Basic KI Construction is used in intransitive clauses to code Manner with which an action is completed. This function appears to be infrequent but has been attested outside of Luo and Koman (see Section V below).

Anywa (Northern Luo)

- s-V [KI NP]
 (9) ēn-ā-rīijó kī ŋēεc
 3.SG-PST-run OBL speed
 ‘S/he ran fast.’ (Reh 1996: 320)

Dholuo (Southern Luo)

- S V [KI NP]
 (10) adhi sikul gi tienda
 1SG-go-IMP school by foot
 ‘I go to school on foot.’ (Okombo 1997: 46)

Time

In these examples, the Basic KI Construction is used to code the Time of an event, most often using words that refer to times of the day (e.g. morning, night). In Anywa, the Basic KI Construction can appear pre-verbally when coding Time (see 12).

Anywa (Northern Luo)

- s-V [KI NP]
 (11) gī-cámó kī wáár
 3.PL-eat:ANTI OBL night
 ‘They eat at night.’ (Reh 1996: 320)

Anywa (Northern Luo)

- [KI NP] V-s
 (12) kī wáʌr réε ā-kwál-é
 KI night RFL:3.SG PST-steal-3.SG
 ‘He stole away at night.’ (lit. ‘He stole himself at night.’) (Reh 1996: 502)

Päri (Northern Luo)

- S V [KI NP]
 (13) lwáak bùt `ò kí wáʌr `
 people sleep-SUF PREP night
 ‘People sleep at night.’ (Andersen 1988: 304)

Shilluk (Northern Luo)

- S V [KI [NP]]
 (14) gé á-bēéd kī cyʌŋ mō gīr
 3.PL PST:E-stay:1.SG KI day REL:PL many
 ‘They remained many days (like this)...’ (Miller & Gilley 2001: 61)

Kwama (Koman)

- V-s [KI NP]
 (15) só-ní-gé gí sùgùn
 spear-1.SG.SBJ.PERVV-3.M.SG.OBJ at night
 ‘I spear it (the pig) at night.’ (Kievit & Robertson 2011: 39)

Comitative

In these examples, the Basic KI Construction is used in both intransitive and transitive clauses to code an animate or inanimate Accompaniment. This may be an Accompaniment of the S, A, or O argument.

Anywa (Northern Luo)

- O V-A [KI NP]
 (16) jáath ā-kág-wá kí cókwā
 tree PST-split- OBL brother:our
 1.PL.EX
 ‘My brother and I split the tree.’ (Reh 1996: 320)

Päri (Northern Luo)

- []o V-s [KI [NP]]
 (17) còow árió á-nēen-á kí dháagò áciéíó
 men two C-see-1.SG PREP woman one
 ‘I saw two men and one woman.’ (Andersen 1988: 303)

Anywa (Northern Luo)

- V O [KI-NP] [KI NP]
 (18) òò kith cèŋŋ-é k-óogùudi [kī tīet-ē]
 and put hands-3.SG KI-bracelets KI feet-3.SG
 ‘(and he) put bracelets on his hands and feet.’ (Reh 1996: 525)

Labwor (Southern Luo)

- S V [KI NP]
 (19) én tíyó kí món
 3.SG work.IMV PREP women
 ‘He works with the women.’ (Heine & König 2010: 76)

Alur (Southern Luo)

- S V [KI [NP]]
 (20) ledit obino kō ter aryō
 person.old 3sg.come with load two
 ‘The elder came with two luggages.’ (Ringe 1949: 37)

Kumam (Southern Luo)

- s-V O [KI NP]
 (21) a=né!nó ióó a óóó tye i jé!rá mwá!kácá kede mótoka
 1SG=PERV:see man REL PST be in jail last year with car
 ‘I saw the man with a car who was in jail last year.’ (Hieda 2011: 80)

Theme of ditransitive

Notably, the Basic KI Construction does not mark the Recipient in ditransitive constructions (or the Goal of movement constructions) in any of the Luo languages or the Koman languages. It is occasionally used to mark the Theme in ditransitive constructions, as seen in (22).² In many of the Luo languages, a separate preposition is used for the Recipient in ditransitive constructions (see 23).

Luwo (Northern Luo)

- R V A [KI T]
 (22) nógé à-mwòj òm-é ké riŋ-ò
 people PERV-give.TR Dimo-ERG OBL meat.SGV
 ‘Dimo gave a piece of meat to the people...’ (Storch 2010: 6)

Anywa (Northern Luo) (NOTE: No KI Construction!)³

- V-IMP T R
 (23) cí óo cipí thír ji wàlárú néé jiedh-é
 go:IMP.SG and give:IMP.SG stick to father:MN:2.PL PURP sharpen-3.SG
 ‘Go and give the stick to your father to sharpen it!’ (Reh 1996: 526)

Theme of antipassive

In these examples, the Basic KI Construction is used with verbs in an antipassive form to code the Theme. This function appears to be restricted to Northern Luo, although some evidence suggests that it is possible in Koman languages (Killian, p.c.).

Anywa (Northern Luo)

- s-V:ANTI [KI NP]
 (24) ēn-ā-còobó kī dīel
 3.SG-PST- OBL goat
 spear:ANTI
 ‘S/he was spearing a goat.’ (Reh 1996: 320)

Päri (Northern Luo)

- s-V:ANTI [KI NP]
 (25) á-póod -ò kī nīkòth
 1.SG-beat:M:ANTI-SUF PREP nyikoth
 ‘I am beating Nyikoth.’ (Andersen 1988: 303)

² Luwo and Shilluk (Northern Luo) have been reported to show some use of the Basic KI Construction to code Location (Storch 2010: 14; Miller & Gilley 2001: 43).

³ The preposition *ji* is used in Anywa to code the Recipient of ditransitive clauses.

Shilluk (Northern Luo)

- S V:_{ANTI} [KI NP]
 (26) jāl-ání á-yēt kī wúnō
 man-REF PST:E-TWIST:ANTI KI rope
 ‘The man twisted some rope.’ (Miller & Gilley 2001: 43)

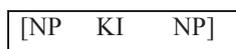
Luwo (Northern Luo)

- V:_{ANTI} [KI NP]
 (27) ø-à-tēdō ké kádō
 3.SG-PERV-cook:ANTI OBL broth.SG
 ‘S/he cooked some broth.’ (Storch 2010: 9)

The Coordinate NP KI Construction

The Coordinate NP KI Construction is used for the coordination of two NPs, including the coordination of numerals for some counting systems. Instances of this construction can also be nested within each other to create a long string of coordinate constituents (see 28). The structure of this construction is represented by the schema in Figure 3. Below are examples of the Coordinate NP KI Construction in the Luo and Koman families.

Figure 3: The Coordinate NP KI Construction.



Anywa (Northern Luo)

- [NP KI NP KI NP]-s V
 (28) mōō kī rídō kī thómōó di-jáábō nāacíel
 flour KI wort KI yeast DEO-mix.PASS together
 ‘The flour, wort, and yeast are mixed together.’ (Reh 1996: 547)

Anywa (Northern Luo)

- [kiNP]
 (29) gēn-ārōŋ k-ōobúurè
 3.PL-PST-dive OBL-children
 ‘He and the child dived.’ (Reh 1996: 536)

Belanda Boor (Northern Luo)

- [NP KI NP]
 (30) yomō kī gifī mu ndingili
 wind KI thing REL round
 ‘tornado’ (von Heyking 2010: 51)

Dholuo (Southern Luo)

- [NP KI NP]_A V O
 (31) otieno gi wuon gero ot
 Otieno and father:POSS build house
 ‘Otieno and his father are building a house.’ (Okombo 1997: 64)

Alur (Southern Luo)

- V-IMP [NP KI NP]-o
 (32) cidh ilwoŋ Bonifasiyo ku Leonardo
 go.IMP and.call B. and L.
 ‘Go call Boniface and Leonard.’ (Ringe 1949: 38)

Päri (Northern Luo)

- A V_{-ANTI} [KI NP] [KI V]
 (39) ùbúr á-póop̃ -ò kí lwáak kí gòor
 ubur C-teach:M:ANTI-SUF PREP people PREP writing
 ‘Ubur taught the people how to write.’ (Andersen 1988: 304)

Luwo (Northern Luo)

- S V [KI V]
 (40) riijen caa ke⁵ ròbo ir nyethenmen
 men resume and talk matter girls
 ‘...(the) men resume (to) talk to the girls.’ (Santandrea 1977: 563)

Shilluk (Northern Luo)

- S V [KI V]
 (41) é bàn kí tèr chanduk
 3.PL refuse KI carry box
 ‘They refused to carry the box.’ (Westermann 1912: 38)

Komo (Koman)

- S V [KI V]
 (42) à=dàw ja-r komà kí pó gı kót
 SG=baboon go-3.SG.M to in garden PURP till
 ‘The baboon went (in) to the garden to till (the ground).’ (Otero, p.c.)

Opuuo (Koman)

- S V [KI V]
 (43) utəni ka:jkaraṭʰe ga nata
 he PERV- agree PART work
 ‘He agreed to work.’ (Kebebw 2010: 80)

Uduk (Koman)

- S AUX V [KI [V]]
 (44) wàk^hkí é miní ó kí c^hāb ē ipābā-m pīnī
 COND 2.SG AUX say KI stay eye PL-father-LNK POSS.2SG
 ‘If you wish to stay with your father...’ (Killian, p.c.)

4 Summary of results

A summary of the constructions detailed in the examples above is provided in Table 1 below. The three right-most columns represent the distribution of the three KI Constructions, respectively. A plus sign (+) indicates that the language uses a particular construction and a minus sign (-) indicates that it does not. A question mark (?) is used where data was insufficient enough to determine the occurrence of a construction or the validity of the data was questionable. If an attested construction exists and is indicated by a plus sign, the form of the cognate KI preposition is provided in parenthesis. If no such construction exists in a given language, then the alternate non-cognate form of a corresponding construction in the language is provided (e.g. Alur *m*).

Nearly all of the languages surveyed in this study, including both the Luo languages and the Koman languages, feature both the Basic KI Construction and the Coordinate NP Construction. Within Northern Luo, both of these constructions appear to use the same cognate KI form, either /ki/, /kɪ/ or /ke/. Southern Luo exhibits a higher amount of variation from language to language and from construction to construction. The use of a voiced velar /gɪ/ in Dholuo is not entirely unexpected, as

⁵ Tone is not marked in Santandrea (1977), so it is assumed here that it is the same as the high tone *ké* described in Storch (2010).

Dholuo is spoken in Kenya in an area that is geographically and politically separated from the other Southern Luo languages in Uganda. The Koman languages also exhibit a higher degree of variability, and the reliability of the data is not equal to that of the Luo languages. Even so, both the Basic KI Construction and the Coordinate NP KI Construction appear to be well represented in all of the Koman languages.

The distribution of the Verbal Complementation KI Construction is limited to the Northern Luo and Koman languages. Notably, the alternate verbal complementation construction that is attested in the Southern Luo languages includes a complementizer of the form /ni/ or /ni/. A cognate complementizer (/ni/, /ni/, /nee/, etc.) is employed by the Northern Luo languages in a more general verbal complementation construction that is less lexically-restricted than the Verbal Complementation KI Construction.

Table 1: KI Constructions in Northern and Southern Luo

family / sub-family	language	Basic KI Construction [KI NP]	Coordinate NP KI Construction [NP KI NP]	Verbal Complementation KI Construction [KI (NOM-)V]
Northern Luo	Shilluk	+(kī)	?	+(kī)
	(Jur) Luwo	+(ké)	+(ké)	+(ké)
	Anywa	+(kī)	+(kī)	+(kī)
	Päri	+(kī)	+(kī)	?
Southern Luo	Dholuo	+(gī) ⁶	+(gī)	-
	Alur	+(kū)	+(kū)	(ni)
	Lango	+(kèdè~kè)	+(kèdè~kè)	(nī)
	Acholi	+(kī) ⁷	+(kī)	(ni)
	Labwor	+(kī) ⁸	+(gín kī)	(ni)
	Kumam	+(kede)	?	(bé)
Koman	Uduk	+(kī, kī)	+(kī) ⁹	+(kī)
	Kwama	+(gì)	+(gì, gù)	-
	Komo	+(gī)	+(gì)	+(gī)
	Oquo	?(go)	?(gu)	?ga

5 Discussion

The development of the Basic KI Construction and the Coordinate NP KI Construction

Based on the widespread distribution within the sampled languages and the general lack of lexical restrictions, both the Basic KI Construction and the Coordinate NP KI Construction appear to be of relatively old origin. Nearly every language in the survey has attested instances of both constructions, which suggests that they can be reconstructed to Proto-Luo and Proto-Koman. The reconstructions of these synchronic KI Constructions are labeled here as the *KI Basic Construction and the *KI Coordinate NP Construction. Their structures are identical to those of the synchronic constructions, as shown by the schemas in Figure 5 and Figure 6.

Figure 5: The reconstructed *ki Basic Construction in Proto-Luo and Proto-Koman.

☆ [*ki NP]

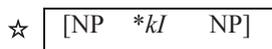
⁶ Dholuo uses *kod* for animate NPs.

⁷ Acholi uses *ki* for inanimate NPs and *ka* for animate NPs (Malandra 1955: 131).

⁸ Labwor uses *ki* for full NPs and *kod* for pronouns.

⁹ Reported by Don Killian (p.c.), but use is limited.

Figure 6: The reconstructed *KI Coordinate NP Construction in Proto-Luo and Proto-Koman.



The fact that both of the reconstructed constructions can be posited for proto-stages indicates that their development must have occurred at an even earlier time. Although there is no relevant attested data from such an early time period, diachronic typology can provide some clues as to the historical development of the Basic KI Construction and the Coordinate NP KI Construction. Stassen (2000) examined a sample of 260 languages and found a critical link between two strategies of noun-phrase conjunction: the Comitative Strategy (e.g. the Basic KI Construction), and the Coordination Strategy (the Coordinate NP KI Construction). He found that, "...with only a few exceptions, all languages in the sample appear to have the possibility of employing the Comitative Strategy." (ibid, 21).

Stassen additionally observed that, "[languages that only have a Comitative Strategy] do not have a Coordinate Strategy, but they would like to have one." (ibid, 26). He found that languages with the Comitative Strategy tended to change features or adopt new features in the direction of the Coordinate Strategy, but never the other way around. These changes often result in a two-strategy system that combines both the Comitative and Coordinate Strategies, as exemplified by the Luo and Koman languages. This is further supported by Heine & Kuteva's (2002: 80) numerous examples of the pathway Comitative -> NP-AND (Noun Phrase Coordination), and also Creissels' (2005: 52) observation that "...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 'comitative strategy'".

The semantic reanalysis of the Comitative function of the Basic KI Construction as coordination of NPs most likely began in either post-verbal O argument coordination (as in 17), post-verbal A argument coordination (similar to 16), or the S argument of the passive (as in 28), because these are contexts where the Comitative Strategy and the Coordination Strategy are structurally identical. Additional evidence from the grammaticalization literature suggests that the functional diversification of the Basic KI Construction traces back to the Comitative (Heine & Kuteva 2002).

The development of the Verbal Complementation KI Construction

The Verbal Complementation KI Construction most likely represents a more recent development than the Basic KI Construction and the Coordinate NP KI Construction, because within the Luo family it can only be reconstructed for Proto-Northern Luo and it exhibits significant lexical restrictions in both the Northern Luo and Koman families. Due to similarity in form and function, it is most likely that the Verbal Complementation KI Construction developed out of the Basic KI Construction through semantic reanalysis involving a small set of lexical verbs.

