Diversity in African languages

Selected papers from the 46th Annual Conference on African Linguistics

Edited by

Doris L. Payne

Sara Pacchiarotti

Mokaya Bosire

language
science
press

Contemporary African Linguistics

Editors: Lee Bickmore, Akinbiyi Akinlabi

In this series:

1. Payne, Doris L., Sara Pacchiarotti & Mokaya Bosire. Diversity in African languages: Selected papers from the 46th Annual Conference on African Linguistics.

Diversity in African languages

Selected papers from the 46th Annual Conference on African Linguistics

Edited by

Doris L. Payne

Sara Pacchiarotti

Mokaya Bosire

language
science
press

Freie Universität Berlin

Contents

Contents

Preface

The 28 papers in this volume were presented at the 46th Annual Conference on African Linguistics, held at the University of Oregon, March 26-28, 2015. The conference featured nearly 120 papers, addressing phonetics, phonology, morphosyntax, semantics, information structure, sociolinguistics, and historical linguistics, as well as a workshop on Luhya languages of East Africa. The selected papers in this volume represent the broad range of topics and language families that characterized the conference, as well as the maturation of research in African linguistics.

The volume editors thank those whose generous support made the conference and this volume possible: the Association of Contemporary African Linguistics (ACAL); and the University of Oregon Linguistics Department, GLOSS (University of Oregon Graduate Linguistics Students organization), College of Arts and Sciences, Division of Equity and Inclusion, Global Oregon, Office of International Affairs, African Studies Program, American English Institute, Yamada Language Center, and Ford Alumni Center. We are grateful to many graduate and undergraduate students, administrative staff, and friends who supported the 46th ACAL conference, notably Ariel Andersen, Linda Campbell, Thiago Castro, Richard Griscom, Manuel Otero, Sanna Parikka, Hugh Paterson, Rebecca Paterson, Jaime Peña, Matthew Stave, and Amos Teo.

Special thanks go to the many reviewers who helped us select and improve the papers in this volume: Enoch Aboh, Colleen Ahland, Michael Ahland, George Akanlig-Pare, Sean Allison, Greg Anderson, Peter Avery, Lee Bickmore, Rosey Billington, Fabio Bonfim Duarte, Leston Buell, Mike Cahill, Robert Carlson, Rod Casali, Tucker Childs, Thera Crane, Nick Danis, Michael Diercks, Gerritt Dimmendaal, Laura Downing, Helen Eaton, Ines Fiedler, Volker Gast, Eitan Grossman, Abbie Hangtan, Martin Haspelmath, Larry Hyman, Kyle Jerro, Jason Kandybowicz, Boniface Kawasha, Mena Lafkioui, Marten Lutz, Michael Marlo, Laura McPherson, Jonathan Moodie, Irina Monich, Edith Moravcsik, Leonard Muaka, Carol Myers-Scotton, Deo Ngonyani, Steve Nicolle, Emmanuel Nikiema, Morgan Nilsson, Derek Nurse, David Odden, Knut Olawski, Paul Olejarczuk, E. Kweku Osam, Thomas Payne, Mary Pearce, Eric Pederson, David Peterson, Kent Rasmussen, Nicholas Rolle, Sharon Rose, Phil Rudd, Josephat Rugemalira, Ken Safir, Ron Schaefer, Anne Schwarz, Frank Seidel, Shahar Shirtz, Irina Shport, Galen Sibanda, Sylvester R. Simango, Keith Snider, Tim Stirtz, Anne Storch, Amos Teo, Harold Torrence, Mark van de Velde, Jenneke van der Wal, Valentin Vydrin, Aggrey Wasike, and Malte Zimmerman. We are also grateful for the support and advice of the Contemporary African Linguistics series editors Lee Bickmore and Akinbiyi Akinlabi, the expertise of Hannah Sande and Emily Clem, and the patient, meticulous, and high-quality technical work of Sebastian Nordhoff and colleagues from Language Science Press.

Part I

(Possible) Niger-Congo

Subpart I-A

Mande

Chapter 1

Tone features revisited: Evidence from Seenku

Laura McPherson

Dartmouth College

Recently, authors such as Hyman (2010) and Clements, Michaud & Patin (2010) have argued that African tone is better modeled with tonal primitives (e.g. H, M, L) than with tonal features. This paper reopens the question with novel data from Seenku, a four-tone Mande language of Burkina Faso (X, L, H, S). I argue that the features [±upper, ±raised] provide a unified analysis of several tonal phenomena, including plural formation, tonal neutralizations, and verbal alternations. First, I argue that plural formation is a case of featural affixation, with a plural suffix [+raised] deriving [-upper,+raised] L from singular X, while underlying [+upper,-raised] H shifts to S. In terms of tonotactics, the two middle tones are treated differently in nouns: [+upper,-raised] is not allowed word-finally and is always followed by X, while derived [-upper,+raised] is allowed. Further evidence for tone features is found in the verbal domain. First, the distinction between S and H in verbs is often neutralized, to S for transitive verbs and H for intransitive verbs. I analyze these neutralizations as default [+raised] assignment to underlying [+upper] verbs in the transitive and [-raised] assignment in the intransitive. In the perfective, S-toned transitive verbs are realized as H while X-toned verbs remain unchanged. A featural account derives this result with the affixation of perfective [-raised]. Finally, complicated argument-head tonal alternations may be more naturally explained under a featural approach. In sum, this paper presents a case where tonal features show an analytic advantage over tonal primitives, suggesting that the debate is not yet over.

1 Introduction

Segments are widely accepted in phonology to consist of phonological features. These features encode parameters such as place ([labial], [coronal], [dorsal]), voicing ([voice]), nasality ([nasal]), or manner ([sonorant], [continuant], [delayed release]). For tone, the situation is much less clear. Unlike segments, tone relies on only one phonetic parameter, f0 (barring secondary features like phonation), which is inherently scalar rather than binary.

Nevertheless, numerous feature systems for tone have been proposed in the literature. Wang (1967) proposed a seven feature system for tone, including three height features ([high], [central], [mid]), and four contour features ([contour], [rising], [falling], [convex]). Later systems abandoned featural specification for contour tones, opting to view contours as sequences of tone levels instead. The most widely accepted systems take four level tones as the base, which can be achieved with two binary features. Yip (1980) proposed a so-called Register feature [±upper], dividing tonal space into two halves further subdivided by a second feature (sometimes called a Tone feature) [±high]. This latter feature was renamed [raised] by Pulleyblank (1986). Other authors such as Clements (1983), Snider (1990), and Hyman (1993) use unary features, [h/l] for Register and [H/L] for Tone, but the resulting systems function in largely the same way.

Despite numerous proposals for both African languages and tone languages elsewhere, recent work has cast doubt on the use of features for tone. Hyman (2010), for instance, points out problems for M tones in featural systems, including featural ambiguity in a three-tone language and the lack of a natural class for M tones in a four-tone language. Clements, Michaud & Patin (2010) echo these criticisms, pointing to the lack of clear natural classes defined by tone features and to the lack of support for assimilation or dissimilation patterns driven by tonal features. For these reasons, both sets of authors suggest that at least African tone is better modeled in autosegmental terms with simple levels (L, M, H, etc.).

This paper has two main goals. The first is to describe the tone system of southern Seenku, a relatively undescribed Mande language of Burkina Faso. The second is to re-open the debate on the featural underpinnings of tone. I will argue that a two-feature system aids in the analysis of Seenku, drawing evidence from plural formation, transitive/intransitive tonal neutralization, perfective formation, and argument-head tonal alternations found in inalienable possession and certain O+V constructions.

The paper is structured as follows: §2 provides background information on Seenku and data sources, and in §3, I give a brief description of Seenku lexical tone. The core arguments for tonal features are given in §4, where I address plural formation (§4.1), transitive/intransitive tonal neutralization (§4.2), perfective formation (§4.3), and argument-head tonal alternations (§4.4). §5 considers alternative analyses and §6 concludes.

2 Language and data

Seenku (ISO 639-3 [sos]) is a Mande language of the Samogo subfamily spoken in southwestern Burkina Faso. It has two main dialects, each named after the main village where the dialect is spoken: northern Seenku (Timiku, literally 'language of Karangasso') with 5,000 speakers and southern Seenku (Gbeneku, literally 'language of Bouendé') with 12,000 speakers (Lewis, Simons & Fennig 2015). The former was the subject of Prost's (1971) *Éléments de Sembla*, a short grammar sketch and lexicon, but the latter has received very little scholarly attention apart from Congo's (2013) Master's Thesis on aspects of the phonology. Since 2013, I have undertaken fieldwork on the southern dialect; all data in this paper are drawn from my field notes.

Like most Mande languages, Seenku shows S Aux O V X word order, where "X" can be occupied by an indirect object, PP, negation, or adverb. Morphologically, it is largely isolating.

3 Sketch of the tone system

Seenku is a four-tone language, with tonal primitives X (extra-L), L, H, and S (extra-H), though with a few exceptions the underlying tonal inventory can be reduced to three (X, H, S); as we will see below, L is commonly the result of plural formation, where it contrasts with singular X, but is rarely found lexically. Minimal sets contrasting even these three underlying levels are remarkably difficult to find, given an apparent tono-tactic restriction on H in word-final position in nouns and many tonal neutralizations found in verbs (see §4.2 and §4.3). In pronouns, we find the following (near) minimal pairs for X vs. H and H vs. S, respectively:[1]

(1) a. à '3sg'
         á '2sg'
     b. mó '1sg'
         mǐ '1pl'

If we include the repair for noun-final H, i.e. epenthesis of X, the following (near) minimal sets can be identified:[2]

(2) Tonal minimal sets contrasting X, H, and S

	X	H(X)	S
a.	kyɛ̀(n)	kyɛ̂(n)	kɛ̋
	'peanut'	'breast'	'fat'
b.	tsὺ	tsû	sű
	'thatch'	'hippo'	'antelope'

Underlying L is limited in the current dataset to one numeral, *nɔ̀* 'five', and a couple of adverbs, *kɔ̀rɔ̀* 'yesterday' and *màa* 'again'. Given the limited nature of numeral and adverbial vocabulary, minimal pairs are not available, but L forms a near minimal pair with X in numerals: *nɔ̀* 'five' vs. *nǎa* 'four', and the f0 of L on 'five' is lower than that

---

[1] Tonal transcription represents X with double grave <à>, L with grave <à>, H with acute <á>, and S with double acute <á̋>. Tone marking for the whole syllable is otherwise only marked on the first vowel, e.g. *bɛ̀ɛ* 'pig' is a long level X. The most common contour tones are HX and XS, represented by circumflex <â> and hacek <ǎ>, respectively. The less common falling contour SX is represented by umlaut <ä>. All of these contour tones are likewise marked on the first vowel only, representing the fact that tone is a property of the syllable rather than the mora, and to maintain identity in tone marking between short and long vowels. All other contour tones are only found through processes of vowel coalescence, and in this case only, each component of the contour is marked on one half of the long vowel, e.g. HS <áá̋>.

[2] In these examples, the <n> in parentheses represents a floating nasal, usually unpronounced in isolation but realized on the following word in connected speech (either by nasalizing a sonorant or prenasalizing a stop).

of H on *sóen* 'one' when pronounced side-by-side, showing that these two middle tones are indeed phonetically distinct.

Contour tones are very common in Seenku, particularly HX (illustrated above) and XS, found on both heavy and light syllables. This distribution suggests that the tone-bearing unit (TBU) in Seenku is the syllable rather than the mora. An example of a minimal pair contrasting HX and XS is given in (3):

(3) kûĩ 'néré seeds'
    kǔĩ 'grass sp.'

Of these, HX is the more common contour, found on all syntactic categories; XS, in contrast, is particularly common on auxiliaries and adjectives, the latter of which may be grammatically assigned.

The other attested underlying contour is the tritonal sequence XHX, as in *däâ* 'basket hanger'.

Other contours are created morphologically or phonologically, as illustrated in the following examples:

(4) Other contour tones and how they are created

Tone	Example	Gloss	Created by…
SX	nï̄ɔ	'has eaten'	Perfect formation
XSX	nä̌ä	'has come'	Perfect formation
HS	móő	'1sg past'	Past tense formation
HL	móò	'1sg genitive'	Genitive formation
XH	ɛ̏ɛ́	'3sg genitive'	Genitive formation

In terms of tone rules, Seenku displays downstep and contour tone simplification, though the domains of these processes and their potential implications for a system of tone features are still under investigation.

4 Evidence for tone features

I propose that tone in Seenku is characterized by the following binary features, using the Pulleyblank (1986) feature system:

(5) Seenku tone features

	X	L	H	S
[upper]	-	-	+	+
[raised]	-	+	-	+

The two binary features produce four potential tone levels, all of which are represented in Seenku.[3] As stated above, L is seldom part of an underlying specification and is instead usually derived by the addition of grammatical tone features (featural affixes).

Evidence for the utility of tone features over tonal primitives is drawn from four sources: plural formation, transitive/intransitive verb tone, perfective formation, and tonal interactions between pronominal arguments of nouns and verbs (inalienable possession and O+V constructions). This featural specification for Seenku tone responds to some of the criticisms of tone features, including providing evidence for natural classes and for assimilation and dissimilation patterns.

4.1 Plural formation

The first piece of evidence for tone features in Seenku comes from nominal plural formation. Here, we see a tone raising process (in addition to vocalic changes that I will not address here), raising X to L and H to S; underlying S in the singular remains S in the plural, since there are no further tone levels to raise to. For example:

(6)   Plural tone raising

    a.   X   →   L
         bɛ̏ɛ   →   bɛ̀ɛ   'pig(s)'

    b.   H(X) →   S
         bî   →   bǐ   'goat(s)'

    c.   S   →   S
         sű   →   sűi   'antelope(s)'

I argue that tone raising is driven by a featural affix [+raised] (McCarthy 1983; Lieber 1987; Wiese 1994; Akinlabi 1996; Wolf 2007, etc.). The addition of [+raised] to an X tone ([-upper, -raised]) yields L ([-upper, +raised]). The addition of [+raised] to H ([+upper, -raised]) yields S ([+upper, +raised]). Finally, the addition of [+raised] to S tone yields no audible difference, since it is already specified as [+raised]. In short, between the singular and the plural, all four possible tonal specifications are attested.

In tonally complex nouns, only the final tone is altered in the plural, suggesting that the plural [+raised] is a suffix. We see this effect in (7a), where the final S absorbs [+raised], leaving the preceding X unaffected, and in (7b), where H(X) raises to S without effecting the preceding X in the contour tone:

---

[3] In the early stages of work, I analyzed the language as a three-tone language, which meant there was ambiguity in the featural specification of the M tone. Nevertheless, differing tonotactic restrictions for erstwhile "lexical M" (now H) vs. the "derived M" (now L) supported this four-way featural distinction. Further fieldwork revealed that the two supposedly M tones are in fact phonetically distinct, with the tone derived by plural formation (L) lower than that found underlyingly (H). The discovery of a small number of underlying L tones corroborate the decision to treat Seenku as a four-tone language, despite the majority of lexical contrasts being created with only three levels. In other words, it is thanks to a featural analysis that I became attuned to the possibility of four distinct levels.

(7)   a. jɔ̀ŋwǎ → jɔ̀ŋwɛ̋ 'cat(s)'
      b. dǎâ → dɛ̀ɛ̋ 'basket hanger(s)'

Looking at (7b), we can see that the feature [+raised] targets H of the tritone XHX contour, not X. This fact is explained if the underlying form is XH, with the final X tone added only if plural formation fails to apply. As mentioned in the last section, there is a systematic absence of level H-toned singular nouns in the lexicon:

(8)   Singular level-tone melodies

	Singular	Plural	Gloss
X	bɛ̀ɛ	bɛ̀ɛ	'pig(s)'
H	–	–	–
S	sű	sűi	'antelope(s)'

Instead, we find an abundance of HX contours that become S in the plural, just as we would expect of a H tone. Examples include:

(9)   /H/ singular → S plural

bî	bǐ	'goat(s)'
kâ	kɛ̋	'yam(s)'
sâ(n)	sɛ̋	'rabbit(s)'
gɔ̂ɔ	gɔ̋ɛɛ	'wood(s)'

If X were part of the underlying representation, then [+raised] would dock to X, creating HL (e.g. *bî* → **bí* ' 'goat(s)').[4] This supports an underlying representation /bí/, which fills in the systematic gap in singular level tone melodies. Anytime a H tone finds itself in noun-final position, an X tone is epenthesized as a repair.[5] If we assume morphology occurs before phonology, then the plural of H nouns would carry a [+raised] feature that alleviates the need for such an epenthetic X:

(10)

UR	/bí$_{sg}$/	/bí$_{pl}$/	
Morphology	—	bǐ	(Addition of [+raised])
Phonotactics	bî	—	
SR	[bî]	[bǐ]	

From a constraint-based perspective, *[+upper, -raised]# would be satisfied in the plural by docking the [+raised] feature and deriving an S tone, whereas in the singular where no such feature is available, X epenthesis is the optimal strategy. In contrast, all other tones (X, L, S) are level, showing that the phonotactic ban is specifically on the featural specification [+upper, -raised].

In sum, plural formation provides evidence for all four tone levels in Seenku, motivated by a single featural affix [+raised].