Within Northern Luo, this reanalysis appears to have occurred in two stages in connection with the coding of the Theme of Antipassive. The antipassive is a common feature of Western Nilotic languages in general (Schröder 2006), but the Northern Luo languages appear to be unique in having some non-derived active-voice verbs that take "demoted" O arguments marked with the KI preposition. Many of these are the same verbs that operate within the Verbal Complementation KI Construction, as seen in the following two examples:

Luo (Northern Luo)

S V [KI NP]

(45) nyakou ayiio ke cwoꞤ

'...the girl agrees on/to her husband.' (when choosing a future husband) (Santandrea 1977: 563)

Anywa (Northern Luo)

S V [KI NP]

(46) bá ēnī kwéeyyó kī ców

but she refuse:FQ OBL men

'But she refused men.' (Reh 1996: 333)

Thus these verbs appear to resemble the Theme of Antipassive Basic KI Construction, but the verbs are not in the antipassive form. Given that the antipassive verbal forms are coded through vowel alternations on the verbal stem and are very likely to have their reconstructed origin in Proto-Western Nilotic, it follows that the examples above reflect a reanalysis of the Theme of Antipassive Basic KI Construction as consisting of a verb (active or antipassive) with a [KI NP] complement. This level of reanalysis is not restricted to PCU and aspectual verbs (see 47).

Anywa (Northern Luo)

- V-s [KI NP]
 (47) wár-í kī dúut
 sing-2SG OBL song
 ‘Sing a song!’ (Reh 1996: 394)

The second stage of reanalysis involves a change in the lexical class of the constituent following the KI preposition in the Basic KI Construction. The original construction is reanalyzed as coding a verbal complement, either through use of nominalization morphology or with a bare verb. There does not appear to be consistency in the pattern of which matrix verbs require nominalization morphology on the verbal complement and which do not.¹⁰ Some evidence points toward dual functionality of the nominalizer as coding both verbal complements and relative clauses (see 48).

Anywa (Northern Luo)

- V-s [KI NOM-V]
 (48) wár-í kī màn-ā-máɲɲ-í
 sing-2SG OBL NMZ-PST-want-2SG
 ‘Sing what you want!’ (Reh 1996: 394)

Parallels in Southern Luo indicate that cognate PCU and aspectual verbs are either irregular, take verbal complements without use of a complementizer, or are used paratactically. Noonan (1992) observes that the Lango verb *dáɡí* ‘refuse’ is irregular in that it does not take an object suffix. The word *caGo* ‘start’ in Lango and Acholi is cognate to Anywa *caGGo* and takes bare verbal complements without a complementizer (Noonan 1992; Crozzolera 1955). Acholi *tyeko* ‘finish’ is used paratactically, as in (49).

Acholi (Southern Luo)

- (49) gi-kayo bel gi-tyeko
 3:PL-harvest corn 3:PL-finish
 ‘They harvested all the corn.’ (Bavin 1989)

Following Givón (1979) and Dahl (2009), it is also possible that some of the Northern Luo verbs functioning within the Verbal Complementation KI Construction have developed out of paratactic constructions like the one exemplified above. This would reflect one of the primary pathways for grammatical evolution: paratactic constructions > syntactic constructions. More evidence is necessary to confirm such a possibility.

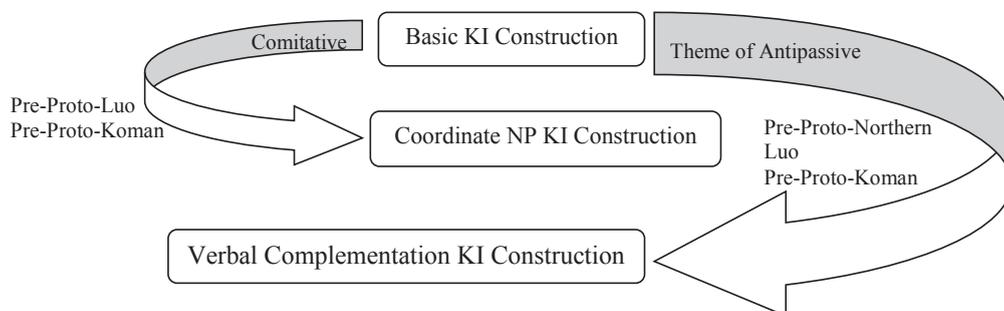
Within the Koman languages, evidence of the development of the Verbal Complementation KI Construction is fewer and farther between. To date, no antipassive verbal forms have been posited for any of the Koman languages, which precludes direct comparison with the developments in Northern Luo. The widespread use of the Verbal Complementation KI Construction with verbs of Utterance represents a stark contrast with the Northern Luo languages (which, in turn, solely use the alternate /ni/ complementizer for verbs of Utterance). Continuing research on the languages of the Koman family will hopefully shed more light on these developments in the future.

In summary, the diachronic development of the three KI Constructions consists of two separate pathways of reanalysis, stemming out of the Basic KI Construction. It is proposed that the Coordinate

¹⁰ The nominalizer in Anywa comes from */mā/, pl. */mō/, which was an alienable noun meaning ‘one’ or ‘thing’ (Reh 1996: 155).

NP KI Construction first developed out of a reanalysis of the Comitative Basic KI Construction, and that the Verbal Complementation KI Construction then developed out of the Theme of Antipassive Basic KI Construction at a later stage. These developments are represented below in Figure 7.

Figure 7: The diachronic development of the KI Constructions.



Implications for language relationships

Given the distribution of the three KI Constructions in Table 1 and the stages of diachronic development in Figure 7 above, four possible explanations for the occurrence of cognate KI Constructions in both the Koman and Luo families are considered.

a. Chance

Ringe (1992) has demonstrated that non-cognate monosyllabic CV morphemes from different languages can easily take the same form by sheer coincidence, and Cambell (2004) has argued that comparison of such forms for the purposes of diachronic linguistics should be avoided. This argument does not apply to the KI forms in the Luo and Koman languages for three reasons: First, the scope of the correspondence is not between just two languages, but rather is a multilateral comparison across numerous languages. Second, the cognate morphemes consist of both an initial consonant as well as a specified vowel (/i/, /e/, or /ɛ/), rather than merely an initial consonant as in Ringe’s multilateral study. Third, the comparison here is between schematic constructions, and not merely individual morphemes. The KI Constructions specify non-segmental information about other atomic units in addition to the cognate KI morpheme (e.g. lexical class, constituent order, etc.), which also reduces the possibility of chance similarity.

b. Inheritance

It is possible that the distribution of shared constructions is primarily due to shared inheritance. The primary issue with this explanation is that the Verbal Complementation KI Construction occurs in only the Koman and Northern Luo families. Based on previous comparative work, it is clear that Northern Luo and Southern Luo are closely-related sub-families and that Koman is more distantly related to both Luo sub-families than the two sub-families are to each other. Thus, for inheritance to be the primary source of these shared constructions, it would be necessary to reconstruct a *KI Verbal Complementation Construction at the Pre-Proto-Luo and Pre-Proto-Koman stage. The absence of the Verbal Complementation KI Construction in the Southern Luo languages would have to be explained as a sub-family-wide loss. This scenario is unlikely for two reasons: First, the restriction of the Verbal Complementation KI Construction to aspectual and PCU verbs in Northern Luo and Koman suggests that it is a recent innovation involving a small subset of verbs and not the remnants of a more widespread phenomenon. Second, there already exists a verbal complementation construction that reconstructs to Proto-Luo that includes the /ni/ complementizer does not appear to be lexically-restricted. Therefore the likelihood is low that the Verbal KI Complementizer was originally more widely used and was then subsequently replaced by the /ni/ constructions.

The distribution of the Basic KI Construction and the Coordinate NP KI Construction, however, clearly indicates that they reflect shared inheritance that reconstructs to Proto-Luo and Proto-Koman. Additionally, there are examples of *kV* constructions that appear to closely resemble the KI Constructions both in form and function. Thus the possibility of higher-level shared inheritance is a distinct possibility, but it is beyond the scope of this paper. A small selection of evidence found so far includes: **Gumuz** (*ká* and *ka*), **Western Nilotic** Nuer, Mayak, and possibly Mabaan, **Eastern Nilotic** Teso, Turkana, Bari, Karimojong, and Lotuko, **Southwest Surmic** Tenneset and Didinga, **Central Sudanic** Yulu, Kresh, and Baka, among others.¹¹

c. Contact-induced change

The distribution of the Verbal Complementation KI Construction can most easily be explained in terms of contact between the speakers of the Northern Luo languages and the Koman languages. Northern Luo is spoken in an area significantly closer to the Koman family than Southern Luo, and Northern Luo exhibits a pattern of KI Constructions that more closely resembles that of Koman. This indicates that, at a minimum the distribution of the Verbal Complementation KI Construction is most likely due to contact-induced change. Significant direct contact between the speakers of these families has not been reported, however, and the possibility of contact-induced phenomena involving the Burun or Dinka-Nuer languages cannot be ruled out. Additional historical and archaeological data is necessary to assess these possibilities. In contrast, the Basic KI Construction and Coordinate NP KI Construction are not best explained by contact, as discussed above. This supports a more nuanced fourth possibility: that the distribution of all of the KI Constructions can be explained by *(d) a combination of inheritance and contact-induced change*.

5 Concluding remarks

Although the Koman and Luo languages are known to be distantly-related, they feature a number of shared grammatical constructions that cannot easily be merely due to coincidence. This paper has shown that three of these constructions, the Basic KI Construction, the Coordinate NP KI Construction, and the Verbal Complementation KI Construction are cognate and can be reconstructed to proto-stages. The co-occurrence of these constructions in the Luo and Koman families is best explained by positing historical stages of both inheritance and contact. It is in this way that the diachronic development of each of the three KI Constructions reflects a unique temporal and interactional backdrop to syntactic change. Further research into the dispersion of cognate KI Constructions outside the Luo and Koman families will shed light on the extent to which these constructions reflect higher-level shared inheritance, and hopefully provide additional insight into the genetic and areal relationships of the Nilo-Saharan languages.

¹¹ Notably, there is not evidence of cognate KI Constructions in Southern Nilotic.

Abbreviations

1	1 st person	NOM	Nominalizer
2	2 nd person	OBJ	Object
3	3 rd person	OBL	Oblique
ANTI	Antipassive	PART	Particle (Kebebw 2010)
C	Dependent Clause Marker (Reh 1996)	PASS	Passive
COND	Conditional	PERV	Perfective
DEO	Deontic	PL	Plural
E	Evidential (Miller & Gilley 2001)	POSS	Possessive
ERG	Ergative	PREP	Preposition
ERG	Ergative	PST	Past
EX	Exclusive	PURP	Purposive
IMP	Imperative	REF	Referential Determiner
IMV	Imperfective	REL	Relative Pronoun
INSTR	Instrument	SBJ	Subject
ITV	Itive / Translocative	SGV	Singulative
LNK	Linker	SUF	Verbal Suffix (Andersen 1988)
M	Multiplicative	TR	Transitive
MN	Modified Noun (Reh 1996)	VEN	Ventive / Cislocative

References

- Andersen, Torben. 1988. Ergativity in Pari, a Nilotic OVS language. *Lingua* 75: 289–324.
- Bavin, Edith L. 1989. The Ki-verb Constructions in Acholi'. *La Trobe working papers in linguistics*, pp. 29–41. Bundoora, Victoria: Linguistics Program, La Trobe University.
- Comrie, Bernard. 1981. *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: University of Chicago.
- Creissels, Denis. 2005. *A Typology of Subject Marker and Object Marker Systems in African Languages*. Amsterdam/Philadelphia: Benjamins.
- Dahl, Östen. 2009. Two Pathways of Grammatical Evolution. In T. Givón & Masayoshi Shibatani (eds.), *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution*, pp. 239–248. Amsterdam/ Philadelphia: Benjamins.
- Dixon, Robert M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Dryer, Matthew. 1997. Are Grammatical Relations Universal? In Joan Bybee, John Haiman & Sandra Thompson (eds.), *Essays on Language Function and Language Type: Dedicated to T. Givón*, pp. 115–143. Amsterdam/Philadelphia: Benjamins.
- Greenberg, Joseph H. 1963. *The Languages of Africa*. Bloomington: Indiana University.
- Givón, Talmy. 1979. From Discourse to Syntax: Grammar as a Processing Strategy. In T. Givón (ed.), *Syntax and Semantics 12: Discourse and Syntax*, pp. 81–112. New York, NY: Academic Press.
- Givón, Talmy. 1984. *Syntax: A Functional-Typological Introduction 1*. Amsterdam/Philadelphia: Benjamins.
- Heine, Bernd & Christa König. 2010. *The Labwor Language of Northeastern Uganda: A Grammatical Sketch* (Studies in Nilotic Linguistics, 1). Tokyo: Research Institute for Languages and Cultures of Asia and Africa.
- Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

- von Heyking, Beatrix. 2010. The Noun Morphology of Belanda Boor. *Studies in Nilotic Linguistics*, 2: 29–60.
- Hieda, Osamu. 2011. *Kumam Vocabulary with Grammatical Notes (Studies in Nilotic linguistics, 4)*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.
- Kebebw, Lemi. 2010. A Grammatical Description of Opo. Master's Thesis, University of Addis Ababa.
- Kievit, Dirk & Erika Robertson. 2011. Notes on Gwama Grammar with Texts and Word Lists. SIL Ethiopia.
- Malandra, Alfred. 1955. *A New Acholi Grammar*. Kampala: Eagle Press.
- Miller, Cynthia L. & Leoma Gilley G. 2001. Evidence for Ergativity in Shilluk. *Journal of African Languages and Linguistics* 22.1: 33–68.
- Okombo, Duncan. O. 1997. *A Functional Grammar of Dholuo* (Nilo-Saharan, 12). Cologne: Köppe.
- Reh, Mechthild. 1996. *Anywa Language: Description and Internal Reconstructions* (Nilo-Saharan, 11). Cologne: Köppe.
- Ringe, Donald A. 1992. On Calculating the Factor of Chance in Language Comparison. *American Philosophical Society, Transactions* 82.1: 1–110.
- Ringe, Peter. C. 1949. *A Simple Alur Grammar and Alur – English – Alur Vocabularies*. Kampala: Eagle Press.
- Santandrea, Stefano. 1977. Jur-Luo Texts and Comments: The Family. *Anthropos* 72: 557–609.
- Schröder, Helga. 2006. Antipassive and Ergativity in Western Nilotic and Surmic. *Annual Publications in African Linguistics* 4: 91–110.
- Stassen, Leon. 2000. AND-Languages and WITH-Languages. *Linguistic Typology* 4: 1–54.
- Storch, Anne. 2005. *The Noun Morphology of Western Nilotic* (Nilo-Saharan, 20). Cologne: Köppe.
- Storch, Anne. 2010. Luwo Morphosyntax. *Studies in Nilotic Linguistics* 2: 1–27.
- Westermann, Diedrich. 1912. *The Shilluk People, their Language and Folklore*. Berlin: Reimer.

LINGUISTIC ETIQUETTE IN A FRONTIER SITUATION: A CASE STUDY FROM CHOPI¹

Anne Storch

1 Introduction

In all languages it is possible to express respect or disrespect towards an addressee or other speech participants. Much of what speakers actually say when they want to convey information on requests, wishes, personal points of view or their own interest needs to be coded in a respectful or polite way, and most of what people then say is said in a rather indirect manner. In other words, we often find it more appropriate to say something like “it would be nice to have some coffee right now”, or “may I have a cup of coffee” instead of “bring me coffee”, and we probably rather say things like “interesting shirt!” or “you look so different today” instead of “I think your dress is ugly”. This has to do with the creation or maintenance of social borders, hierarchies, and so on. In politeness theory, indirect communication is often understood as an act of face-keeping which gives a lot of agency to the addressee: at least, it is the addressee who can still decide to interpret indirect communication in various ways (Brown & Levinson 1987).