---

[4] HL is never found on a light syllable in Seenku, so no single diacritic is employed to represent it, the circumflex already being used to represent HX.

[5] This is either a case of lexical class-specific tonotactics or X is itself morphological, perhaps encoding singular (though not on S nouns). I leave this question to future work.

4.2 Transitive and intransitive verbal tone

The next piece of evidence for features comes from transitive and intransitive verbal tone. On the surface, most verb stems show only a two-way tone contrast, with neutralization of S and H (though as we will see later, there is a contrast between these two underlyingly).[6] For transitive verbs, S and H verb stems neutralize to S tone, as highlighted in Table 1, with a dummy 3sG object *ǎ*.

Table 1: Transitive verb stems

Underlying tone	Surface form	Gloss
X	*ǎ sǎ̀*	'buy it'
	ǎ gyɔ̌̀	'grill it'
	ǎ fɔ̌	'uproot it'
H	***ǎ kǔ̃ɔ***	'bite it'
	ǎ sɔ̌ɔ	'sell it'
	ǎ gǎa	'pull it'
S	***ǎ bǎ̃***	'hit it'
	ǎ dzǐ̃	'put it'
	ǎ nǐ̃ɔ	'eat it'

Both a lexically H-toned stem like /sɔ́ɔ/ 'sell' and a lexically S-toned stem like /níɔ/ 'eat' have S tone on the surface in constructions where verbal tone is not perturbed by either aspect (see §4.3) or the presence of an object in the irrealis mood (see §4.4), namely the progressive and the immediate past, to be expanded upon below.

For intransitive verbs, S and H verb stems neutralize to H. However, since the underlying distinction between the two only emerges in the presence of a direct object, it is impossible to determine the underlying tone of intransitive verbs in most cases. (11) gives surface forms only:

(11)   Intransitive verb stems

		Surface form	Gloss		Surface form	Gloss
a. X		kà	'go'	b. H	sɔ́	'arrive'
		nà	'come'		tsí̃	'jump'
		kì	'die'		sú	'get up'
		kwàa	'farm'		gyɔ́ɔ	'return'

The neutralization of S and H is a dynamic process that results in alternations. For instance, an ambivalent stem *gyəra* 'spill' surfaces as *gyɔ̌rǎ* when used transitively and

---

[6] Recent fieldwork has unearthed some irregular verbs that do not follow these tonal patterns, including a few S-toned intransitives, but the majority of verbs do undergo the neutralizations described here.

gyɔ́rá when used intransitively. X-toned verb stems, on the other hand, always surface with X. Thus, intransitive *kwǎa* 'farm' is still X-toned *kwǎ* when used transitively.[7]

I analyze these patterns as the result of morphological neutralization rules targeting [+upper] tones and shifting their registers to either [+raised] for transitive or [-raised] for intransitive verbs:

(12)   a.   [+upper] → [+raised] when V$_{transitive}$
       b.   [+upper] → [-raised] when V$_{intransitive}$

For (12a), the change to [+raised] in [+upper, +raised] S verbs is vacuous, since they already carry this specification, while the change to [+raised] in [+upper, -raised] H verbs results in a [+upper, +raised] S tone. Similarly, for (12b), the change to [-raised] in [+upper, -raised] H verbs is vacuous, but this same change in [+upper, +raised] S verbs results in [+upper, -raised] H tone. In both cases, the tonal distinction is neutralized. We know that these neutralizations are the result of more restricted rules and not general floating featural morphemes (e.g. [+raised] for transitive, [-raised] for intransitive), since the concatenation of [+raised] with an X verb in the transitive would raise it to L, a change we do not see.

These featural alternations are most likely related to another tonal change we see in the same realis verb forms: In the periphrastic progressive and immediate past, both of which employ the verb stem followed by the postposition *nɛ*, transitive verbs are followed by an S tone and intransitive verbs are followed by an X tone. This tone is most often realized solely on the postposition, leaving transitive verbs followed by S-toned *nɛ̋* and intransitive verbs by X-toned *nɛ̏*, but intransitive verb stems with a long vowel allow the X tone to dock, creating a HX contour on H-toned stems.[8] For example:

(13)   Addition of transitive S and intransitive X in postpositional forms

	a.	Transitive		b.	Intransitive	
		à sɔ̋ɔ nɛ̋	'sell it'		kà nɛ̏	'go'
		à sǎ̰ nɛ̋	'buy it'		sá nɛ̏	'cry'
		à kpɔ̰̋ɔ̰ nɛ̋	'sew it'		gyɔ̂ɔ nɛ̏	'return'

While it is tempting to view the neutralizations as the synchronic result of partial assimilation to the added tone, this analysis is not supported by the data. First, we might expect under this view that X-toned transitive verb stems might also raise, which they do not; explanations along the line of parasitic harmony (Cole & Trigo 1988) would hold only of transitive verbs (where H raises to adjacent S), and not of intransitive verbs where it is the maximally different tone (S) that lowers. Second, and more importantly, certain

---

[7] The vowel length distinction may be due to an assimilated antipassive suffix in the intransitive form.

[8] Presumably, the same docking principles would hold true for transitive verbs as well, but the only audible contour that could be created is an XS rising tone, and Seenku displays progressive tonal absorption (Hyman & Schuh 1974) when a rising tone is followed by an S tone. This results in simplification back to X. Evidence that a rising tone is in fact created on X-toned transitive verbs can be found in the xylophone surrogate language (McPherson 2016), where contour simplification is not encoded; musicians play these verbs as rising tones.

idiosyncratic verbs like *ŋáǎ nɛ̃́* 'yawn' display a HS contour on the surface before an S-toned postposition, showing that there is no reason such contours could not be created by the addition of S to H-toned transitive verbs. In other words, raising of H to S before another S is not automatic. Instead, I argue that the tonal neutralizations shown above may be the grammaticized result of phonetic raising or lowering due to the following tone but cannot be analyzed purely on these grounds from a synchronic perspective.

Summarizing this section, the use of tonal features allows us to clearly capture patterns of neutralization in two ways. First, the feature [+upper] defines a natural class of tones in Seenku, namely S and H, that is affected by the rules of neutralization. Second, the neutralization itself can be explained in featural terms as the change to [+raised] in transitive verbs and to [-raised] in intransitive verbs.

4.3 Perfective formation

We find another case of featural affixation in the perfective, though unlike the plural, its effects are only audible in one type of verb, namely transitive S-toned verbs. In the transitive, we see a lowering of surface S-toned verb stems to H; X-toned verb stems show no change:

(14) Perfective forms of transitive verbs

		Progressive	Perfective	Gloss
a.	S	à sɔ̌ɔ nɛ̃́	à sɔ́ɔ	'sell it'
		à níɔ nɛ̃́	à níɔ	'eat it'
		à bɑ̌́ nɛ̃́	à bɑ́	'hit it'
b.	L	à sɑ̀̌ nɛ̃́	à sɑ̀̌	'buy it'
		à ɡyɔ̀̌ nɛ̃́	à ɡyɔ̀̌	'grill it'
		à fɔ̌ nɛ̃́	à fɔ̌	'uproot it'

Intransitive verbs, like X transitive verbs, show no tonal change in the perfective (apart from the last case, where the absence of the X tone and postposition allows the verb stem 'return' to surface as level H):

(15) Perfective forms of intransitive verbs

		Progressive	Perfective	Gloss
a.	L	kà nɛ̀	kà	'go'
		nà nɛ̀	nà	'come'
		kǐ nɛ̀	kǐ	'die'
		kwǎa nɛ̀	kwǎa	'farm'
b.	H	sɔ́ nɛ̀	sɔ́	'arrive'
		tsɪ́ nɛ̀	tsɪ́	'jump'
		sú nɛ̀	sú	'get up'
		ɡyɔ́ɔ nɛ̀	ɡyɔ́ɔ	'return'

I analyze the perfective as a featural affix [-raised]. Added to [+upper, +raised] S, this affix derives [+upper, -raised] H. Added to [+upper, -raised] H or [-upper, -raised] X, it has no audible effect. Because of this, it is indeterminable whether perfective formation

applies before or after tonal neutralizations; the resulting forms would be the same either way.

Thus, the existence of featural affixes in Seenku is corroborated by data from the perfective. Without tone features, we would have to propose an arbitrary rule of S-tone lowering in the perfective, whereas the affixation of [-raised] explains both the cases where the affix is audible and those where it is not.

4.4 Alternations with pronominal internal arguments

The final argument for tone features is more speculative and is made based on a series of complicated tonal alternations that arise between either a verbal or nominal head and its internal argument (direct object or possessor) when that argument is pronominal. The contexts in which these alternations take place are summarized in (16):

(16)   a.   A pronominally possessed **inalienable** noun.
       b.   A transitive verb in **irrealis mood** (future, imperative, habitual) with a
            pronominal object.

When the verb is realis (including when it is perfective), it does not interact tonally with the object.

Before we turn to the alternations, the inventory of Seenku pronouns is summarized in Table 2.

Table 2: Seenku pronouns

Person	Singular	Plural
1	*ń/mó*	*mǐ*
2	*á (wó)*	*í (yó kwɛ̌)*
3	*ǎ wǒ*	*ǐ/kwɛ̌*

Where there are slashes in Table 2, the form on the left is the basic (unfocused) form and the form on the right is the focused form; similarly, elements in parentheses are added after pronouns when they are focused. As we can see, all three basic tones (X, H, S) are attested on pronouns, while L is absent.

When a noun or verb takes a pronoun as its argument, it follows the pronoun and displays tonal alternations depending on both its own underlying tone and on the tone of the pronoun. It is here that we see the three-way tonal contrast on verb stems emerge despite its neutralization in other contexts. Table 3 summarizes the alternations, which are the same for both nouns and verbs. The body of the table displays the resulting tonal form of the head noun or verb based on its underlying form (top row) following pronouns of varying tonal forms (leftmost column).

One pattern is clear and straightforward: all head tones are neutralized to S tone after an S-final pronoun (1PL, focused 2PL and 3PL). The pattern with X-final pronouns (3SG,

Table 3: Summary of tonal alternations

Final tone of pronoun	Underlying tone of head		
	X	H	S
X	X	X	H
H	S	X	X
S	S	S	S

unfocused 3PL) is likewise fairly consistent: it triggers lowering on the head, with S becoming H, H becoming X, and X remaining X (the opposite pattern of that seen in the plural). The pattern with H-final pronouns (1SG, 2SG, unfocused 2PL) is the most challenging: there is polarity of underlying X and S tone, and H lowers to X.

How can tone features help us make sense of this situation?

First, it is important to note that after non-pronominal (nominal) arguments, the tone of the head always takes on the final tone of its argument; that is, it is always neutralizing.[9] We can see this in (17), where the same verb *sǎ* 'buy' takes an X-, L-, and S-toned object:

(17) a. bɛ̀ɛ sǎ 'buy a pig!'
    b. bɛ̀ɛ sɑ̀ 'buy pigs!'
    c. bǐ sɑ̋ 'buy goats!'

The examples in (18) show the neutralization of verbal lexical tone after a X-toned object:

(18) a. /X/ bɛ̀ɛ sǎ 'buy a pig!'
    b. /H/ bɛ̀ɛ sɔ̀ɔ 'sell a pig!'
    c. /S/ bɛ̀ɛ bɑ̀ 'hit a pig!'

Multiple analyses are possible for the distinction between nouns and pronouns as the argument of the head. One possibility relies on underspecification: tone spreading or copying only takes place after fully specified tones. Under this approach, S-toned pronouns would be necessarily fully specified as [+upper, +raised], whereas X- and H-toned pronouns would be missing one of the tonal features. The problem with this approach is that in cases with complex arguments (compound nouns or possessive phrases as the object of a verb), the verb undergoes the same tonal alternations as it would after a pronoun, despite the complex argument arguably having full tonal specification. A second possibility is that differences result from phrasing or domain assignment: arguments

[9] This is reminiscent of tonal compounding processes elsewhere in Mande, e.g. *compacité tonale* in Bambara (Creissels 1978; 1988; 1992; Dumestre 1984; Green 2010), tonal compounding in Susu (Grégoire 1978; Green, Anderson & Obeng 2013), as well as in other Western Mande languages (deZeeuw 1979).

and heads seek to form a unified, possibly binary domain, in which the initial element is tonally dominant. Nouns are prosodically stronger than pronouns and are able to fully fill this role, while pronouns cannot overpower the tone of the head. The exception is with S-toned pronouns, which tend to be CV in shape rather than V or N (impossible syllable shapes in nouns); the combination of the "strong tone" S and their nominal shape allows them to behave like a regular noun. The similarities in tonal effects between pronominal and complex arguments would come from the fact that in both cases the argument is a non-ideal tonal head: in the former case, it is too light, and in the latter case, it is too heavy.

I will leave these explanations for future work and offer here only some preliminary thoughts on why we find the particular tonal alternations described in Table 3 as opposed to any others. I will show that tone features may indeed hold the key.

Whether we fully specify X-toned pronouns as [-upper, -raised] or underspecify them as [-raised] alone, alternations with these pronouns follow straightforwardly from the spread of the [-raised] feature as in (19). The pronoun and the verb here are linked with the feature [-raised], which causes the verb to lower from S to H.

(19) Example of [-raised] spreading

  [-upper]   [+upper]
     | |
     a   bã
     |_ _ _ _ _ _ _ _ ⌐
  [-raised]   [+raised]
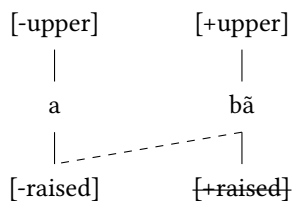
  'hit him!'

H-toned pronouns are trickier, since they neither consistently raise nor consistently lower the head. However, I argue that these alternations, too, can be understood as a preference for spreading [-raised] of [+upper, -raised] combined with an OCP effect against adjacent H tones. When a H-toned pronoun precedes an S head, [-raised] spreads. However, this creates a sequence of two H tones; to resolve this sequence, [+upper] on the head dissimilates to [-upper], yielding an X tone:

(20) Example of [-raised] spreading from H to S

  [+upper]   [+upper] ⟶ [-upper]
     | |
    mo   bã
     |_ _ _ _ _ _ _ _ ⌐
  [-raised]   [+raised]
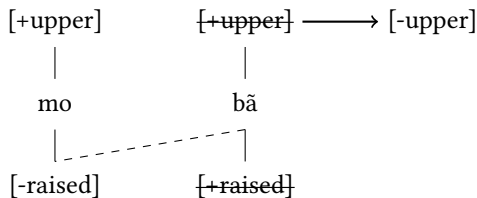
  'hit me!'

With an underlyingly H-toned head, the OCP restriction comes into effect right away, triggering the same repair of dissimilating [+upper] to [-upper]. This results in an X tone once again, as in (21).

(21)   Example of [+upper] dissimilating in a sequence of two Hs

[+upper]   ~~[+upper]~~ ⟶ [-upper]
   | |
   mo   sɔː
   | |
[-raised]   [-raised]

   'sell me!'
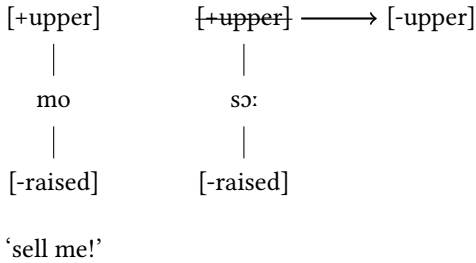
Taking H to be a middle tone in Seenku, this kind of M-tone dissimilation has support in other African languages, such as Leggbo (Paster 2003). Alternatively, the dissimilation could be driven by an OCP constraint (e.g. McCarthy 1986) on [+upper] rather than on the sequence of two Hs specifically. However, a similar dissimilation pattern is arguably at work with X-toned heads. Here, rather than spreading [-raised] onto a tone already designated as [-raised], the non-homophonous [+upper] spreads instead. This creates once again a sequence of two H tones, and here it is [-raised] that dissimilates on the head to [+raised], creating an S tone:

(22)   Example of [+upper] spreading from H to X

[+upper]   ~~[-upper]~~
   ⌐ – – – – – – – – ⌐|
   mo   sã̃
   | |
[-raised]   ~~[-raised]~~ ⟶ [+raised]

   'buy me!'
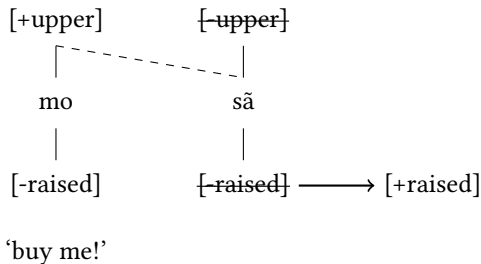
These results can be unified by the following informally conceived constraints: 1. The argument and the head should be linked tonally, preferably by [-raised]. 2. This linking should be of a non-homophonous tonal feature. 3. Two H tones may not follow one another (or, there is an OCP constraint on [+upper] and [-raised]).

Pronoun-head configurations are still under investigation in Seenku, but the use of tone features brings us closer to understanding how we can get cases of partial assimilation (only one feature spreads rather than both) and why we get the particular changes that we do. We further find promising cases of featural dissimilation of [+upper] and [-raised], driven either by the features themselves or by the larger tonal complex (H) in which they are found.