Ritual communication, such as greetings and other access rituals are other means of creating or maintaining social borders. They usually differ according to social roles: Sometimes we greet people by handshake and sometimes we just nod, and this may refer to differences in our relationship to a person, status, social hierarchy, formality of a situation, and so on.

The construction of honorific expressions, the use of indirect communication in the language of politeness, be it through metaphorical language or non-verbal communication (such as in hosting somebody, cf. Stoller & Olkes 2005, or adornment and dress), and the creation of manipulated languages all occur in specific forms and refer to different, specific social and cultural contexts relevant to them in the different languages. This contribution sets out to explore whether we can identify any socio-cultural context that nurtures the creation and use of particular forms of polite and honorific language and whether this helps us to understand why a particular language is associated with particular ways of speaking, in contrast to other languages that may form part of its speakers’ repertoires. The language we look at in some detail here is called Chopi, a Southern Lwoo language (Nilotic) spoken in northwest Uganda. We will see further below what kind of language this is.

Politeness and honorific constructions are understood in this contribution as salient expressions of language ideologies. They provide insight into how social praxis shapes the ways in which speakers use their language and the ways in which linguistic forms develop (or, are developed).

Politeness and/or honorifics as an important part of communication have been studied in some African languages such as Wolof (Irvine 1974), ChiBemba (Irvine 1998, Irvine & Gal 2000), Nyakyusa (Kolbusa 2000), Kambaata (Treis 2005), and a few others, but rarely in Nilotic languages. Of course, not all languages have conventionalised linguistic forms that we could call honorifics – deictic forms that express social difference, such as a specialised set of pronouns or a dedicated use of the plural. In Western Nilotic, to which Chopi belongs, such forms seem to be missing. However, language ideologies and ways of speaking in these languages haven’t been much explored yet so that a different picture may emerge once more documentary work has been done. At least, a number of interesting observations have been made in Western Nilotic languages. Evans-Pritchard (1948) provides data on the use of personal names in Nuer, addressing taboo and honorifics in a complex system of naming. Reh (1996) and Heusing (2004) have demonstrated how speech registers in Anywa and Southern Lwoo, respectively, refer to social hierarchy and to its provocative negation in youth culture.

¹ I’m grateful to Norbert Kirya Oryem and Betty Nyama Abalo for their kind assistance and patience when working with me in Kampala, and to Angi Mietzner and Friederike Lüpke for their many stimulating and helpful comments on earlier versions of this text.

Furthermore, Storch (2011a) explores the variety of expressions of linguistic ideology and creativity in language in a number of Western Nilotic languages.

2 Forms and constructions

2.1 Typological outline

Chopi² shares with its Southern Lwoo relatives and neighbours most of the (so-called) basic vocabulary, which reaches up to 82 per cent of common retentions (Blount & Curley 1970). The differences in the phonology and morphosyntax, however, are obvious (Heusing 2004). Similar to Adhola, Alur, Kumam, and Dholuo, Chopi has nine vowels, which fall into two sets, namely [+ATR] vowels /i, e, a, o, u/ and [-ATR] vowels /ɪ, ɛ, ʊ, ɔ/. The consonant inventory of Chopi is an innovative one, which exhibits four fricatives, being rather unusual for most Northern Lwoo and other Western Nilotic languages. Like in Acholi, Lango and Kumam, the original dental consonants have merged with alveolar consonants. There are no nasal-stop clusters with the exception of loanwords from Bantu. The consonant inventory can be summarised as /p, t, c, k, ʔ, b, d, j, g, f, s, sh, x, m, n, ŋ, r, l, w, y/. Note that the fricatives tend to appear in intervocalic position, e.g. *òfùk* ‘tortoise’, *àsham* ‘leftside’, *wàxó* ‘outside’. As far as the presently analysed data suggests, Chopi distinguishes three tones, namely high [á], mid [a], low [à]. Syntactically, Chopi is AVO/SV. The noun morphology is characterised by prefixed number-marking morphemes which are historically derivative morphemes. The language also exhibits remnants of prefixed classifiers. Verb inflection primarily operates by affixation, whereby the conjugated verb basically consists of a pronominal prefix, an affixed TAM-morpheme and in some forms an auxiliary.

2.2 Address terms and access rituals

Speakers use specific terms and constructions to mark the beginning and end of an encounter. Greetings, polite remarks and ritualised enquiries about well-being, work, family and so on are access rituals that help the speakers to enter into a type of communication that increases closeness.

Chopi speakers claim that the design of access rituals has another important function, namely to mark differences between communities on the one side and create openness for all participants on the other. Whereas neighbouring communities who speak Bantu languages such as RuNyoro use complex greeting formula in which honorifics are salient, Chopi speakers say that greetings in their community are simple so that everybody can integrate. There are a number of conventionalised forms which refer to a situation, mostly the time of the day, or the circumstances of the encounter:

- | | | | | |
|-----|------------------------------|--------|----------------|----------|
| (1) | A | | B | |
| | i-búr-ú | needi? | à-búr-ú | à-béè |
| | 2SG-sleep-PERF | INTER | 1SG-sleep-PERF | NOM-good |
| | ‘how did you sleep?’ | | ‘I slept well’ | |
| | | | | |
| (2) | i-rí-ó | needi? | | |
| | 2SG-spend_time-PERF | INTER | | |
| | ‘how did you spend the day?’ | | | |

The encounter may include a question after somebody’s well-being, greetings and good-bye:

- | | | | | |
|-----|-----------------|-------|---------------|-----|
| (3) | A | | B | |
| a. | wác-á | ŋɔɔ? | wác | pee |
| | problem-PL | INTER | problem | NEG |
| | ‘any problems?’ | | ‘no problems’ | |

² Up to now, no linguistic work on Chopi is available, apart from a very brief description of its noun morphology in Storch (2005). A description that covers the grammar and the lexicon of Chopi is currently prepared by the author.

- b. *cì* *móóg-í*
greet.IMP other-POSS.2SG
'greet the others'
- c. *dè* *beet à-béè* *ín* *béné* *béét* *à- béè*
wish stay NOM-good 2SG also stay.IMP NOM-good
'wish [you] stay well' 'you also stay well'

In principle, this is what opens and closes an act of communication, and most speakers claimed that this was easily learned by outsiders so that they could feel as part of the society.

In the dialogues and texts I was able to collect, greetings never turned out to be more complex as this. Greeting oneself into an encounter or a situation, getting into touch with each other and establishing a relation is never made difficult, for speakers and semi-speakers of Chopi alike. It seems as if Chopi is thought as a broker language in which everybody can freely interact, not making any mistake in choosing the wrong register. There is no utter pressure on speakers to avoid any mistakes, as shaming and teasing are not part of access rituals.

However, a different picture emerges once we take address terms into account. In order to maintain face in the community, one needs to be greeted with the correct address term, which implies that speakers actually need to know the members of their community in terms of their clan-membership and family background, something that outsiders would hardly be capable of.

A respectful way of greeting would involve uttering the respective clan's praise name, which is an epithet signifying group membership, association to a particular place and a specific history. Clan praise names are used in a similar way as personal names in greetings, but are not identical with them. Hence, while greeting a person is called by the clan praise name rather than by his or her given name, e.g.:

- | | |
|--|--|
| <p>(4) A
<i>ì-búr-ú</i> ín <i>needì</i> <i>àmóòti?</i>
2SG-sleep-PERF 2SG INTER [NAME]
'how did you sleep, oh Amoti?'</p> <p><i>ì-wór-ú</i> ín <i>needì</i> <i>àkiiki?</i>
2SG-arrive-PERF 2SG INTER [NAME]
'how did you arrive, oh Akiki?'</p> | <p>B
<i>à-búr-ú</i> <i>à- béè</i> <i>àdyèèrí</i>
1SG-sleep-PERF NOM-good [NAME]
'I slept well, Adyeri'</p> <p><i>à-wór-ú</i> <i>à-béè</i> <i>àbwóòì</i>
1SG-arrive-PERF NOM-good [NAME]
'I reached well, Abwoi'</p> |
|--|--|

What is especially interesting here is that A's greeting formula always includes the free form of the 2sg pronoun, which serves as an honorific marker. Honorifics based on deictics occur in those Lwoian languages that are spoken by groups who have or had sacred kings. In Shilluk, for example, the use of 3sg pronouns for EGO has been reported to be part of the royal register. But here, in Chopi, not self-elevating functions are covered, but obviously speakers emphasise the role of the addressee. Unlike honorifics in Bantu languages, where such strategies have to do with deferential style (Irvine 1998:60 f.), in Chopi no deferential meaning is implied.

Such greetings are usually only then exchanged when members with in-group knowledge meet, people who consider themselves "real Chopi" and who claim to be member of the clans of the old Chopi land:

- (5) *biètó* royal clan
jó píina 'mason wasp' clan
lú gáyá 'finger millet' clan
lú bup 'elephant' clan
bá fùmá 'buffalo' clan
bà sénjà 'short-feathered chicken' clan

Besides these, it is appropriate to use the praise names for those who have been displaced from their original land, or for groups who settle abroad, such as:

- | | | |
|-----|----------------|--|
| (6) | pàluò jó àbòrà | displaced from Murchison Falls National Park |
| | liira pàluò | Chopi from Lira |
| | pàluò bóxe | Chopi from Kibanda |
| | pàwii ~ kidôpò | Chopi from Mutunda on River Nile |

In case the relevant praise name is not known, a speaker is supposed to use clan-neutral address terms of respect. These are:

- | | | |
|-----|---------|--|
| (7) | múkáyí | ‘elder’ |
| | nà-ròót | ‘female address term’ (lit. ‘child of king’) |
| | nà-tòón | ‘male address term’ (lit. ‘child of bull’) |

All of these terms imply that the addressee is a person with whose clan or family specific understandings concerning the exchange of daughters and joking relations exist. Other people are greeted the way it has been sketched above in (1) – (3), unless they have been included in the community by marriage or adoption.

3 Language attitudes and ideologies

In a society that largely does without complex, time-consuming access rituals one may assume that strangers play a simple role. As already hinted above, Chopi speakers see a positive meaning in directness and the minimalised exchange of polite expressions, claiming that this keeps an encounter uncomplicated and makes their society open and easily accessible (in contrast to Bantu-speaking communities in the neighbourhood). This can be easily understood when taking the situation of the language and its community into account:

According to the 2002 census, there are about 20,000 “ethnic” Chopi, whereby not all of them may be speakers of the Chopi language. The Chopi of Masindi immediately neighbour the Banyoro, and most are at least bilingual in Chopi and in RuNyoro. This tendency has not only resulted in increasing language shift, but also in a relatively new concept of Chopi as a language and community being distinct from that of Acholi. As the *Ethnologue* entry (Lewis 2009) of “Acholi” still illustrates, Chopi is considered one of several Acholi dialects by many linguists, as well as by influential bodies such as SIL (who are responsible for the *Ethnologue*). As a consequence of such steady shifts and changes – we will learn more about them below – Chopi is spoken by a highly multilingual community, whereby repertoires are largely unstable. Speakers know Acholi and RuNyoro, likely English, and a number of local languages plus sometimes older vernacular languages such as Swahili and Juba Arabic.

Moreover, very few clans actually originate from a “Luo” population. Besides those mentioned in (5), there are more clan names. Like the ones listed above, most of them refer to groups that have been incorporated through marriage, or that have migrated into Chopi land and vice versa. There are some clans – *niúágí*, *pàbít*, *òtác*, *geyá*, *kòc* – that relate to Luo communities in Acholi land, but these seem to be a minority among the Chopi clans. The different Chopi communities have extremely diverse histories and consist of families with a large variety of origins. Coherence in the interethnic conglomerates is never permanently achieved, and individual speakers may identify themselves as Chopi, Banyoro, Acholi, and so on, depending on circumstance and context.

The minimal form in access rituals in this respect seems to be an ideal compromise, easy to learn and simple to use. In many aspects, Chopi looks like a language of the Frontier which is used to negotiate relations and social hierarchy in small, multidiverse communities. What is interesting in terms of linguistic change is that the speakers who reside in such multilingual settings, being exposed to larger high-prestige groups and their languages, on the one hand use grammatical strategies that resemble those of neighbouring Bantu languages, but on the other consciously foster typically “Nilotic” properties of their language. For example, Chopi exhibits a relatively large number of noun prefixes which is rather atypical for Western Nilotic languages but characteristic for Bantu, and it has largely restructured its number-marking system (e.g. by innovating a periphrastic singulative; e.g.

Storch 2011b), all of which could well be a result of linguistic accommodation towards Bantu (Dimmendaal 2001b). But all prefixes on nouns exhibit very conservative functions and meanings, and there is no indication of Bantu calques or the spread of Bantu morphological and/or classificatory patterns into Chopi. Moreover, Bantu prefixes (as in *mù-nóórò* ‘Nyoro man’) are re-analysed as masculine sex-indicating morphemes, thus totally deviating from their original meaning in Benue-Congo. Consider the following examples:

(8)	SOURCE	DERIVED FORM		CATEGORY
támò	‘think’	à-támò	‘thought’	abstract noun
tíí	‘be old’	ò-tíí	‘old age’	abstract noun
mééró	‘be drunk’	à-mééró, PL lù-mééró	‘drunkard’	agent noun
kòjò	‘help’	là-kòjò, PL lù-kòjò	‘helper’	profession
ʔúmò	‘cover’	là-ʔúm	‘cover of pot/granary’	instrument
cam	‘eat’	kà-sham	‘eating place, restaurant’	locative noun
jòok	‘spirit, god’	à-jòok (FEM), ò-jòok (MASC)	‘born with six fingers/toes’	personal name
àchólí	‘Acholi’	ná-rachólí (FEM), wôd-achólí (MASC)		ethnonym
nóórò	‘Banyoro’	ná-mùjòórò (FEM), mù-nóórò (MASC)		ethnonym
jò-yugáandà	‘Ugandans’	ná-jò-yugáandà (FEM), wôd-jò-yugáandà (MASC)		group term
cól	‘black’	mà-shól	‘black cattle/chicken’	livestock term
lúò	‘Luo’	pà-lúò	‘Chopi area’	locative noun

It seems as if by constructing innovative forms that resemble those in the dominant contact languages, Chopi converges to northern interlacustrine Bantu. This tendency is also found in some domains of the phonology, where convergence with Benue-Congo is obvious (Storch 2011b). But instead of creating something like an “emergency language”, or experiencing a process of creolisation, speakers otherwise basically cling onto Chopi (and Nilotic) concepts and features. Cherishing their group identity has become rather crucial for survival, in the sense of creating some distance between themselves and other Northerners such as the Acholi. Moreover, the implementation of identity-marking institutions such as dancing troupes help to support social networks that clearly aid individuals in finding work, receiving financial support for education, and so on. This makes Chopi a rather “vital” language that is full of meaning and fulfills several important social functions for its speakers.