5 Feature-less alternatives

If Hyman (2010) and Clements, Michaud & Patin (2010) are correct that tone should not be modeled with features, then alternative approaches must be found for Seenku. In this section, I briefly consider two possibilities, showing where each is successful and where it falls short.

5.1 Tonal primitives (X, L, H, S)

Under this approach, tones are indivisible elements. The tonal neutralizations found in transitive and intransitive verb tone could be explained by differential phonotactics or reduced inventories: transitive verbs only allow S and X and intransitive verbs only allow H and X.

However, the other tonal effects do not emerge as easily. First, we might try to explain the tone raising chain shift in the plural with the affixation of S, where X+S yields L and H+S yields S, but seeing as the language allows contour tones, there is no principled reason why these tone mergers should take place; the situation is the same for the lowering effect of the perfective. Second, there is no natural explanation for the restricted nature of L. Under a two feature system, four categories are automatically available, and L is derived naturally by grammatically manipulating these features. Under a tonal primitive analysis, this fourth category would need to be specifically posited and then restricted to (mostly) derived environments. Finally, the tonal alternations found between pronouns and their lexical head would require even more stipulated tone rules without the availability of features.

5.2 Scalar tone

A more promising alternative is the use of a scale for tone, shown in (23):

(23)   Seenku scalar tone

X	L	H	S
1	2	3	4

Raising in plural formation would be easily accounted for in this system by a rule of [+1] ($1 \rightarrow 2$, $3 \rightarrow 4$). Perfective formation would be a rule of [-1], but only in transitive verbs and only after the neutralization rules that raise /H/ to S. As above, this would require that we stipulate reduced tonal inventories for transitive and intransitive verbs. The lowering effect with X-toned pronouns, however, would be problematic, since a rule of [-1] would create a L tone from a H tone rather than the attested X. Further, the tonal effects with H-toned pronouns do not follow naturally, since tone level 1 raises to 4, while both 4 and 3 lower to 1.

Thus, like the tonal primitive approach, this approach faces a number of difficulties that are more elegantly solved under the featural account.

6 Conclusion

To sum up, tone features have been rejected based in part on the following criticisms:

1. No evidence for tonal natural classes, as in segmental phonology.
2. No evidence for assimilation or dissimilation patterns.
3. They give rise to ambiguity in M tones for three-tone languages.
4. Everyone employs them differently; there is no accepted standard.

In this paper, I have argued that some of these criticisms need to be reconsidered. First and foremost, the existence of tonal features allows us to posit featural affixes for tone, which allow for the elegant analysis of a number of phenomena in Seenku. The data thus far give evidence of a [+raised] featural affix marking plural, a [-raised] featural affix marking perfective, and possible [+raised] and [-raised] marking transitive and intransitive, respectively, on underlying [+upper] verb stems. On this point, the existence of a tone rule or featural affix targeting only [+upper] verb stems responds directly to criticsm 1: Seenku provides evidence for tonal natural classes.[10]

In response to criticism 2, we may find evidence for both assimilations and dissimilations in pronoun/head alternations. Specifically, there may be an OCP effect of [-raised] and [+upper] sequences, triggering dissimilation on the second feature, while feature spreading of [-raised] could be viewed as an assimilatory process.

Criticism 3 is a bit difficult to assess, given Seenku's four-tone nature. As I have shown, however, the vast majority of lexical contrasts are produced with only three tones, with the second "middle tone" restricted to contexts derived by manipulating tone features of the other three. I take this as evidence that the availability of four categories under a feature system may actually be a boon not only for analysis but also for the development of a four-tone system out of what was presumably a system with fewer contrasts historically (as evidenced by related Mande languages).

Finally, criticism 4 is a valid point: there is no accepted standard for tonal features or their geometry. However, I do not view this as reason to abandon the hypothesis. Either we simply have not examined enough languages yet in light of tonal features to reach a consensus, or, as Odden (2010) argues, feature systems need not be phonetically-grounded and universal. They may be deduced by speakers from the learning data, leading to different systems and analyses in different cases.

If languages like Seenku continue to respond to these criticisms, then it may not be time to close the book on tonal features just yet.

---

[10] It is interesting to note that all of the featural affixation required for Seenku involves the feature [raised]. In this light, we might take [raised] to be the register feature, as in Snider (1990; 1998), and thus think of Seenku morphological processes as manipulating register. Future work will explore this topic further, focusing on the relationship between downstep (an attested process in Seenku phonology) and the featurally defined tones presented in this paper.

Laura McPherson

Acknowledgments

References

Akinlabi, Akinbiyi. 1996. Featural affixation. *Journal of Linguistics* 32. 239–289.

Clements, George N. 1983. The hierarchical representation of tone features. In Ivan Dihoff (ed.), *Current approaches to African linguistics*, 145–176. Dordrecht: Foris.

Clements, George N., Alexis Michaud & Cédric Patin. 2010. Do we need tone features? In John A. Goldsmith, Elizabeth Hume & Leo Wetzels (eds.), *Tones and features: Phonetic and phonological perspectives*, vol. 107 (Studies in Generative Grammar), 3–24. Berlin: De Gruyter Mouton.

Cole, Jennifer & Loren Trigo. 1988. Parasitic harmony. In Harry van der Hulst & Neil Smith (eds.), *Features, segmental structure and harmony processes*, 19–38. Dordrecht: Foris.

Congo, Rasmane. 2013. *L'esquisse d'analyse phonologique du sembla (parler de Bouendé)*. Ouagadougou: Université de Ouagadougou. (Master's thesis.)

Creissels, Denis. 1978. A propos de la tonologie du bambara: Realisations tonales, système tonal et la modalité nominal 'défini'. *Afrique et langage* 9. 5–70.

Creissels, Denis. 1988. *Etudes manding no. 1 - Le système tonal du bambara standard*. Grenoble, France: Centre de Linguistique Africaine de l'Université des Langues et Lettres de Grenoble.

Creissels, Denis. 1992. Tonologie du bambara: Bilan et perspectives. *Mandenkan* 24. 1–45.

deZeeuw, Peter Howard. 1979. *Western Mande compound tone rules*. East Lansing, MI.

Dumestre, Gerard. 1984. Contribution à l'étude des tons du bambara. *Mandenkan* 7. 63–74.

Green, Christopher R. 2010. *Prosodic phonology in Bamana (Bambara): Syllable complexity, metrical structure, and tone*. Bloomington, IN: Indiana University. (Doctoral dissertation.)

Green, Christopher R., Jonathan Anderson & Samuel Obeng. 2013. Interacting tonal processes in Susu. *Mandenkan* 50. 61–84.

Grégoire, Claire. 1978. La tonalité des mots composés et celle du présent progressif en soso. *Journée d'études langue et linguistique manding* 1. 75–86.

Hyman, Larry M. 1993. Register tones and tonal geometry. In Harry van der Hulst & Keith Snider (eds.), *The phonology of tone: The representation of tonal register*, 75–108. Berlin: Mouton de Gruyter.

Hyman, Larry M. 2010. Do tones have features? In John A. Goldsmith, Elizabeth Hume & Leo Wetzels (eds.), *Tones and features: Phonetic and phonological perspectives*, vol. 107 (Studies in Generative Grammar), 50–80. Berlin: De Gruyter Mouton.

Hyman, Larry M. & Russell G. Schuh. 1974. Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry* 5(1). 81–115.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the world*. 18th edn. Dallas: SIL International. http://www.ethnologue.com.

Lieber, Rochelle. 1987. *An integrated theory of autosegmental processes*. Albany: SUNY Press.

McCarthy, John. 1983. Consonantal morphology in the Chaha verb. In *Proceedings of the West Coast Conference on Formal Linguistics*, vol. 2, 176–188.

McCarthy, John. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17. 207–263.

McPherson, Laura. 2016. Seenku phonology in the Sembla xylophone surrogate language. (Paper presented at the 90th annual meeting of the Linguistic Society of America.)

Odden, David. 2010. Features impinging on tone. In John A. Goldsmith, Elizabeth Hume & Leo Wetzels (eds.), *Tones and features: Phonetic and phonological perspectives*, vol. 107 (Studies in Generative Grammar), 81–107. Berlin: De Gruyter Mouton.

Paster, Mary. 2003. Tone specification in Leggbo. In John Mugabe (ed.), *Linguistic typology and representation in African languages*, vol. 5 (Trends in African Linguistics), 139–150. Trenton, NJ: Africa World Press Inc.

Prost, André. 1971. *Eléments de sembla: Phonologie - grammaire - lexique*. Lyon: Afrique et Langage.

Pulleyblank, Douglas. 1986. *Tone in lexical phonology* (Studies in Natural Language and Linguistic Theory). Dordrecht: D. Reidel.

Snider, Keith. 1990. Tonal upstep in Krachi: Evidence for a register tier. *Language* 66(3). 453–474.

Snider, Keith. 1998. Phonetic realization of downstep in Bimoba. *Phonology* 15(1). 77–101.

Wang, William S-Y. 1967. Phonological features of tone. *International Journal of American Linguistics* 33(2). 93–105.

Wiese, Richard. 1994. Phonological vs. Morphological rules: On German umlaut and ablaut. *Journal of Linguistics* 32. 315–404.

Wolf, Matthew. 2007. For an autosegmental theory of mutation. In Leah Bateman, Michael O'Keefe, Ehren Reilly & Adam Werle (eds.), *Papers in Optimality Theory III* (University of Massachusetts Occasional Working Papers in Linguistics 32), 315–404. Amherst, MA: Graduate Linguistics Student Association (GLSA), University of Massachusetts.

Yip, Moira. 1980. *The tonal phonology of Chinese*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation.)

Subpart I-B

Gur

Chapter 2

Intonation and emotions in Kɔnni: A preliminary study

Michael Cahill
SIL International

One of the paralinguistic functions of intonation is the use of gradient changes of pitch and duration to indicate emotional states of the speaker. This study examines the difference in pitch of Kɔnni native speakers' speech which accompanies several different emotions. A neutral utterance was compared to the same sentence uttered as if the speakers were surprised, bored, angry, "contemptuous", and wanted to emphasize the sentence. Base pitch level, pitch range and overall duration of the sentences were measured and compared to the neutral statement. The results of this study are compatible with those found in other languages, and add to the knowledge of how tone languages are able to express paralinguistic intonation in a systematic way.

1 Introduction

The term *intonation* does not have a universally agreed on definition. Some researchers either explicitly define it in terms of pitch alone or seem to assume such a definition (Lieberman 1967; most papers in Bolinger 1972; Gussenhoven 2004). As Hirst & Di Cristo (1998: 3) note, the terms *intonation* and *prosody* have often been used interchangeably in the literature. These authors spend significant time discussing the ambiguities of various terms, distinguishing INTONATION PROPER, which deals with pitch, from the broader concept PROSODY, which also includes intensity and quantity. Ladd (2008: 4) gives a useful definition which we will assume here: "the use of *suprasegmental* phonetic features to convey 'postlexical' or *sentence-level* pragmatic meanings in a *linguistically structured way*" (his emphasis). Though pitch will be the most common measure referred to in this paper, duration will also be examined.

A particular instance of intonation can be either STRUCTURAL or PARALINGUISTIC (Gussenhoven 2004; Ladd, Scherer & Silverman 1986; Ladd 2008). Structural intonation is categorical and phonological, indicating linguistic boundaries or morphosyntactic functions. Paralinguistic intonation involves gradient phonetic values of pitch, as well as duration and intensity, often indicating emotions and attitudes. Kɔnni has cases of each

of these, and a broader range of intonational patterns is examined in Cahill (2016), but this paper's focus will be on paralinguistic intonation.

Intonation in tone languages has not been studied nearly as much as in non-tonal languages, probably on the assumption that lexical and grammatical tone would override any pitch differences attributable to intonation.[1] The papers on various languages in Hirst & Di Cristo's (1998) survey include a few tone languages (Thai, Vietnamese, and Beijing Mandarin), and the papers on Thai and Vietnamese have some detailed remarks on the topic of this paper: how emotional states influence intonation. However, the overall literature on emotions and intonation in tone languages is still sparse. Green's (2009) work titled "Prosody and Intonation in Non-Bantu Niger-Congo Languages: An Annotated Bibliography" includes 125 works on individual languages, of which only five deal at all with intonation, and none with the emotion/intonation issues addressed here.

Tone languages often use particles or other morphosyntactic strategies, rather than pitch, to indicate grammatical functions which are indicated by pitch in other languages. Focus will serve to illustrate this difference. Narrow focus in English is indicated intonationally, with pitch as a major component: "You DROVE to the store"(that is, you didn't WALK…). Cruttenden (1997: 73) notes that tone languages are less likely to use intonation as a means of focus than non-tone languages, and several recent studies affirm this. In Awutu (Lomotey 2014), a deliberate attempt to have speakers focus on one part of a sentence produced almost no pitch variation. Schwarz (2009) writes that Kɔnni and two closely related languages (Buli and Dagbani) use only morphosyntactic structure to indicate focus. Harley (2009) notes five strategies that Tuwili uses for focus. Four are morphosyntactic, though there is also a pitch-accent strategy. Even in the non-tonal African language Wolof, focus is marked by morphosyntactic means, not by intonation (Rialland & Robert 2001). Fiedler & Jannedy (2013) show that Ewe's most reliable prosodic cue for focus is not F0, but duration of the focused element. In light of this, the natural question that arises is how intonation can function in a tone language, since both intonation and tone affect pitch. Cruttenden (1997: 9–10) notes four ways that tone languages may implement what he terms *superimposed intonation*:

- The pitch level of the whole utterance may be raised or lowered.
- The range of pitch may be narrower or wider.
- The normal downdrift of a sentence may be suspended.
- The final tone of the utterance may be modified.

The first two of these, and sometimes the others, are paralinguistic expressions of intonation, and these are more common than structural intonation in African languages. We will see the first two of these – change in pitch level and change in pitch range – exemplified in the present study on Kɔnni on the interaction of pitch with states of emotions in Kɔnni.

Kɔnni ([kma], Gur family) has two underlying tones, High (H) and Low (L). These may combine as rising (LH) and falling (HL) contours on single syllables. A second

---

[1] An exception to this is a a volume entirely devoted to intonation in African languages, Downing & Rialland (2016). This includes my broader review of several intonation patterns in Kɔnni.

High may be downstepped from the preceding High (HꜜH). This sequence can appear on adjacent syllables or on a single syllable as a second type of falling tone. The tone-bearing unit is the syllable, to which one or two autosegmental tones may associate. A detailed presentation and analysis of Kɔnni tone can be found in Cahill (2007).

Cahill (2012) gives an examination of Kɔnni polar question intonation. This phenomenon is structural: the tone of the final syllable of the utterance is lowered in one of several distinct ways by adding tonal autosegments. For a final noun ending in High tone, either a L autosegment is added, resulting in a falling HL tone as in Table 1, example (a); or LH autosegments are added, resulting in a falling HꜜH tone as in example (b). Which pattern applies appears to be a lexical choice. If the final noun ends in a Low tone, HLH autosegments are added, in effect raising the tone before it is lowered, as in example (c). The final vowel of the syllable is also categorically lengthened.[2]

Table 1: Statements with corresponding polar questions

	Kɔnni	gloss
a.	*ù sìé gìlìnsìèlé*	's/he is dancing gilinsiele dance'
	sìé gìlìnsìèléè	'is s/he dancing gilinsiele dance?'
b.	*ὸ ŋmìà gúúmꜜbú*	's/he is rolling the rope'
	ὸ ŋmìà gúúmꜜbúꜜú	'is s/he is rolling the rope?'
c.	*ὸ dàwá níígè*	's/he has bought a cow'
	ὸ dàwá nííꜜgéꜜé	'has s/he bought a cow?'

Sometimes polar questions are the response given when the speaker is asked to act surprised, as we will see in some situations in this paper.

2 Methodology

The data for this study was gathered by Mr. Konlan Kpeebi of the Ghana Institute of Linguistics, Literacy, and Bible Translation (GILLBT). It was a small part of a broader project which gathered data on several aspects of Kɔnni intonation (Cahill 2016). I provided detailed instructions but was not personally present for the data gathering. Kpeebi recorded Mr. Naaza Solomon Dintigi and Mr. Mumuni Salifu Barnabas, both men in their 20s and native speakers of Kɔnni. This was done in a recording studio in Tamale, Ghana. Specifics of the recording hardware are not available, but the recording quality was free of roosters and other outside noises so frequently encountered in field recording situations, and the quality was more than adequate for pitch and duration analysis. I am extremely grateful to all of them for their input and expertise.

---

[2] Word order in Kɔnni is, like English, SVO, so the Kɔnni and their English translations here can be matched essentially word for word.

These two Kɔnni speakers were verbally given a natural Kɔnni sentence, and told to first say it normally (termed the *neutral* intonation here), then to repeat the same sentence, saying it as if they were experiencing various emotional states. Instructions were given in English, in which the speakers are both fluent. They repeated each utterance three times. Solomon produced one sentence with its emotional variants, and Salifu produced that sentence as well as six additional sentences with emotional variants, seven in all.