4 Strategies of group formation in a Frontier context

This is where the importance of names and linguistic etiquette comes in: even though there exists a multidiverse background and speakers have fluid linguistic identities, there are ideologically motivated strategies that come into play when negotiating and creating community. Although in Chopi communities, diversity and change are managed through the creation of social cohesion and positively evaluated ties between members, whereby foreigners are encouraged to marry into the community, and orphaned children are integrated into families regardless of their ancestry, there are certain aspects of knowledge and linguistic capacity that distinguishes core-group people from others. The following proverb helps to illustrate this:

- (9) dánò má-pee ɲéy-ò jámì má-còon pee tweeró lòk
 person REL-NEG know-PERF thing:PL REL-PAST NEG be_unable-PERF talk
 kí dàdíit
 DIR big_one
 ‘a person who doesn’t know things of the past cannot talk to a big somebody’

It is only through the knowledge on clans, praise names and internal relations between community members that a speaker can interact adequately. An outsider may learn how to do this, but this task takes time to be accomplished. And that the integration of outsiders isn't free from conflict is illustrated in various folktales. In one of them, a young man wants to marry a girl of unclear origin. The beautiful bride looks like a dog to all of his relatives, who want to chase it away. It takes many nights of hiding the dog and convincing the parents until they agree:

- (10) gù-wác-í: íí áan wôd ì-kél-ó gîn mà-ràc
 3PL-say-IMPER INTER 1SG.POSS son 2SG-bring-PERF thing REL-bad
 ì-kwaan-ó gwók ên mù-kélè jákò mà-cí!
 2SG-bring-PERF dog CONS NOM-bring girl REL-nice

'They say: oh! Son of mine, you brought a bad thing. You brought a dog, and then [you are] a bringer of a beautiful girl!'

Only after she is recognised as a real person, the young bride is offered a place in the house and may become a member of the community. She is not uttering a single word throughout the entire tale, but remains silent. This may change as she becomes part of the society.

Meanwhile, change continues to take place: Ongoing processes of cultural change among the Chopi are driven by social marginalisation and economic interests. A major reason for many people's interest in changing alliances and moving from Acholi into Bunyoro territory is the fact that many villages suffered dramatically from the attacks of the Lord's Resistance Army (LRA) during the last decade. As the LRA is basically associated with the Acholi by most Chopi, and as they have been unable to receive as much support and aid after the destruction of their villages as other groups, many people have decided on changing group affiliations and leaving their home areas.

This tendency has not only resulted in increasing language shift, but also in a relatively new concept of Chopi as a language and community being distinct from that of Acholi. The Chopi, therefore, increasingly consider their language and group identity as NON-ACHOLI and NON-LUO, as this sets them apart from the North and its violent conflict, which presented a lethal threat to them as a minority group. By emphasising Chopi identity, the choice of becoming part of the Bunyoro area is conceptualised and marked as a strategy that helps the Chopi to escape identification with those who suffered most from the civil war and continuing threat, namely the Lwoo-speaking inhabitants of Uganda's north.

This process, together with the testimonies of the speakers themselves is exactly the situation as we find it throughout the African Frontier. Here, language in a setting of superdiversity and salient multilingualism poses a crucial question: where should we draw boundaries between languages, how do we dare define delimitations between them? For example, the Chopi as a group of multilingual speakers are associated to the Nilotic language Chopi, but ascribed identities such as BEING/SPEAKING CHOPI are multiple and fragile, and in a fuzzy-bordered area, are subject to ever-changing ideas of group membership. Kopytoff (1987) proposes a model that reconciles the observation of constant change, migration and upheavals and the apparent continuity of exchange and interaction between persisting groups. His African Frontier model has helped to understand the particular processes that have been at work in the context of the establishment of multilingual societies. Kopytoff observes that the history of these societies has been more complex than formerly suggested by conventional models, namely a continuous creation of frontiersmen, and the constant mobility of groups. Instead of the creation of fixed ethnic groups and nations that comprise of a specific people, area and language, such as in nineteenth century Europe, the societies of our area of study have undergone a different process. They existed as or alongside with empires and kingdoms, but these were never stable constructions. Hegemonies were flexible and hardly had real borders; their focal point in Kopytoff's model is the periphery and not the centre, the frontiers and marginalities. This is owed to the insight that a preferred conflict solution strategy has been, and often continues to be, migration and retreat. Conflict, but also crop failure and hunger crisis and other situations of crisis tended to result in the migration of inferior groups or clans into the bush or areas that belonged to other hegemonies. There, new settlements could be founded and new communities established. These later eventually turned into centres of power themselves.

Power in these communities was achieved through attraction of adherents. The more followers one had the safer and more powerful was a village. With the immigration of other people into the area, a canonic history (focusing on a common ethnic origin) was established. This concept of history clearly had to do with a construction of common identity, but still it co-existed with differing histories of kin-groups (focusing on the story of ancestors who originate from various places). Interestingly, Kopytoff correlates these processes with the emergence of an institutionalised and socially most meaningful multilingualism. He claims that besides the public language of the consolidated community, the languages of the separate immigrant groups continued to be spoken.

The exchange between communities, both in form of migrating people and in form of ritual exchange, creates a regional context which incorporates areal features: communities were accepting hegemonies by adopting regional patterns and practices (e.g. shrines and deities) within extremely dynamic borders and over a long period of time. Whenever stabilising frontier societies became metropolises themselves and produced new borders, the process and emergence of new frontier groups experienced its continuous repetition.

Hence, the continuity of mobility was guaranteed by continuous crisis and collapse; instability can thereby be understood as an equilibrium which generates convergence. As proposed by Dixon (1997), linguistic areas emerge through long periods of equilibrium during which contact between these languages is balanced and power relations between minority groups are insignificant. Any incident like the immigration of other people, as in the colonial context, or globalisation pushes for example, are punctuations of the equilibrium. These punctuations can interrupt a convergence process, change it, or stop it.

5 Some conclusions

In a community where people speak several languages on a daily basis, they do not only interact in a multilingual setting, knowing different languages, but they will have to acquire knowledge on how to speak to whom and when in these different languages. Learning these languages and ways of speaking may take a lifetime, and linguistic biographies may differ significantly among the members of a given community (Lüpke & Storch 2013). But the decisive moment is that it is not only multilingualism in a broad sense that plays a role here, but also the knowledge and management of the relevant speech varieties and sociolects of all the different languages known to individual speakers. This holds true for specialised vocabularies, ritual form, and the use of names in greetings.

It is exactly in the context of the Frontier theory that I want to locate Chopi linguistic etiquette, which simply helps to include as many people as possible in an open, fragile community of people with most fluid identities. The contact phenomena that appear to be so salient in this language's phonology, morphology, syntax and lexicon have emerged through the processes that took place in an "equilibrium of constant changes". Since most if not all our theories on linguistic change are based on ideas about language use, a deep study of how a language such as Chopi is actually used will shed much light on Western Nilotic history.

That the situation described here is characteristic for Lwoo-speaking groups elsewhere is supported by comparative data. Reh (2001) describes that for 'courting', another important and conventionalised access ritual, a term is used that has an older meaning as 'speak in a foreign language'. Her conclusion is that the expansion of the early Southern Lwoo groups largely occurred by exchanging daughters with local groups. Other evidence comes from Northern Lwoo: in Sudanese Luwo ("Jur"), we can observe almost as many contact-induced changes of the morphology and syntax as in the south. And most Luwo speakers I talked to emphasised how much their ancestors and present families liked to marry people from outside (Storch 2014). There was a story about abandoned children of missionaries in the 1950ies, and about escaped Arabic concubines in the 1880ies. They all have become somebody's grandmother or aunt, and they are all reported to have been treated in a straight but most polite way, regardless of their not-so nice white skin.

Abbreviations

CONS	consecutive	NOM	nominaliser
DIR	direction	PAST	past tense
FEM	feminine	PERF	perfective
IMP	imperative	PL	plural
INTER	interrogative	POSS	possessive
MASC	masculine	REL	relativiser
NEG	negation	SG	singular

References

- Blount, Ben & Richard T. Curley. 1970. The southern Luo languages: A glottochronological reconstruction; *Journal of African Languages* 9.1: 1–18.
- Brown, Penelope & Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Evans-Pritchard, Edward E. 1948. Nuer modes of address. *The Uganda Journal* 12: 166–171.
- Heusing, Gerald. 2004. *Die südlichen Lwoo-Sprachen* (Nilo-Saharan). Cologne: Köppe.
- Irvine, Judith T. 1974. Strategies of status manipulation in the Wolof greeting. In Richard Bauman & Joel Sherzer (eds.), *Explorations in the Ethnography of Speaking*, pp. 167–191. Cambridge: Cambridge University Press.
- Irvine, Judith T. 1998. Ideologies of honorific language. In Kathryn A. Woolard & Paul V. Kroskrity (eds.), *Language Ideologies. Practice and Theory*, pp. 51–67. Oxford: Oxford University Press.
- Irvine, Judith T. & Susan Gal. 2000. Language ideology and linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of Language*, pp. 35–83. Santa Fe: School of American Research Press.
- Kiessling, Roland & Maarten Mous. 2004. Urban youth languages in Africa. *Anthropological Linguistics* 46.3: 303–341.
- Kolbusa, Stefanie. 2000. Ingamwana. Nyakyusa-Schwiegermeidung. Universität zu Bayreuth.
- Kopytoff, Igor. 1987. The Internal African Frontier: The Making of African Political Culture. In Igor Kopytoff (ed.), *The African Frontier: The Reproduction of Traditional African Societies*, pp. 3–84. Bloomington/Indianapolis: Indiana University Press.
- Lüpke, Friederike & Anne Storch. 2013. *Repertoires and Choices*. Berlin: Mouton de Gruyter.
- Reh, Mechthild. 1996. *Anywa Language* (Nilo-Saharan, 11). Cologne: Köppe.
- Reh, Mechthild. 2001. Migration, polygamy and interethnic marriage as causes of language birth: The Southern Lwoo languages. In Rainer Vossen, Angelika Mietzner & Antje Meissner (eds.), *Mehr als nur Worte*, pp. 583–595. Cologne: Köppe.
- Storch, Anne. 2005. *The Noun Morphology of Western Nilotic* (Nilo-Saharan, 21). Cologne: Köppe.
- Storch, Anne. 2011a. *Secret Manipulations*. New York: Oxford University Press.
- Storch, Anne. 2011b. Grammatical categories of the noun in Chopi. *Studies in Nilotic Linguistics* 3: 37–51.
- Storch, Anne. 2014. *A Grammar of Luwo. An Anthropological Approach*. Amsterdam: Benjamins.
- Treis, Yvonne. 2005. Avoiding their names, avoiding their eyes. How Kambaata women respect their in-laws. *Anthropological Linguistics* 47.3: 292–320.

IS AROID NILO-SAHARAN OR AFRO-ASIATIC? SOME EVIDENCES FROM PHONOLOGICAL, LEXICAL AND MORPHOLOGICAL RECONSTRUCTIONS

Moges Yigezu

Abstract

The classification of Aroid as the southern branch of Omotic appears to be the received opinion among many linguists, though the genetic position of this group remains controversial. Fleming (1974, 1976, and 1988) maintains the position that the Aroid group constitutes the southern branch of the Omotic family (see also Tsuge (1996) for a similar view). Lamberti (1991 and 1993) insists on the separation of the Aroid group from west Cushitic as a separate branch of Cushitic. Bender (1988: 131) states that Aroid languages may originally be non-Nilotic Nilo-Saharan that were integrated at a much earlier date and then developed independently. Later, Bender (2000, 2003) in his comparative study of Omotic reiterates that Aroid is perhaps not Afro-Asiatic. Zaborski (2004), on the other hand, strongly argues, on morphological grounds, that Aroid or South Omotic languages belong to Nilo-Saharan. Similarly, Yigezu (2007), based on the reconstruction of the vowel system of Proto-Aroid, concludes that the Proto-Aroid vowel system is more similar to neighboring Nilo-Saharan languages such as Surmic and Nilotic than to the rest of the Omotic family and consequently suggests that the Aroid group may have a Nilo-Saharan origin. A more recent paper by Theil (2012) entitled “Omotic” questions the classification of the whole family of Omotic under Afro-Asiatic and argues that the Omotic family as a whole should be treated as an independent language family since no convincing evidence has ever been presented to show that Omotic is indeed Afro-Asiatic. We may label the different positions on the genetic classification of Aroid languages as follows: the “Omotic theory”, the “Cushitic theory”, the “Nilo-Saharan theory”, and the “Non-Afro-Asiatic theory”.

The objective of the current contribution is to revisit the genetic position of the Aroid group and present evidence in favor of the Nilo-Saharan theory based on phonological, lexical and morphological reconstructions of the Aroid group. The study provides further evidence by comparing Proto-Aroid against Proto-Surmic and Proto-Nilotic forms.

1 Introduction

The Aroid group consists of **Kara, Hamar-Benna, Aari, Galila** and **Dime** (Bender 1988, Fleming 1988). Geographically, all languages and dialects are spoken in the South Omo zone of the Southern Regional State of Ethiopia. Previous attempts in the reconstruction of the Proto-Aroid include Fleming (1988) and Tsuge (1996). Fleming gives a preliminary analysis of the Proto-South-Omotic sound system. Tsuge is a follow up to Fleming's analysis and tries to clarify some phonetic correspondences within South Omotic. Both pioneering attempts are preliminary in nature and, in some cases, based on secondary and insufficient data. The database used for this study covers over 300 basic lexical items collected by the author for all languages in a recent fieldwork.

The objective of the present contribution is to present some historical-comparative evidences regarding the genetic position of the Aroid group of languages by making systematic phonological correspondences between the languages via the comparative method. In doing so, this study will focus on the following: (a) reconstruct the phonemic system of Proto-Aroid and describe some of the sound changes that occurred in the history of these languages, (b) make a lexical reconstruction of the Proto-Aroid forms based on the 300 basic lexical items, and (c) provide morphological reconstruction of independent pronouns, demonstratives and some verbal affixes. Based on the phonological, lexical and morphological data a comparison of Proto-Aroid forms will be made against Proto-Surmic and Proto-Nilotic in order to revisit the current controversy on the genetic position of Aroid languages.

2 Genetic Classifications

The internal unity of Aroid languages has not been questioned so far. Although the classification of Aroid as the southern branch of Omotic seemingly looks as the received opinion among many Africanists, the genetic position of this group remains controversial.

Fleming (1988) maintains the position that the Aroid group constitutes the southern branch of the Omotic family. Bender (1988: 131) has expressed doubt by saying that South Omotic languages may originally be non-Nilotic Nilo-Saharan who were integrated at a much earlier date and then left to develop independently but supports the Omotic theory. Lamberti (1991 and 1993) insists on the separation of “Aroid” group from west Cushitic as a separate branch of Cushitic. Zaborski (2004), on the other hand, strongly argues, on morphological grounds, that Aroid or South Omotic languages belong to Nilo-Saharan. Similarly, Yigezu (2007), based on the reconstruction of the vowel system of Proto-Aroid, concludes that Proto-Aroid vowel system is more similar to some neighboring Nilo-Saharan languages such as Surmic and Nilotic than the rest of the Omotic family suggesting that the Aroid group may have a Nilo-Saharan origin. A more recent paper by Theil (2012) entitled “Omotic” questions the classification of the whole family of Omotic under Afroasiatic and argues that the Omotic family as a whole should be treated as an independent language family since no convincing evidence has ever been presented to show that Omotic is indeed Afro-Asiatic. Particularly, Theil (2012: 376) concludes, largely based on data from Dime (Seyoum 2008), that Aroid languages along with Dizoid “must be treated as unrelated to Omotic until the opposite has been proved with traditional comparative method”. The author further argues that the most plausible explanation for the lexical similarities between Aroid and the rest of Omotic is borrowing. We may label the different opinions and positions on the genetic classification of Aroid languages as follows: the “Omotic theory”, the “Cushitic theory”, the “Nilo-Saharan theory” and the “Non-Afro-Asiatic theory”.

3 Methodological issues

In historical-comparative analysis, there are different methods and techniques used particularly in determining the genetic relationship between languages and in grouping them into language families. Within Omotic and the Aroid group the most commonly used methods are *mass comparison*, *lexicostatistics* and *morphological comparisons*.