Table 2: Sentences produced by speakers

Both speakers		ʊ̀ dìgìwó nyʊ́à	's/he has cooked yams'
Salifu only	a.	ʊ̀ gàwá ˈnyʊ́ŋ	's/he has gone to market'
	b.	ʊ̀ dààwá kpááŋ	's/he has bought oil'
	c.	ù dìgìwó gílà	's/he has cooked eggs'
	d.	ù chʊ̀ŋwá gílà	's/he has fried eggs'
	e.	ù dùùwó ˈsááŋ	's/he has eaten TZ (porridge)'
	f.	ʊ̀ chɔ̀gìsìwó ˈbólíŋ	's/he has fetched fire'

The emotional states chosen were as if the speakers were *surprised*, *bored*, *angry*, *contemptuous*, and finally, *emphatically* (emphasizing that the statement is what really happened). Studies on emotions and intonation have covered a very wide and inconsistent list of emotions, even including irony and admiration (Đỗ Thế Dũng, Trần Thiên Hương & Boulakia 1998: 402). All the emotions in this study (plus several others) were included in the study on Thai by Luksaneeyanawin (1998), and several other studies in Hirst & Di Cristo (1998) included emphasis. Surprise, anger, boredom, and emphasis were all mentioned by Ladd (2008) as emotions that have been the subject of intonation studies. Considering my previous background in Ghana, as well as these fairly common prior mentions of these emotions in intonation studies, the choice of particular emotions in this study were a reasonably practical subset of possible states to elicit.

The response sentences had the same word order and lexical and grammatical tones as the input neutral sentences, with one partial exception. The surprise response often resulted in the speakers' producing a polar question. That is, 'S/he went to market,' expressed with surprise, became 'S/he went to market?' These are discussed somewhat separately in this study.

To test the phonetic variation of pitch in the utterances, the pitch range and base pitch level were measured.[3] This was done by measuring the frequency at two positions in each utterance: the initial Low tone of the sentence, and the first High tone in the

---

[3] All recordings were analyzed using SIL's Speech Analyzer program, available as free download at http://www.sil.org/resources/software_fonts/speech-analyzer.

sentence.[4] The initial Low is labeled as the base level, and the difference between this and the first High is the range. All the input utterances of this study had a Low-toned pronoun sentence-initially, and a High-toned verb suffix two or three syllables later, as exemplified in Figure 1. The syllables which were measured are bolded and underlined.
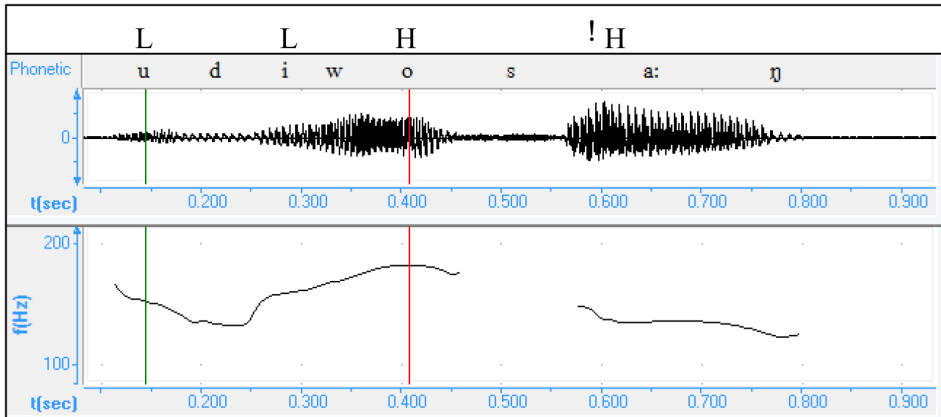
Figure 1: **_ù_** _dìi-_**_wó_** _⸍sááŋ_ 's/he ate TZ' (a type of porridge)

The frequency, in Hertz [5] is read directly off a part of the graph not included in Figure 1. The frequency was read at either the stable part of the vowel or, lacking a flat portion of the frequency, at the midpoint of the vowel. The *base pitch level* in Figure 1 is the frequency at the left cursor, i.e. at the Low toned [ù]. The *pitch range* is the difference between this Low and the High of the [wó] syllable at the right cursor.

Duration has also been found relevant in studies of other African languages (Hyman & Monaka 2011; Fiedler & Jannedy 2013), even when pitch is not directly involved. So the duration of the entire sentence was also measured, the distance between two cursors again being read directly off a part of the graph not included in Figure 1.

Regular and systematic differences were found between the neutral form of the utterance and the various emotional states for which data was gathered. We turn now to these.

---

[4] An alternative would be to measure the maximal pitch range, that is, the highest and lowest pitch in the sentence. This was not done because of downdrift. As is common in African languages, there is a continual downdrift of High tones after a Low, so that in a H-L-H-L-H-L-H sequence the last H is considerably lower than the first H, and in a longer sentence the last High may even be a lower pitch than the initial Low tone. Downdrift is also the reason why an average pitch was not taken across the sentence; the longer the sentence, the lower the average pitch.

[5] Reading the data in semitones is also an option in Speech Analyzer, and would be reported if there had been a mixed gender sample.

3 Results

All the numbers reported in the following tables are averages of three utterances of that particular sentence. Frequencies are reported in Hertz (Hz), and duration is reported in milliseconds (ms). Other abbreviations in the tables are:

- EXP= expanded range (a larger L-H difference than the neutral)
- CONT= contracted range (a smaller L-H difference than the neutral)
- "l" means that it's only slightly more of that quality, noticeable but with perhaps marginal significance.

As noted before, the requested *surprise* intonation often elicited a polar question as the response. In terms of structural vs. paralinguistic intonation, the polar question exhibits both. As briefly mentioned above and detailed in Cahill (2012), a polar question in Kɔnni is not only phonetically raised in pitch (paralinguistic), but is analyzable phonologically in terms of autosegments added to the neutral sentence (structural), and has one of several varieties of a falling tone on the final syllable. That syllable is categorically lengthened, and this accounts for the total duration of the *surprise* intonation being lengthened in all the measurements to follow (thus the label "longer" rather than "slower").

We begin with a detailed examination of results from one sentence, with separate charts for the two speakers. Each figure in the cells is the average from three repetitions. The columns L (Hz), H (Hz) and duration are all direct measurements, with the range (H-L) being derived from the first two. The last column sums up, in general terms, the difference between that emotional state and the neutral base form.

Table 3: 'S/he has cooked yams' *ù dìgìwó nyɔ̌à* (Solomon)

	L (Hz)	H (Hz)	range (H-L)	duration	compared to neutral
neutral	128	153	25	798	—
bored	128	157	29	810	l-EXP, l-slower
angry	152	182	30	697	higher, l-EXP, faster
contemptuous	147	177	30	738	higher, l-EXP, faster
emphatic	135	172	37	743	l-higher, EXP, faster
surprise (no Q)	156	185	29	703	higher, l-EXP, faster

The first thing to note is that the two speakers had a few seemingly categorial differences in their expressions. Especially noteworthy is that the bored expression was slower than the neutral one for Solomon and faster for Salifu. Also, Solomon's angry and emphatic expressions were faster than Salifu's. There was other minor variation, but the main difference between speakers was speed in three utterances.

Table 4: 'S/he has cooked yams' *ù dìgìwó nyɔ́à* (Salifu)

	L (Hz)	H (Hz)	range (H-L)	duration	compared to neutral
neutral	134	155	21	680	—
bored	140	164	24	629	l-higher, l-EXP, faster
angry	142	172	32	682	higher, EXP
contemptuous	139	167	28	679	l-higher, l-EXP
emphatic	142	180	38	706	higher, EXP, slower
surprise (ques)	146	186	40	774	higher, EXP, longer

But more broadly, both speakers had results consistent with each other in that:

- Bored was slightly expanded in range, a definite but not robust result.
- Angry was definitely higher with an expanded range.
- Contemptuous was slightly higher and slightly expanded, again definite but not robust.
- Emphatic was higher, with an expanded range.

On the surprise intonation, Salifu consistently responded by turning the statement into a polar (yes/no) question ('She has cooked yams?'). Solomon, however, uttered a non-question surprise intonation, which was higher and faster. It seems likely that pragmatically, the polar question is a more natural response to a surprising situation, but this cannot be verified at this point.

Also, the pitch in a polar question in isolation is higher than in the corresponding statement, but the pitch in a polar question when someone is surprised is yet higher (Cahill 2012), and these two are quite distinguishable. The second situation is that which was produced and examined here.

Next we turn to a variety of input sentences, with the results of speaker Salifu. These are the same ones listed in Table 2.

Figure 2 shows the aggregate results for pitch of the six sentences that Salifu repeated with neutral intonation and various emotional states. For this, the raw data measurements were used and combined. (Sentence-by sentence summary tables are in the Appendix.) For each emotional state, the bottom measure is the initial Low tone of the sentence, and the second measure is the first High tone. Bars represent one standard deviation above and below the average.