Fleming (1969) bases his analysis on the methods of lexicostatistics and grammatical comparison. In his 1974 article Fleming uses the methods of morphological and lexical comparison considering them as two separate methods of comparisons. Bender (2000, 2003) largely depends on morphological comparisons, and using the same method he came up with four Proto-Omotic lexical cognates and 25 grammatical morphemes in support of the Omotic theory. Similarly, Lamberti gives some comparisons of grammatical morphemes from various Cushitic and Omotic languages and further argues that morphological comparison “represents the most conservative and intimate part of a language” (1991: 556).

Mass comparison, as described by Campbell (1997: 210), is looking at lexical similarity shared across many languages, as evidence of genetic relationship with no methodological considerations deemed relevant. Campbell contends that since resemblances detected in mass comparison are subject to investigation whether they are due to inheritance from a common ancestor or whether they result from borrowing, accident, sound symbolism, etc. the results frequently have proven erroneous or at least highly controversial. The method of lexicostatistics involves measuring the percentage of words with similar sound and meaning (cognates) in different languages on the basis of lists of basic vocabulary. The larger the percentages of cognates, the closer the languages being compared are presumed to be related. Lexicostatistics, as a historical-comparative method, has its own weaknesses and no serious conclusions can be drawn on the genetic relationship between languages unless lexical data are presented. On the other hand, assuming the historical stability of morphology cannot be taken for granted. Thomason (1980), cited in Campbell 1997: 222 – 223), shows that “morphology is by no means so stable as to justify the assumption that lexical cognates may vanish almost entirely while the morphology holds firm” (1980: 360). Depending solely on morphological correspondences without having the corresponding regular phonological and semantic correspondences would be of little value. Methodologically, what was lacking, in the previous comparisons made on different branches of the

Omotic family in general and the Aroid group in particular, is establishing a systematic phonological correspondences between the languages in question via the comparative method. This is the task the present study is undertaking.

4 Phonological systems of Aroid languages

As a background to the forthcoming discussions, in this section we shall have a brief look at to the phonological systems of the Aroid languages. Linguistic descriptions on these languages include: Hayward (1990), Menuta (2011) and Seid (2011) on Aari; Lydall (1976) on Hamar; Fleming (1990), Seyoum (2007, 2008) on Dime and Yigezu (2007) on Kara.

4.1 The vowel systems

The whole group of Aroid seems to have a ten-vowel system. Hamar has been described as a ten-vowel system (Lydall 1976) grouped into two harmonic sets usually described as ATR system in many African languages such as Eastern Nilotic and Surmic. Kara has been described as having a similar ten-vowel system (Yigezu 2007). Aari (Hayward 1990) and Dime (Fleming 1990), on the other hand, are described as having a five-vowel system. As shown in Yigezu (2007), however, the ten-vowel system is the basic configuration for the whole genetic unit, a system that is also very similar to the neighboring Nilo-Saharan languages such as Surmic (Yigezu 2001) and Nilotic (Dimmendaal 1988).

As suggested by Hayward (1990), Bender (1988) and Yigezu (2007) breathy voice quality is a characteristic feature of the vowel system of Aroid languages. Particularly, remnants of breathy voice quality in vowels are still apparent in Aari and Kara. On the other hand, although all languages have demonstrated a ten-vowel system prevalent in many African languages vowel harmony is virtually absent in all the systems.

4.2. The consonant systems

The consonant system of Aroid languages is fairly complex in the stop series having voiced stops [b, d, g], voiceless stops [p, t, k, ʔ], front articulated implosives [ɓ, ɗ], and largely back articulated ejectives [s', tʃ', k'/q']. Hamar has also the velar implosive [ɠ] in addition to the front articulated implosives while Dime has only the alveolar implosive [ɗ] and the whole series of ejectives [p', t', k', s', tʃ']. Two palatal affricates [dʒ and tʃ] are also recorded but the contrast between the alveolar affricates [ts, dz] is almost non-existent (cf. North Omotic languages) except in Aari and Dime where the voiceless counterpart [ts] occur. Fricatives range from labio-dental to velar and glottal point of articulations, [f, s, z, ʃ, ʒ, x, ɣ, h]. Dime has two additional voiced fricatives, namely, [ɣ] and [β] in its inventory. Liquids (lateral [l] and trill [r]), nasals (bilabial [m] and alveolar [n] as well as [ŋ] in Hamar and Dime) and semi-vowels are also part of the phonemic system of these languages. The detailed descriptions of each consonant will be given in conjunction with the comparative analysis in the following section.

5 Phonological reconstructions

5.1 Proto-Aroid vowel phonemes

As shown in Yigezu (2007), the ten-vowel system is the basic configuration for the whole genetic unit: *i, *ɪ, *u, *ʊ, *e, *ɛ, *o, *ɔ, *ɤ and *a. This system is very similar to neighboring Nilo-Saharan languages such as Surmic (Yigezu 2001) and Nilotic (Dimmendaal 1988). In Aari, the breathy vowel /ü/ as the structural counterpart of the plain vowel /u/ is lacking. Similarly, the same breathy vowel is absent in Nuer (Yigezu 1995), Dinka (Malou 1988) and Anuak (Keefer et al. 1976; Lusted 1976). Other authors such as Anderson (1987) for Dinka and Reh (1996) for Anuak state that these languages do have such a vowel.

5.2 Proto-Aroid consonant phonemes

In the comparative data, the comparison was made for C1 (stem-initial consonant), V1 (stem-initial vowel), C2 (stem-medial consonant) and C3 (stem-final consonant). The comparison was based on the basic form of a noun or a verb stripped off any affix. Note that the reconstruction of vowel phonemes is not the concern of this paper as it is already dealt with separately (see Yigezu 2007).

Voiced stops

The voiced stop consonants are **b**, **d** and **g**, all are found in all languages and dialects occurring at all positions. Tables 1 – 3 illustrate the correspondences at different positions within a word.

Table 1: Sample of Proto-Aroid correspondences for *b

	‘beard’	‘untie’	‘smoke (n.)’	‘rain’	‘other’
Kara	booci	bul-	c’ubi	doobo	ab
Hamar-Benna	buufi	bul-	c’uba	doobi	ab
Aari	buuci	bul-	c’uba	doobi	ab
Galila	buuci	bul-	c’uba	doobi	ab
Dime	[s’ire]	bul-	c’ubu	diibi	[balub]
Proto-Aroid	*b-	*b-	*-b-	*-b-	*-b

Table 2: Sample of Proto-Aroid correspondences for *d

	‘grind’	‘rain’	‘hold’	‘count’	‘man/human being’
Kara	des-	doobo	yΛd-	payd-	edi
Hamar-Benna	des-	doobi	yΛd-	payd-	edi
Aari	deys-	doobi	yΛd-	payd-	eed
Galila	deys-	doobi	yΛd-	payd-	eed
Dime	dis-	diibi	yiidi	fayd-	[iɾye]
Proto-Aroid	*d-	*d-	*-d-	*-d-	*-d

Table 3: Sample of Proto-Aroid correspondences for *g

	‘claw’	‘canoe’	‘crocodile’	‘blow’	‘male/husband’
Kara	gufo	gaagi	gurgur	pug-	aŋge
Hamar-Benna	gufo	gɔŋgɔla	gurgur	fug-	aŋgi
Aari	guɸma	gonga	gurgur	fug-	?aŋg
Galila	gusa	gonga	gurgur	fug-	?aŋg
Dime	guɸ	gongu	guuru	[fitse-]	[goɸtu]
Proto-Aroid	*g-	*g-/*-g-	*g-/*-rg-	*-g-	*-g-

The voiced stops ***b**, ***d** and ***g** have been reconstructed as part of the Proto-Aroid consonant system occurring at stem-initial, medial and final positions. Note, however, that Dime is not only lacking the word-final voiced stops but also cognates are not found for the above correspondence sets. In Kara and Hamar-Benna, although cognates are found, the voiced alveolar and velar stops do not seem to occur word-finally. Whereas in Aari and Galila, the northern members of the group, the alveolar and velar stops appear word-finally as well.

Voiceless stops

The voiceless stops ***p**, ***t**, ***k** have been reconstructed as part of the Proto-Aroid consonant system as these are straight forward cases having complete cognate series in the daughter languages. See Tables 4 – 6 below.

Table 4: Sample of Proto-Aroid correspondences for *p

	‘count’	‘mouth’	‘fat’	‘moon’	‘knife’	‘cry’
Kara	payd-	ʔapo	durpi	arpi	halpa	ʔep-
Hamar-Benna	payd-	ʔapo	durpi	arpi	halpa	ʔep-
Aari	payd-	ʔapo	durpi	arpi	halpa	ʔef-
Galila	payd-	ʔapo	durpi	arpi	[mʌʃa]	ʔef-
Dime	fayd	ʔape	[nʌrʒɔb]	irpa	halpi	ʔiip-
Proto-Aroid	*p-	*-p-	*-rp-	*-rp-	*-lp-	*-p-

In Dime, in the word for ‘count’, the voiceless bilabial fricative /f/ corresponds to the voiceless bilabial stop /p/ in other languages. Dime has both phonemes whereas the other Aroid languages lack the bilabial fricative. In the northern Aroid languages, Aari and Galila, as can be seen in the word ‘cry’, there is a regular sound change taking place at stem-final position, i.e. /p/ is changed into /f/ stem-finally. This is a recurrent sound change observed elsewhere as well. See Table 10 below, item for ‘wet’; Table 12, item for ‘eye’ (the change in this particular case took place word-finally); and Table 15, item for ‘get’. Compare Table 21, item for ‘fat’, where the same sound change does not take place because of the different phonetic environment, i.e. the consonant /p/ occurs medially before a consonant.

Table 5: Sample of Proto-Aroid correspondences for *t

	‘cut’	‘ten’	‘we’	‘seven’	‘return’	‘climb’
Kara	tʌx-	tʌʃi	woti	tsobäh	maat-	wut-
Hamar-Benna	tax-	tʌʃi/taʃi	wodi	toba	maat-	ʔut-
Aari	tʌx-	tamma	wota	tʌbza	maat-	[eb-]
Galila	tʌc-	tamma	wota	tʌbza	maat-	[eb-]
Dime	[qars’-]	tamme	wotu	tusom	[womɔ]	wut-
Proto-Aroid	*t-	*t-	*-t-	*t-	*-t-	*-t-

Table 6: Sample of Proto-Aroid correspondences for *k

	‘dig’	‘divide’	‘nose’	‘donkey’	‘cattle’
Kara	koy-	kʌʃ-	nuki	ukuli	waaki
Hamar-Benna	koy-	kʌʃ-	nuki	ukuli	waaki
Aari	koy-	kʌʃ-	nuki	okulɪ	waaki
Galila	koy-	kʌʃ-	nuki	okulɪ	waaki
Dime	koy-	kiʃ-	nuku	[yʌɾɪ]	wooxɪn
Proto-Aroid	*k-	*k-	*-k-	*-k-	*-k-

Implosives

Synchronically, Dime is quite different from the rest of the Aroid languages. One area of difference is the lack of implosives in Dime. The other languages of Aroid exhibit the front articulated implosives /ɓ/ and /ɗ/. Fleming (1990: 505) recognizes the bilabial and alveolar implosives as phonemes of Dime but in our data implosives are missing altogether. Seyoum (2008: 9) recorded the alveolar implosive consonant. Except for Hamar, for all languages, in our data, there is no velar implosive consonant. Thus, we may postulate that Proto-Aroid had two implosives in its phonemic system, /*ɓ/ and /*ɗ/. The following data in Tables 7 and 8 illustrate this point.

In Aari and Galila, in the item for ‘seven’, the bilabial implosive /ɓ/ changes into the voiced stop /b/ when it occurs before a consonant word-medially. Whereas at intervocalic position as in Table 5, item for ‘ten’, the bilabial implosive /ɓ/ corresponds to a bilabial nasal /m/.

Table 7: Sample of Proto-Aroid correspondences for *ḡ

	‘seed’	‘seven’	‘animal blood’	‘pour’	‘hat’	‘long’
Kara	benta	tsoḃäh	zḡmbi	q’olḃ-	q’obe	gudḡḃ
Hamar-Benna	ḃeta/ḃenta	toḃa	zḡmbi	q’olḃ-	q’oḃi	gudḡḃ
Aari	ḃeta	ḡḃza	zḡmbi	q’olḃ-	q’ḡḃ	gazimi
Galila	[mafa]	ḡḃza	zḡmbi	q’olḃ-	q’ḡḃ	[k’almi]
Dime	[miʃit]	[tuḡum]	[maḡse]	q’olḃ’-	q’op’u	gudumup
Proto-Aroid	*ḃ-	*-ḃ-	*-mḃ-	*-ḃ-	*-ḃ-	*-ḃ

In Table 7 above, item for ‘pour’, the bilabial implosive /ḃ/ changes to a glottal stop /ʔ/ stem-finally. Note also that when the implosive is homorganic to the preceding consonant no sound change applies as in the item for ‘animal blood’. All the changes are regular sound changes attested across languages and phonetically motivated. Depending on the phonetic environment in which it occurs the bilabial implosive very often changes into its voiced stop counterpart or to a bilabial nasal (usually in its geminated form which might result from some kind of assimilation, *ḃm > mm) or to a glottal stop stem-finally or medially after a consonant. See Table 12, item for ‘leaf’.

Table 8: Sample of Proto-Aroid correspondences for *ḡ

	‘tie’	‘know’	‘show’	‘buttocks’	‘hot’
Kara	ḡax-	ḡas-	ḡaw-	tudḡ	ʔoyḡa
Hamar-Benna	ḡax-/ḡax-	ḡes	ḡoy-/ḡay-	tudḡ	ʔoyḡ
Aari	ḡax-	es-	ḡaw-	tudḡ	ʔoyḡ
Galila	[ʔak-/zok-]	es-	[ʃeci-]	tudḡ	[sol]
Dime	[zuk-]	---	[ʔeḡs-]	[goyo]	[suulum]
Proto-Aroid	*ḡ-	*ḡ-	*ḡ-	*-ḡ-	*-ḡ

The alveolar implosive /*ḡ/ occurs at all positions within a word and based on the correspondence sets given in Table 8 above, it has been reconstructed as part of the phonemic inventory of Proto-Aroid.

Ejectives

In Kara, Hamar-Benna, Aari and Galila we have the back articulated ejectives such as /c’/ and /q’/ as well as the fricative ejective /s’/. Surprisingly, Dime is so divergent from the rest of the languages in this respect having the full series of ejectives in its phonemic inventory. These include: /p’/, /t’/, /c’/, /k’/, /q’/ and /s’/. Hence, comparisons are made for the ejectives commonly found in all languages, namely /c’/, /q’/ and /s’/.

Table 9: Sample of Proto-Aroid correspondences for *c’

	‘rotten’	‘smoke’	‘wet’	‘root’	‘laugh’
Kara	c’afi-	c’ubi	c’api	c’ac’i	hanc’-
Hamar-Benna	c’afi-	c’uba	c’api	c’ac’i	hanc’-
Aari	c’af-	c’uba	c’afi	c’ac’i	yinc’-
Galila	c’af-	c’uba	c’afi	c’ac’i	yinc’-
Dime	c’uc’up-	c’ubu	[ʔmub]	c’nc’i	yinc’-
Proto-Aroid	*c’-	*-c’-	*c’-	*-c’-	*-c’-

As shown in Table 9 above, there are correspondence sets for the palatal ejective /c’/ word-initially, word-medially and stem-finally. We can, therefore, safely reconstruct the phoneme /*c’/ as part of the phonemic inventory of Proto-Aroid.

Table 10: Sample of Proto-Aroid correspondences for *q

	‘hat’	‘pour’	‘tree’	‘water’	‘frog’
Kara	q’obe	q’olb-	haq’a	nunq’o	panaq’o
Hamar-Benna	q’obi	q’olb-	haaq’a	noq’o	panaq’
Aari	q’ɔb	q’olʔ-	haχa	noq’o	fanq’a
Galila	q’ɔb	q’olʔ-	haq’a	noq’a	fanq’a
Dime	q’op’u	q’olp’-	ʔaaχe	naχe	(madɪ)
Proto-Aroid	*q’-	*q’-	*-q’-	*-q’-	*-q’-

The uvular ejective /q’/ is attested word-initially and word-medially but not word-finally. In Dime, in the word for ‘water’, the uvular ejective is weakened to a velar fricative intervocalically. The same process of weakening took place for the word ‘tree’ in Aari between two low vowels but not between two high vowels as in the case of the item for ‘water’. The uvular ejective /*q’/ has been reconstructed as a phoneme of Proto-Aroid word initially and medially.

Table 11: Sample of Proto-Aroid correspondences for *s’

	‘vomit’	‘short’	‘milk’	‘shoot’	‘close’
Kara	s’a-	s’eedi	raas’i	kʌs’-	dis’-
Hamar-Benna	s’a-	[orɔ]	raat’i/raas’i	kʌs’-/kas’-	dis’-
Aari	s’a-	s’eedi	raas’i	kas’-	zis’-
Galila	s’a-	s’eedi	raas’i	kas’-	zis’-
Dime	s’aa-	s’edob	[jiʔi]	qɪs’-	zis’-
Proto-Aroid	*s’-	*s’-	*-s’-	*s’-	*-s’

The fricative ejective /*s’/ has been reconstructed as a member of the phonemic inventory of Proto-Aroid based on the correspondence sets given in Table 11 above. Note that in Hamar-Benna the alveolar ejective /t’/ alternates with the fricative ejective /s’/ in a few lexical items. The former has no phonemic status in the language.

The glottal stop /ʔ/, on the other hand, has a defective distribution in all languages. It does occur frequently word-initially and less frequently word-medially but not word-finally. As can be seen in Table 12 below, item for ‘human blood’, the glottal stop tends to weaken intervocalically in Aari, Galila and Dime. Nevertheless, this is not a regular sound change attested elsewhere; it is more of a sporadic change.

Table 12: Sample of Proto-Aroid correspondences for *ʔ

	‘give’	‘eye’	‘human blood’
Kara	ʔim-	ʔapi	maʔas
Hamar-Benna	ʔim-	ʔapi	(zombi)
Aari	ʔom-	ʔafi	maχas
Galila	ʔim-	ʔafi	maχas
Dime	ʔim-	ʔaape	maχse
Proto-Aroid	*ʔ-	*-ʔ-	*-ʔ-

Fricatives

All languages of Aroid have the following fricatives: /f/, /s/, /z/, /ʃ/, /χ/, and /h/. But Dime once again radically diverges from the rest of the languages and adds more fricatives into its phonemic inventory. These are: /ʒ/ and /fi/. Since the rest of the languages do not have correspondences for these phonemes attested only in Dime the reconstruction would be made for the commonly found fricatives in all languages.

Table 13: Sample of Proto-Aroid correspondences for *f and *s

	‘rotten’	‘blow’	‘human blood’	‘calabash’	‘grind’	‘sweep’
Kara	c’afi-	pug-	maʔas	gusi	des-	saa-
Hamar-Benna	c’ʌfi-	fug-	[zɔmbi]	koysi	des-	sah-
Aari	c’ʌfi-	fug-	maχas	gusi	deys-	[ʔaxiz-]
Galila	c’ʌfi-	fug-	maχas	gusi	deys-	[ʔeqiz-]
Dime	c’uc’up-	[fitse]	maχse	[mange]	dis-	sah-
Proto-Ariod	*f-	*f-	*-s	*-s-	*-s	*s-

While the bilabial fricative /f/ is attested only word-initially and medially the alveolar fricative /s/ is found word-initially, medially as well as stem-finally. Both phonemes are reconstructed as phonemes of Proto-Aroid.

Table 14: Sample of Proto-Aroid correspondences for *z, *ʃ and *χ

	‘animal blood’	‘body’	‘claw’	‘divide’	‘human blood’
Kara	zɔmbi	[biʃi]	guʃo	kʌʃ-	maʔas
Hamar-Benna	zɔmbi	zʌʃa	guʃo	kʌʃ-	[zɔmbi]
Aari	zɔmbi	zʌna	guʃma	kʌʃ-	maχas
Galila	zɔmbi	zʌna	gusa	kʌʃ-	maχas
Dime	[maχse]	zʌʃe	guʃ	kiʃ-	maχse
Proto-Ariod	*z-	*z-	*-ʃ-	*-ʃ-	*-χ-

The voiced alveolar fricative /z/ looks like a marginal consonant in all languages. A correspondence set could be found only for word-initial position. Whereas for the palatal fricative /ʃ/ and the voiceless uvular /χ/ correspondence sets are found only word-medially. Nevertheless, all of them are considered as phonemes of Proto-Aroid only in certain positions within a word.

Table 15: Sample of Proto-Aroid correspondences for *h

	‘descend’	‘get’	‘six’	‘sweep’
Kara	hanc-	haap-	lä(h)	saa-
Hamar-Benna	hanʃ-	haf-	läh	sah-
Aari	hac-	haf-	la	[ʔaxiz-]
Galila	hac-	haf-	la	[ʔeqiz-]
Dime	yic-	yʌp-	lahe	sah-
Proto-Ariod	*h-	*h-	*-h-	*-h

Based on the comparative data given in Table 15 above, the glottal fricative consonant /*h/ has been reconstructed as a member of the Proto-Aroid phonemic system word-initially and word-medially as well as stem-finally.

Affricates

Synchronically, the only affricate recorded in all languages is the voiceless palatal affricate /c/ - an IPA equivalent of it is /tʃ/. Dime has additional affricates such as the alveolar affricates /ts/ and /dz/, which are absent in the rest of the Aroid languages but are typical features of North Omotic languages. It has also the voiced palatal affricate /j/ - represented by the IPA symbol /dʒ/. Due to the divergence of Dime from the rest of the Aroid languages, the reconstruction of affricate is limited to the voiceless affricate /c/.

Table 16: Sample of Proto-Aroid correspondences for *c

	'beard'	'chicken'	'chin'	'guest'	'cold'
Kara	booci	baaca	mooca	cooci	k'Λji
Hamar-Benna	buufi	baaʃa	moʃ	foʃi	k'Λji
Aari	buuci	baaca	mooca	foci	k'aʒi
Galila	buuci	baaca	mooca	soci	k'aʒi
Dime	[s'ire]	[kɔyz]	[gahrc'ɛ]	[yifɔb]	[bahzim]
Proto-Aroid	*-c-	*-c-	*-c-	*-c-	*-j/*-ʒ-

In Hamar-Benna the palatal affricate /c/ is weakened to the palatal fricative /ʃ/ but this change is not a recurrent sound change. As can be seen in Table 16 no cognate has been found for Dime though the phoneme /c/ is part of the phonemic inventory of the language. The voiceless palatal affricate /*c/ has been reconstructed as a phoneme of Proto-Aroid based on the correspondence sets given in Table 16 above.

Nasals

There are three nasal phonemes in all the languages, namely /m/, /n/ and /ŋ/. All the nasal consonants *m, *n and *ŋ have been reconstructed as phonemes of Proto-Aroid as the correspondence sets are relatively straight forward.

Table 17: Sample of Proto-Aroid correspondences for *m

	'human blood'	'chin'	'animal blood'	'egg'	'female'	'give'
Kara	maʔas	mooca	zɔmbi	muqayo	mee	ʔim-
Hamar-Benna	(zɔmbi)	moʃ	zɔmbi	muqa	maa	ʔim-
Aari	maχas	mooca	zɔmbi	muχa	mana	ʔum-
Galila	maχas	mooca	zɔmbi	muqa	maa	ʔim-
Dime	maχse	[gahrc'ɛ]	[maχse]	molu	[ʔamzi]	ʔim-
Proto-Aroid	*m-	*m-	*-m6-	*m-	*m-	*-m

Table 18: Sample of Proto-Aroid correspondences for *n

	'fire'	'frog'	'girl'	'hand'
Kara	nɔ	panaɔo	anza	ʔan
Hamar-Benna	nuu	panaɔ	anza	ʔan
Aari	nɔha	fanqa	anza	ʔaani
Galila	nɔha	fanqa	anza	ʔaani
Dime	nuunu	[madɪ]	[wodɪ]	ʔaani
Proto-Aroid	*n-	*-nc-	*-nc-	*-n-

Table 19: Sample of Proto-Aroid correspondences for *ŋ

	'five'	'canoe'
Kara	dɔŋ	gaagi
Hamar-Benna	dɔŋ	gɔŋɔla
Aari	dɔŋɔ	gɔŋga
Galila	dɔŋɔ	gɔŋga
Dime	[ʃinne]	gɔŋgu
Proto-Aroid	*-ŋ[c]	*-ŋc-

Liquids

The lateral /*l/ and the trill /*r/ consonants have also been reconstructed as phonemes of Proto-Aroid. The correspondence sets are given below in Tables 20 and 21.

Table 20: Sample of Proto-Aroid correspondences for *l

	‘bone’	‘call’	‘corpse’	‘donkey’	‘six’
Kara	lapo	[gissim-]	[dʌmbi]	ukuli	lä(h)
Hamar-Benna	leepi	ɛl-	leesi	ukuli	läh
Aari	lafi	ɛl-	leysi	ökölɪ	la
Galila	lafi	ɛl-	leysi	ökölɪ	la
Dime	[kʰus]	ɛl-	[duuku]	[yʌɾɪ]	lahe
Proto-Aroid	*l-	*-l-	*l-	*-l-	*l-

Table 21: Sample of Proto-Aroid correspondences for *r

	‘day’	‘fat’	‘refuse’	‘plant’
Kara	roro	durpi	gar-	kɔɾ-
Hamar-Benna	roro	durpi	---	kɔɾ-
Aari	ror	durpi	gɪr	kɔɾ-
Galila	ror	durpi	gɪr	kɔɾ-
Dime	[ʔoxtu]	[narzob]	gɪr	kor-
Proto-Aroid	*r-	*-rc-	*g-/*-r-	*k- /*-r-

Semivowels

The semi-vowels /*w/ and /*y/ have been reconstructed as part of the phonemic inventory of Proto-Aroid. The correspondence sets are given below in Tables 22 and 23.

Table 22: Sample of Proto-Aroid correspondences for *w and *y

	‘cattle’	‘dry’	‘dig’	‘count’	‘stand’
Kara	waaki	[sʰedi]	koy-	payd-	wʌy-
Hamar-Benna	waaki	wɔfi [wɔcci]	koy-	payd-	wʌy-
Aari	waaki	wɔci	koy-	payd-	wʌy-/ wʌʔ-
Galila	waaki	wɔce	koy-	payd-	wʌy-/ wʌʔ-
Dime	wooxin	wucim	koy-	fayd-	wiy-
Proto-Aroid	*w-	*w-	*-y-	*-yc-	*w-/*-y

Table 23: Sample of Proto-Aroid correspondences for *w and *y

	‘hold’	‘meat’	‘sun’	‘you (sg.)’
Kara	yʌd-	waa	hayo	yʌti
Hamar-Benna	yʌd-	waa	ayi	yʌdi
Aari	yʌd-	waa/waaha	hayi	yʌta
Galila	yʌd-	waaha	hayi	yʌta
Dime	yidi	wɔhɔ	yɪye	yʌse
Proto-Aroid	*y-	*w-	*-y-	*y-

The consonant phonemes given in Table 24 below have been reconstructed as phonemes of the Proto-Aroid. Although most phonemes are reconstructed in all positions within a word or stem, some are reconstructed only in some positions. Consonants reconstructed only initially and medially include: *glottal stop*, *uvular ejective*, *voiceless bilabial fricative* and the *lateral* consonant. The *velar nasal* is reconstructed only medially and finally. The consonants that are reconstructed only medially are: the *voiceless palatal affricate*, the *voiceless uvular fricative* and the *voiceless palatal fricative*. The *voiced alveolar fricative* has been reconstructed only initially.

Table 24: Proto-Aroid Phonemes

	Labials	Alveolars	Palatals	Velars	Uvular	Glottal
Voiced stops	*b	*d		*g		
Voiceless stops	*p	*t		*k		*ʔ
Implosives	*ɓ	*ɗ				
Ejectives		*s'	*c'		*q'	
Voiceless fricatives	*f	*s	*ʃ	*χ		*h
Voiced fricative		*z				
Voiceless affricate			*c			
Nasals	*m	*n		*ŋ		
Liquids		*l, *r				
Semivowels	*w			*y		

Looking into the general picture, Proto-Aroid makes a four-way distinction of stop consonants – all occurring stem initially, medially and finally; only the glottal stop is attested stem-initially and medially. The velar implosive is missing (except in Hamar) and only the front-articulated implosives are fully attested. This is identical to some neighboring Nilo-Saharan (Surmic: Koegu, Me'en) patterns. Front-articulated ejectives such as /p'/ and /t'/ are missing as compared to North Omotic languages. Generally, implosives tend to favor front articulations while ejectives tend to favor back articulations once again typical to Nilo-Saharan Surmic patterns (Yigezu 2001).

Perhaps the most striking outcome of the phonological reconstruction is that Dime does not simply fit into the pattern given in Table 24 and is much more divergent than Fleming (1988: 163) has thought. Fleming in his preliminary study of “Proto-Somotic” consonant phonemes states that in the family tree Dime is distant from the rest of the Aroid languages than they are from each other, and regarded Dime as one of the sister languages of the Aroid group. Nevertheless, the phonological correspondences and the reconstruction of Proto-Aroid phonology given above do not support this hypothesis. It rather suggests that Dime may be left out of the Aroid group altogether. The rationale are: (1) in the stop series Dime lacks all implosives (except the alveolar implosive) attested in other languages but has exhibited a full series of ejectives, a feature lacking in the other Aroid languages, (2) it has a rich inventory of fricatives (Seyoum 2008), half of which are absent in the other Aroid languages. Phonologically, therefore, Dime radically diverges from the rest of the Aroid languages lacking two of the implosive consonants and having more consonant phonemes in its inventory. If we consider the additional consonants Dime has in its inventory it makes it rather closer to languages such as the Omotic cluster. Whether the additional consonants are innovations in Dime or retention of older forms requires further research in other neighboring languages as well. Thus this study suggests that Dime should be left out of the Aroid group and better considered as an isolate on its own until we learn more about the historical developments in Dime and other languages in the area.

On the other hand, Proto-Aroid phonology has clearly some Nilo-Saharan features such as (1) the ten-vowel system (unlike any other Omotic group), (2) the presence of remnants of breathy voice quality in some of the languages such as Aari and Kara (see Yigezu 2007 for the details), and (3) the stop series are more similar to neighboring Nilo-Saharan languages such as Koegu and Me'en than the rest of the Omotic languages. Dime shares the first two features related to vowels but not the third one.

6 Sound changes

Some sound changes, not so regular, were observed in the reconstruction of Proto-Aroid consonant phonemes. These can be grouped into two types: weakening of stops into fricatives and glides and simplification of pre-nasal stops. The process of weakening affects word-initial and word-medial stop consonants and changes them to fricatives or glides. The change of *p > *f in particular is frequent. Examine the following data:

Table 25: Examples of the process of weakening

‘count’	* payd- in Proto-Aroid > * fayd- in Dime
‘blow’	* pug- in Proto-Aroid > * fog- in Hamar and Aari
‘cloth’	* apala in Proto-Aroid > * afala in Aari
‘cotton’	* putta in Proto-Aroid > * futta in Dime
‘climb’	* wut- in Proto-Aroid > * ?ut- in Hamar
‘cattle’	* waak- in Proto-Aroid > * woox- in Dime

Simplification of pre-nasal stops is another sound change apparent in the system. While the Proto-Aroid form has the pre-nasal stops [-nC-] word medially, one of the daughter languages Aari has dropped the nasal element, i.e. ***nt** > ***t as in: ‘I’, *?inta** in Proto-Aroid > ***?ita**; ‘seed’, ***benta** in Proto-Aroid > ***beta**.

7 Lexical reconstructions

According to Bender (2003: 224) only three Aroid forms have plausible link to Nilotic: buffalo (***meek-**), head (***mat-**) and tail (***go/ul-**) and rightly suggested that for the source of the majority of Aroid lexicon which does not agree with the rest of Omotic, one must look elsewhere than Nilotic. Apparently Surmic gives more positive results than Nilotic. The reconstructed Proto-Aroid forms were checked against Proto-Surmic forms (Yigezu 2001) and some 42 lexical correspondences were attested between Proto-Aroid and Proto-Surmic, some with traces in one or more Surmic languages (See Annex I below). Annex II contains some 20 additional lexical correspondences of Proto-Aroid with Koegu alone. Koegu data should be used with cautions because the language is at the verge of extinction and has a few hundred (around 400) speakers who are shifting to either the Aroid Kara or the Nilotic Nyangatom.

Many of the correspondences between Proto-Aroid and Surmic can be described as ‘safe’ or plausible; some others could possibly be ‘probable’ since traces are found in one or more Surmic languages rather than in Proto-Surmic form. Still few items could be taken as ‘questionable’, and put under question marks in the data. Some of the plausible links to Surmic may also be seen as cases of ‘symbolic items’ such as the nursery term ‘**father**’, ‘**mother**’ (35), ‘**urine**’ (cf. Proto-Aroid: ***jaan** (28)).

Another set of item which is known as “Wanderwörter” may also occur. These items usually have cultural origin from the domain of technology or domestic animals. Such forms show up in widespread languages of several genetic families and appear to be the result of diffusion but with no assignable source (see Bender 2003: 256). These include items such as **cotton** (widespread in Cushitic and Omotic), **mud** and **ten** (widespread in N-S), **donkey** (widespread in N-S and Omotic). There could also be apparent loans from known sources. Bender states that ‘urine’ in Proto-Aroid ***jaan** is possibly borrowed from Amharic **fint**; but it is not easy to ascertain how this Amharic form has been borrowed by several languages of Aroid and Surmic.

In terms of wider connections, Proto-Aroid alleged to have some cognates in common with Proto-Omotic and beyond. According to Bender, Proto-Aroid shares Proto-Omotic cognates such as *eye*, *smoke*, *count*, *root*, *ten*, *man*, *rat*, and *cry* (Bender 1988). The item *ten* has already been cited as a case of “Wanderworter”; the item *man* has a cognate in Surmic (SWS) ***eeet** (cf. Proto-Aroid ***eed**). Chadic and Cushitic correspondences include items for *person/man*, *four* and *ten* (Bender 1988: 155). Two of the cognates (**person/man: Aroid *eed**; and **ten: Aroid *tamm-**) have plausible links with SWS ***eeet** and SES ***tammən** respectively.

Bender (2003: 260) listed 107 Proto-Aroid items, which are restricted to Aroid reflecting the relative independence of the Aroid lexicon. Of these around 10 items have some kind of relations with Surmic lexicon (see Annex I below: boat (canoe) (1), fire (10), mountain (45), grass (43), drink (8), person (21), snake (23), sand (49), sew (31) and hoe (13)). See also Annex II for the following 6 items: claw (59), burn (56), dog (74), seed (53), work/make (64), plant (51). The remaining 91 Proto-Aroid items do not have cognates in Omotic and possibly either they are innovations within Proto-Aroid or retentions of older forms. In any case, this is yet a strong lexical evidence against the Omotic theory and in favor of the independence of Aroid.

Fleming (1974) also gives 21 Omotic words with Afro-Asiatic cognates. Of these Afro-Asiatic cognates, Fleming claims that Proto-Aroid shares items such as *all*, *blood*, *eat*, *mouth*, *tooth* and *donkey*. But these cognates have been rejected by Theil (2012) who argued that the cognates presented by Fleming lack phonological and semantic correspondences in support of the cognates. Neither the phonological correspondences shown in this study support the cognates given in Fleming (1974). Rather the lexical reconstructions given in this study strongly favor the “Nilo-Saharan” theory that the Aroid group may be originally Nilo-Saharan with a strong Omotic influence.

8 Morphological reconstructions

8.1 Independent Pronouns

It has already been claimed that Aroid pronouns resemble Nilotic (Bender 2000, 2003). Bender (2000) has also compared Aroid pronouns with western Nilotic Nuer, Dinka and Anuak as well as Eastern Nilotic Teso and concluded that Aroid Pronouns are similar to Nilotic. Examine the following comparative data from Aroid languages:

	Aari	Hamar	Kara	Dime	Proposed Proto-Aroid forms	Proto-Omotic forms (Bender)
1sg	ʔinta	ʔinta	ʔinta	ʔate	*ʔinta	*in
2sg	ääna	ya	ya	yaai	*ya-	*ne
3sg (m)	nöö	kisi	nöh	nuu	*nöh/*nuu	*is/*isi
3sg (f)	naa	kosi	nää	naa	*naa	*is/*isi
3sg (n)	yī	yi	yi	---	*yi	---
1pl	wota	wosi	wota	wotu	*wot-	*nu
2pl	yata	yasi	yata	yasi	*yāt-	*int
3pl	keta	kisi	keta	kete	*ket-	*ist

Note that the **w-**, **y-**, **k-** prefixes in the plural forms are person markers typical to Nilotic. Aari second person singular **ääna** must be the result of loss of the initial consonant that is followed by compensatory lengthening, i.e. *hāna > āna > äāna; the third person singular masculine in Aari **nöö** must have passed through the same process, *nöh > nō > nöö (Yigezu 2005). Note that the Hamar third person singular and plural forms are identical in form and are akin to Nilo-Saharan forms as pointed out in Zaborski (2004: 181). See also Ehret (2001: 510) for the reconstruction of Nilo-Saharan *si ‘this’.

None of the proposed Proto-Aroid forms matches the Proto-Omotic forms given by Bender. The third person singular forms in Hamar (**kisi** and **kosi**) are assumed to be a loan from Nilotic by Bender (2000: 198). But within the synchronic grammar of Hamar the shortened forms of the independent pronouns usually appear before various dependent verbs as: **ki** in Aari, Hamar and Kara (for instance, in Hamar as in **ki-na** ‘his’; **ko-na** ‘her’ and **ke-na** ‘them’) and **kin** in Dime. The initial consonants function as a third person marker in all Aroid languages. The **k-**element recurs in various types of pronominal as well as deictic constructions in all Aroid languages.

Zaborski (2004: 181), on the other hand, rejects Bender’s (2000: 199 – 201) assumption that Aroid pronouns have been borrowed from “nearby Nilotic languages”, most likely the Nilotic Nyangatom, saying that these pronouns prove that Aroid languages are in fact Nilotic. He further contends that there is no evidence as far as pronouns are concerned that historically the ancestor of any Nilotic language had strongly influenced the Aroid languages.

The similarity of Aroid pronouns to that of Surmic is also interesting. Note that Proto-SWS form for first person singular is *aneetta while the Proto-SES form is *apne. The first person plural form in Proto-SWS is *aggetta while it is *agge in Proto-SES. The independent pronoun ‘I’ in Baale, for instance, is **anda** that must have undergone the process of syncope, as in *aneetta > *an-tta > **anda**, creating a word-medial pre-nasalized stop (Yigezu & Dimmendaal, 1998: 199). The fact that the Aroid independent pronouns are more similar to Surmic than to the rest of Omotic languages verifies the identity of independent pronouns in Aroid languages.

8.2 Verbal person suffixes

A typical feature of Hamar and Kara verb system is that there is no morphological marking for person in the verb form while Aari has person marking elements in the verb form. Fleming (1974: 86) notes that the reconstruction of Proto-Aroid verbal suffixes marking person rests heavily on Galila (a dialect of Aari) which is the only Aroid variety with proper paradigm of person marking inflections. Dime, Fleming states, “has enough left of an earlier paradigm to make it plausible”. Fleming’s reconstruction of verbal suffixes marking person is essentially based on a single member of the Aroid group, and his proposal has also been rejected by Theil (2012: 380) on the ground that no systematic phonological comparisons have been made between the grammatical morphemes in Omotic and other branches of Afro-Asiatic. As can be seen from Table 26 below it is hardly possible to make phonological correspondences between the verbal person marking elements within the Aroid group.

Table 26: Aroid verbal person suffixes

	Aari	Hamar	Kara	Dime	Proto-Aroid
1sg	..it	-----	-----	..t	
2sg	..ot	-----	-----	..n	
3sg	..aj	-----	-----	..n	
1pl	..et	-----	-----	..t	
2pl	∅	-----	-----	..n	
3pl	..ek	-----	-----	..n	

Concerning the verbal derivation, the only plausible cognate that can be reconstructed at the Proto-Aroid level is the causative marker **-si-*. This form has also been suggested by Fleming (1974: 84) as the Proto-Omotic form for the causative and also noted that it is “almost universal”. Hayward (2000: 93) also states that the “transitivizing/causative” is found in all the six families of Afro-Asiatic, as well as in Bantu (Niger-Congo) and reconstructed for Proto-Central Khoisan. Hence, the causative suffix may not be useful evidence as it is a widespread form found across language families. For all the other derivational markers such as the passive, reciprocal and reflexive it is difficult to find phonological correspondences due to the varied forms occurring the Aroid languages, except that Hamar and Kara has identical forms.

Table 27: Aroid verbal derivation markers

Verbal derivations					
	Aari	Hamar	Kara	Dime	Proto-Aroid
Passive	-rimʔ-ri	-ɕi-	-ɕi-	-int ⁱ	
Causative	-si	-si-	-si-	-isʔ-s	*-si-
Reciprocal	-mi	ki-	ki-	-sim	
Reflexive		-mal	-mal		

8.3 Case markers

Nominative case in all Aroid languages is not morphologically marked while accusative, genitive, dative, locative and instrumental cases are marked.

Table 28: Aroid case markers

	Aari	Hamar	Kara	Dime	Proto-Aroid
Nominative	∅	∅	∅	∅	*∅
Accusative	-im	-ta	-ta	-im	*-ta/*-im
Genitive	-ta	-sa/-n	-sa	-ko	*-sa
Dative	-ken/-kenta	-dar	-dar	-in	
Locative	-war	-ra	-ra/-dar	-se/-o	*-ra
Instrumental	-ka/-k	-nka	-ka	-ka	*-ka

Fleming (1974: 85) suggests that Dime and Galia, have /-m/ accusative and /-n/ dative-benefactive especially on pronouns. In our data, in Hamar and Kara, the accusative is marked by **-ta**, **-n** and **-ɕan**; while **-ta** attaches only to nouns and the other two forms can be suffixed to both nouns and pronouns. In any case, the only clear phonological correspondence attested is the instrumental case suffix.

9 Conclusions

The phonological and lexical reconstructions presented in this study strongly suggest that Aroid languages have a Nilo-Saharan origin with strong Omotic influence. Particularly, the lexical correspondences between Proto-Aroid and Proto-Surmic forms are concrete attestations to the Nilo-Saharan theory. As Zaborski (2004: 173) noted “Aroid ...pronouns have not been borrowed from Nilo-Saharan as alleged by Bender – they are a good proof of the Nilo-Saharan origin of these languages”. Although the morphological reconstruction is sketchy and awaits an in depth analysis of the topics under discussion, the reconstruction of independent pronouns has clearly shown that the system is evidently not an Omotic one but closely akin to Nilo-Saharan pronominal system.

References

- Anderson, Torben. 1987. The phonemic system of Agar Dinka. *Journal of African Languages and Linguistics* 9: 1–27.
- Bender, M. Lionel. 1988. Proto-Omotic phonology and lexicon. In Marianne Bechhauus-Gerst & F. Serzisko (eds.), *Cushitic-Omotic: Papers from the International Symposium on Cushitic and Omotic Languages*, pp. 121–159. Hamburg: Buske.
- Bender, M. Lionel. 2000. *Comparative Morphology of the Omotic Languages*. Munich: Lincom.
- Bender, M. Lionel. 2003. *Omotic Lexicon and Phonology*. Carbondale: Southern Illinois University.
- Campbell, Lyle. 1997. *American Indian Languages: The Historical Linguistics of Native America*. Oxford Studies in Anthropological Linguistics, 4. Oxford: Oxford University Press.
- Dimmendaal, Gerrit J. 1988. The lexical reconstruction of Proto-Nilotic: a first reconnaissance. *Afrikanistische Arbeitspapiere* 16: 5–67.
- Ehret, Christopher. 2001. *A Historical-Comparative Reconstruction of Nilo-Saharan* (SUGIA Supplements, 12). Cologne: Köppe.
- Fleming, Harold C. 1969. The classification of West Cushitic within Hamito-Semitic. In Daniel McCall, Norman Bennett & Jeffrey Butler (eds.), *Eastern African History*, pp. 3–27. New York/Washington/ London: Praeger.
- Fleming, Harold C. 1974. Omotic as an Afroasiatic family. *Studies in African Linguistics*, 5: 81–94.
- Fleming, Harold C. 1976. Omotic overview. In M. Lionel Bender (ed.), *The Non-Semitic Languages of Ethiopia*, pp. 299–323. East Lansing: African Studies Centre, Michigan State University.
- Fleming, Harold C. 1988. Proto-South-Omotic or Proto-Somotic consonant phonemes: stage one. In M. Bechhaus-Gerst & F. Serzisko (eds.), *Cushitic-Omotic: Papers from the International Symposium on Cushitic and Omotic Languages*, pp. 163–175. Hamburg: Buske.
- Fleming, Harold C. 1990. A grammatical sketch of Dime (Dime-Af). In Richard J. Hayward (ed.), *Omotic Language Studies*, pp. 494–583. London: School of Oriental and African Studies.
- Hayward, Richard J. 1990. Notes on the Aari language. In Richard J. Hayward (ed.), *Omotic Language Studies*, pp. 425–493. London: School of Oriental and African Studies.
- Hayward, Richard. 2000. Afroasiatic. In Bernd Heine & Derek Nurse (eds.), *African Languages: An Introduction*, pp. 74–98. Cambridge: Cambridge University Press.
- Keefer, James, Aurelia Keefer & Charles V. Taylor. 1976. Anyuak. In M. Lionel Bender (ed.), *Language in Ethiopia*, pp. 164–70. London: Oxford University Press.
- Lamberti, Marcello. 1991. Cushitic and its classification. *Anthropos* 86: 552–561.

- Lamberti, Marcello. 1993. The Ari-Benna group and its classification. *Studi Italiani di Linguistica Teorica e Applicata* 22.1: 39–87.
- Lusted, Marie. 1976. Anywa. In M. Lionel Bender (ed.), *The Non-Semitic Languages of Ethiopia*, pp. 495–512. East Lansing: African Studies Center, Michigan State University.
- Lydall, Jean. 1976. Hamar. In M. Lionel Bender (ed.), *The Non-Semitic Languages of Ethiopia*, pp. 393–438. East Lansing: African Studies Center, Michigan State University.
- Malou, Job. 1988. *Dinka Vowel System*. Publications in Linguistics, 82. Dallas: The Summer Institute of Linguistics, University of Texas at Arlington.
- Menuta, Fekede. 2011. *Aari Language: A Comparative Grammar*. Addis Ababa: Bole Printing Enterprise.
- Reh, Mechthild. 1996. *Anywa Language: Description and Internal Reconstructions*. (Nilo-Saharan, 11). Cologne: Köppe.
- Seid, Melkineh. 2011. Documentation of Aari language. MA thesis, Department of Linguistics, Addis Ababa University.
- Seyoum, Mulugeta. 2007. Some notes on the personal and demonstrative pronouns in Dime. In Rainer Voigt (ed.), “*From Beyond the Mediterranean*”: *Akten des 7. internationalen Semitohamitistenkongresses*, pp. 253–264. Aachen: Shaker.
- Seyoum, Mulugeta. 2008. A grammar of Dime. PhD thesis, Netherlands Graduate School of Linguistics.
- Theil, Rolf. 2012. Omotic. In Lutz Edzard (ed.), *Semitic and Afroasiatic: Challenges and Opportunities*, pp. 369–384. Wiesbaden: Harrassowitz.
- Thomason, Sarah G. 1980. Morphological instability, with and without language contact. In Jacek Fisiak (ed.), *Historical Morphology*, pp. 359–372. The Hague: Mouton.
- Tsuge, Yoichi. 1996. On the consonant correspondences of South Omotic languages. In Shun Sato & Eisei Kurimoto (eds.), *Essays in Northeast African Studies* (Senri Ethnological Studies 43), pp. 163–188. Osaka: National Museum of Ethnology.
- Yigezu, Moges. 1995. The Nuer vowel system. *Journal African Languages and Linguistics* 16: 157–170.
- Yigezu, Moges & Gerrit J. Dimmendaal. 1998. Notes on Baale. In Gerrit J. Dimmendaal & Marco Last (eds.), *Surmic Languages and Cultures* (Nilo-Saharan, 13), pp. 273–317. Cologne: Köppe.
- Yigezu, Moges. 2001. A comparative study of the phonetics and phonology of Surmic languages. PhD thesis, Universite Libre de Bruxelles, Belgium.
- Yigezu, Moges. 2005. Some observations on the pronouns of Hamar: a comparative-historical perspective. In Yohachi Tsuge (ed.), *Cushitic-Omotic Studies*, pp. 113–132. Japan: Kanazawa.
- Yigezu, Moges. 2007. The vowel system of Kara from a historical-comparative perspective. In Rainer Voigt (ed.), “*From Beyond the Mediterranean*”: *Akten des 7. internationalen Semitohamitistenkongresses*, pp. 245–252. Aachen: Shaker.
- Zaborski, Andrzej. 2004. West Cushitic: a genetic reality? *Lingua Posnaniensis* 46: 173–186.

Annex I: Aroid vs. Surmic Lexical Correspondences

(Abb.: SWS = Southwest Surmic; SES = Southeast Surmic)

No.	Gloss	Aroid	Surmic
1	boat (canoe)	Aari: gonga ; Hamar: gonggala ; Kara: gaagi ; Dime: gongu [Aroid:*gongul]	Baale: gonggól Chai: gonggúl ; Mursi: gonggúl ; Me'en: gónggul ; Koegu: gagi [SES:*gonggul]
2	call (v)	Aari: el- ; Hamar: el- ; Kara: gis- ; Dime: el- [Aroid:*el-]	Mursi: elley- ; Chai: eley-
3	calabash	Aari: gusi ; Kara: gusi [Aroid:*gusi]	Baale: gussi ; Chai: gussi ; Mursi: gussi ; Koegu: gufa
4	cloth	Aari: abla [afla]; Hamar: afala ; Kara: apala ; Dime: apal [Aroid:*apala]	Me'en: apala Koegu: apala
5	cotton	Aari: putta ; Hamar: puddo ; Kara: putta ; Dime: futtu [Aroid:*putta]	Didinga: guttu ; Me'en: puta Koegu: putta
6	crocodile	Aari: gurgur ; Hamar: gurgur ; Kara: gurgur ; Dime: guuru [Aroid:*gurgur]	Tennet: oggol ; Me'en: gur Majang: ugul [SWS:*kugul]
7	divide (v)	Aari: kAʃ- ; Hamar: kajf- ; Kara: kAʃ- ; Dime: kif- [Aroid:*kAʃ-]	Koegu: kajf- ; Majang: kaas-
8	drink (v)	Aari: wutʃ'- ; Hamar: wutʃ'- ; Kara: wutʃ'- ; Dime: wutʃ'- [Aroid:*wutʃ'-]	Narim: wod- ; Murle: wót- [SWS:*kod-]
9	egg	Aari: muk'a- ; Hamar: wuk'a ; ɓula ; Kara: muk'a- ; Dime: molu [Aroid:*muk'-; mol-]	Tennet: ɓoro ; Narim: ɓooro ; Didinga: ɓooro ; Baale: ɓura ; Chai: ɓurray ; Mursi: ɓurray ; Me'en: mulatʃ ; Koegu: moogu
10	fire	Aari: nōha ; Hamar: nuu ; Kara: nɔ ; Dime: nuunu [Aroid:*nōha ~ nuu]	Tennet: guo ; Narim: guo ; Me'en: gu ; Koegu: go [SWS:*gʷuo; SES:*go]
11	father	Aari: baabi ; Hamar: imba- ; Kara: ʔimba ; Dime: baabe [Aroid:*baab] - symbolic?	Baale: babba ; Chai: babbá ; Mursi: dadda ; Me'en: baba ; Koegu: baaba ; Majang: baabe [Surmic:*babba]
12	hand/arm	Aari: ʔaani ; Hamar: aan ; Kara: ʔan ; Dime: ʔani [Aroid:*ʔan]	Baale: aji ; Majang: ʔari
13	hoe	Aari: gofa ; Hamar: gaita ; Kara: gaita ; Dime: gait [Aroid:*gaita]	Baale: gaiða ; Chai: gaisa ; Me'en: gaita ; Koegu: gaita ; [SES:*gaita]

14	hold	Aari: yAd-; Hamar: yAd-; Kara: yAd-; Dime: yiid- [Aroid:*yAd-]	Majang: yodi-
15	I	Aari: ?ita; Hamar: inta; Kara: ?inta; Dime: ?ate [Aroid:*?inta]	Narim: anetta; Tennet: anítta; Murle: anetta; Baale: andá; Koegu: áán [SWS:*anetta]
16	liver	Aari: turi; Hamar: tiraḅo; Kara: toraḅo; Dime: tayte [Aroid:*to/ur-]	Chai: tārrá; Mursi: tárrá; Me'en: tara [SES:*tarra]
17	navel	Aari: gula?/guubi; Hamar: gulf-; Kara: gungusi; Dime: gufa [Aroid:*gulf-]	Baale: guldú; Chai: gūldú; Koegu: guóbo
18	nine	Aari: sakal; Hamar: sAl; Kara: sAl; Dime: ?okalʃ [Aroid:*sakal]	Chai: sakal; Mursi: sākkāl; Me'en: saal; Koegu: sál [SES:*sakkal]
19	people	Aari: eedna; Hamar: eena; Kara: eena; Dime: nyap [Aroid:*eedna]	Narim: eetta; Didinga: éetta; Baale: eeḅa [SWS:*eetta]
20	person	Aari: eed; Hamar: eedi; Kara: eedi; Dime: nye [Aroid:*eed]	Tennet: éét; Narim: éét; Didinga: éét; Murle: éét; Baale: éé [SWS:*eet]
21	show	Aari: daw-; Hamar: day-; Kara: daw-; Dime: ?eχs- [Aroid:*daw-]	Mursi: dól-; Me'en: dól-; Koegu: dual- [SES:*dol-]
22	snake	Aari: guni; Hamar: guni; Kara: guno; Dime: gunu [Aroid:*gun-]	Mursi: kuno; Me'en: kono; Chai: kono [SES:*kono]
23	sow	Aari: bAk-; Hamar: dok'-; Kara: dok'-; Dime: deeχ- [Aroid:*dok'-]	Murle: dók-; Baale: dug-; Chai: dug-; Mursi: dug-; Me'en: du?- [SES:*dug-]
24	spit	Aari: tAs'-; Hamar: pas'-; Kara: maamm-; Dime: tuf- [Aroid:*tAs'-; pas'-] cf. Bender: *t'up-	Murle: tór-; Baale: tór; Chai: tuh-; Mursi: tuh-; Me'en: tuy- [SWS:*tor-; SES:**tuh-]
25	ten	Aari: tamma; Hamar: tAbi; Kara: tAbi; Dime: tamme [Aroid:*tamm-; tAb-]	Chai: tammən; Mursi: tammən; Me'en: tammən; Koegu: tammən [SES:*tammən]
26	tobacco	Aari: dampa; Hamar: tampo; Kara: dampo; Dime: dampu [Aroid:*dampo]	Baale: dambu; Chai: dambu; Mursi: dammu; Me'en: dampu; Koegu: dampo [SES:*dampu]
27	urinate	Aari: jaan-; Hamar: ʃAn-; Kara: jaan-; Dime: jan- [Aroid:*jaan-]	Chai: tʃorō-; Mursi: tʃoro-; Me'en: tʃoro-; Koegu: tʃoʔo- [SES:*ʃoro]

28	urine	Aari: ʃaan ; Hamar: ʃAAN ; Kara: ʃano ; Dime: ʃaan [Aroid:*ʃaan]	Murle: zɔɔ ; Baale: sɔɔ ; Chai: ʃörrá ; Mursi: ʃörrá ; Me'en: ʃórwá ; Koegu: ʃoʔoté [SES:*ʃoro]
29	road	Aari: gɔgi ; Kara: goyi [Aroid:*gɔgi]	Mursi: gɔrɔ ; Me'en: gɔrɔ ; Koegu: guar [SES:*gɔrɔ]
31	sew	Aari: dʒak- ; Hamar: dʒaag- ; Kara: dʒaag- ; Dime: dʒig- [Aroid:*dʒaag-]	Mursi: tʃeg- ; Chai: tʃey- ; Me'en: tʃik- ; Koegu: tʃik- [SES:*tʃvk-]
32	rainy season	Aari: bargi ; Hamar: bargi ; Kara: bargi ; Dime: birgi- [Aroid:*bargi]	Me'en: bergu ; Majang: ban̄gi
33	chicken	Aari: baatʃa ; Hamar: baaʃa ; Kara: baatʃa ; Dime: koyz [Aroid:*baatʃa ~ koyz] ?	Koegu: baatʃa ; Baale: kowu ; Chai: kowi
34	split	Aari: zal-/ʔaʔ- ; Hamar: ʃat- ; Kara: pals- ; Dime: p'al- [Aroid:*p'al-] ?	Chai: bel- ; Me'en: bel- [SES:*pel-]
35	mother	Aari: ʔindi ; Hamar: inda ; Kara: inda ; Dime: ʔindɪ [Aroid:*ʔind-]	Majang: indii
36	mouth	Aari: ʔapa ; Hamar: apo ; Kara: ʔapo ; Dime: ʔape [Aroid:*ʔap-]	Majang: átó
37	jump	Dime: ʔotl-	Majang: utal-
38	mountain	Dime: edi	Majang: eden
39	pierce	Hamar: tors- ; Kara: tors- ; Dime: tussu- [Aroid:*tors-]	Me'en: tuuna ?
40	grass	Dime: haayi	Me'en: haay
41	thing	Dime: lók	Chai: lɔgɔ ; Mursi: lɔgɔ
42	beard	Dime: s'ire	Mursi: tʃɔre ; Me'en: tʃ'ore
43	bite (of bee)	Dime: ʔank'-	Tennet: ʔák- Murle: ʔak-
44	mud	Dime: təl	Baale: ḏór Chai: ḏór ; Mursi: ḏór
45	rope	Dime: gontu	Me'en: gontiit
46	stick	Dime: kalitʃi	Narim: kalli
47	stomach	Dime: tʃ'olay	Mursi: tʃolloy
48	sand	Aari: ʃaami ; Hamar: ʃami ; Kara: ʔazo ; Dime: ʃayɪ [Aroid:*ʃaam-]	Me'en: ʃebá
49	shoot	Aari: kas'- ; Hamar: kas'- ; Kara: kas'- ; Dime: gɪs'- [Aroid:*kas'-]	Baale: kaḏ-
50	hot	Aari: sol (Galila dial.); Dime: suulum	Mursi: ʃalo

Annex II: Aroid vs. Koegu Lexical Correspondences

51	plant	Aari: kor-; Hamar: kor-; Kara: kɔr-; Dime: kor- [Aroid:*kor-]	Koegu: kor-
52	animal	Aari: dabɪ; Hamar: dabɪ; Kara: dʌbi-; [Aroid:*dabɪ-]	Koegu: daban
53	seed	Aari: βeta; Hamar: βenta; Kara: βenta; Dime: mɪʃit [Aroid:*βenta]	Koegu: βenta
54	untie	Aari: bul-; Hamar: bul-; Kara: bul-; Dime: bul- [Aroid:*bul-]	Koegu: bul-
55	blow	Aari: fog-; Hamar: fug-; Kara: pug-; [Aroid:*pug-]	Koegu: pugu-
56	burn	Aari: ʔats-; Hamar: ati-; Kara: ʔat-; Dime: ʔats- [Aroid:*ʔats-] ?	Koegu: at-
57	charcoal	Aari: tʃ'ilʃi; Dime: qɪtʃ'i; Kara: tʃ'ilʃi [Aroid:*tʃ'ilʃ-]	Koegu: k'iltʃ'i
58	chicken	Aari: baatʃa; Hamar: baafa; Kara: baatʃa [Aroid:*baatʃ-] ?	Koegu: baatʃa
59	claw	Aari: guʃma; fukuma (for animals); Kara: guʃo [Aroid:*guʃ-]	Koegu: fukuma
60	darkness	Aari: ɗommi	Koegu: ɗurum
61	defecate	Aari: fan	Koegu: tʃa-
62	donkey	Aari: okolɪ; Hamar: ukuli; Kara: ukuli [Aroid:*ukul-] ?	Koegu: ukkuli
63	hat	Aari: kɔβ; Dime: k'op'u; Kara: k'obe [Aroid:*kɔβ-]	Koegu: k'obe
64	make	Aari: ʔakʃ-; Hamar: aʃk- Kara: aʃk- [Aroid:*ʔaʃk-]	Koegu: aʃk-
65	push	Aari: haŋk'-; Kara: yik'-	Koegu: hik'-
66	smoke (n)	Aari: tʃ'uba; Hamar: tʃ'uba; Kara: tʃ'ubi [Aroid:*tʃ'ub-]	Koegu: tʃubui
67	arrow	Hamar: bankar; Kara: bankar [Aroid:* bankar] ?	Koegu: bakár
68	chief	Hamar: bitɪ; Kara: bitɪ [Aroid:*bitɪ]	Koegu: bitti
69	dry season	Hamar: bona; Kara: bɔna [Aroid:*bona]	Koegu: bona
70	open	Hamar: bul-; Kara: bul- [Aroid:*bul-]	Koegu: bul-

71	salt	Hamar: sok'o ; Kara: sok'o [Aroid:*sok'o]	Koegu: sok'o
72	bark of a tree	Dime: ʔufumu	Koegu: ʔuom
73	bury	Dime: duuk- ; Kara: duk- [Aroid:*duuk-] ?	Koegu: tutuk-
74	dog	Dime: kem	Koegu: kiani (cf. Gamo)
75	six	Aari: la ; Hamar: läh ; Kara: läh ; Dime: lahe [Aroid:*läh] ?	Koegu: leh