As can be seen, the bored and contemptuous states have approximately the same starting pitch as neutral, but with the High of the sentences slightly lower than the neutral, they have a slightly narrower range. The angry and emphatic states start slightly higher than neutral, but have a significantly larger pitch range. The surprise intonation starts at the highest pitch of all (recall that Salifu turned this into a question), and also has a significantly larger pitch range than the neutral. At this point, as a rough approximation,
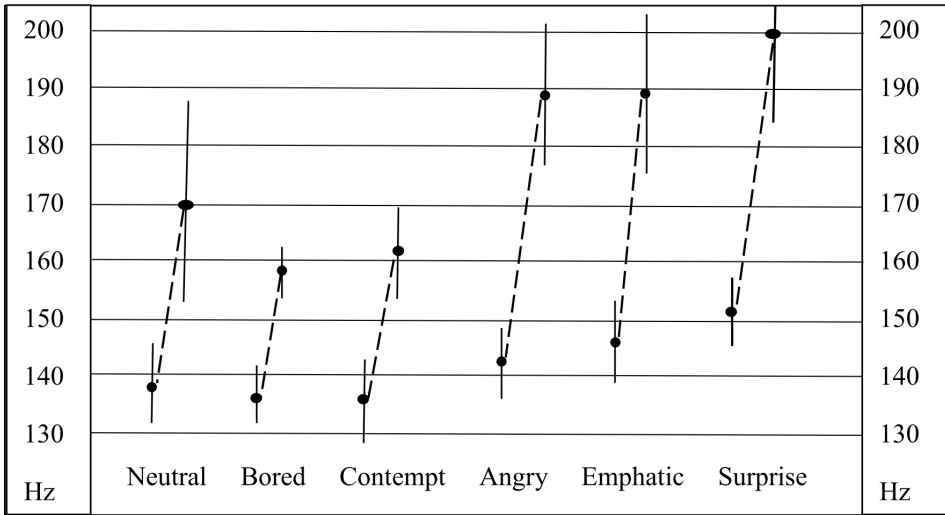
Figure 2: Base pitch and first H tone – a measure of height and of range

the bored and contemptuous intonations appear quite similar to each other, as do the angry and emphatic, while surprise stands somewhat apart.

Measurements of duration must be done sentence by sentence, since the target sentences are not uniform in syllable count. We would expect *ʊ̀ chɔ̀gìsìwó ˈbólíŋ*, with seven syllables, to have an inherently longer duration than the four syllables of *ʊ̀ gàwá ˈnyɔ́ŋ*, and indeed in the neutral form they average 801 vs. 651 ms respectively. Thus the *ratio* of the various emotive sentences to the neutral one is what is revealing, and these ratios are presented in Figure 3.
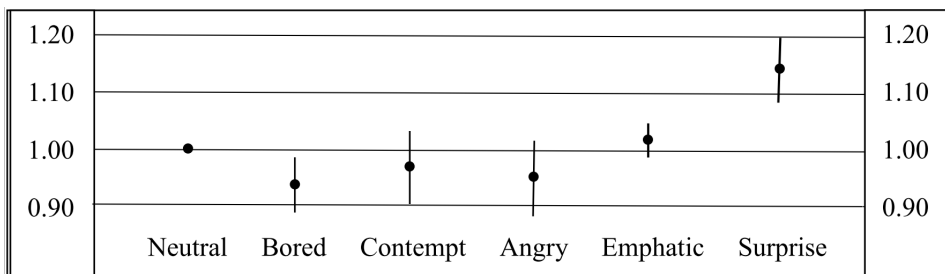
Figure 3: Ratio of duration of emotive sentences to corresponding neutral sentence

The duration of the surprise question is somewhat due to the extra mora added at the end of a polar question, as illustrated in Table 1. If 100 ms is subtracted from the average duration to account for this extra vowel, the duration of the surprise sentences drops closer to the range of the neutral.

The duration of the emphatic sentence is worth singling out, as it contrasts with the other sentences (ignoring surprise for now) as being longer in duration than neutral.

Several observations can be made on the basis of Figure 2 and Figure 3. Again, actual data for these sentences is found in the Appendix.

- The bored expressions in Salifu's speech were consistently faster than the neutral ones, and most of the time had a contracted range. No consistent pattern of raising or lowering the base pitch was found.
- The contemptuous expressions in Salifu's speech varied in speed, but were generally faster than neutral, and most of the time had a contracted range. Again, no consistent pattern of raising or lowering was found.
- The bored and contemptuous patterns thus were quite similar to each other.
- The angry expression was sometimes higher than neutral, mostly faster, but always with an expanded range.
- The emphatic expression was always higher, always with expanded range, and almost always slower.
- The surprise was always higher, with an expanded range, and longer. I use "longer" rather than "slower" because there is always an extra mora added.

4 Summary and discussion

A summary of generalizations is displayed in Table 5, with the caveat that these highest level generalizations conceal some detail. Besides previous measures, I also add some non-quantitative notes on volume/intensity , based on observations of the wave forms.

Table 5: Properties of emotions in Kɔnni, compared to neutral sentence intonation

	base pitch	range	speech rate	volume
bored	same	contracted	faster	quieter
contemptuous	same	contracted	varied/faster	quieter
angry	little higher	expanded	faster	same
emphatic	higher	expanded	slower	louder
surprise	higher	expanded	longer	louder

Comparing the properties of the emotions in Table 5 with Solomon's production of his one sentence and variants in Table 3, we see that the angry sentence had the same qualities for both speakers. The others were similar, but varied in one or more characteristics. As noted in the discussion after Table 4, the main speaker differences were in duration of the utterances.

The similarity between the bored and the contemptuous patterns may suggest that these are closely related in Kɔnni speakers' minds. It is easy to imagine that someone

who is expressing contempt would act as if he were bored. However, other language studies have sometimes found overlap in pitch characteristics of unrelated emotions (see discussion of Thai in and around Table 6 below), so phonetic likeness does not necessarily entail emotional similarity or identity. In light of the fact that there are different expressions for bored and contemptuous in Kɔnni (see discussion below), it seems more probable that these are merely phonetic overlaps of unrelated emotions.

If bored/contemptuous is counted as one intonation pattern, there are four distinct configurations of intonation that indicate emotional states in Kɔnni:

- A bored/contemptuous sentence is generally pronounced at a lower volume than neutral, with a contracted range, and faster than neutral. These all conspire together to reduce the overall prominence of the sentence.
- An angry sentence is raised a bit, but its main characteristic is the expanded range and faster speed than neutral.
- An emphatic sentence is significantly higher pitch than neutral, with expanded range, and a slower speed, the latter two of which presumably helps the hearer to clearly identify every word. The slower speed distinguishes this from the angry sentence.
- A surprise sentence, even excluding the additional tonal and vocalic autosegments added to the final syllable, is also raised in pitch, with expanded range. It somewhat resembles the emphatic sentence in these respects.

As noted in the title, this is must be regarded as quite a preliminary investigation; the results are suggestive and compatible with studies in other languages, but are too limited to be considered definitive. There are two obvious limitations to this study, one of which is more amenable to attack than the other. One limitation is limited data from a limited number of speakers, which could be remedied if time and conditions permitted. Secondly, and ideally, recordings would be made in natural settings rather than the artificial acting out that was done for this study. However, in such a situation, the amount of data collected in a reasonable amount of time would be a challenge, and control of the variables (the same sentence, number of repetitions, same speaker) would be very difficult. I know of no study which has actually put this into practice.

An obvious follow-up study is to see if other Kɔnni speakers can reliably identify the emotion acted out. It has been demonstrated that people can recognize the intonation patterns of some emotions, even cross-linguistically. Gussenhoven (2004: 72) cites Van Bezooijen (1984: 128) for a case of this. Taiwanese and Japanese speakers identified sadness, anger, and surprise by Dutch speakers at above chance levels. However, contempt and shame were not recognized. On the other hand, Gårding (1998: 12) notes that Swedish emotional intonations are not so well established, and experiments have shown that Swedish speakers cannot distinguish happiness from anger by intonation alone. With the long-distance setup for this study, testing recognition of these emotional states by native Kɔnni speakers was impractical. Testing by speakers of other languages, particularly related African languages, would be a possible next step.

For an experiment of this type, where the aim is to collect data on intonation of different emotions from another language, a relevant question is if speakers of that language actually have that emotional category in their language, and words or expressions for it. Not all emotions translate directly, and if a speaker is aiming for an emotion that he or she is uncertain about, the results will also be uncertain and not dependable, both for that study and for comparative cross-linguistic purposes. In the midst of this writing, I inquired about this, and Konlan Kpeebi was able to confirm the following terms with another Kɔnni speaker, Mr. Ben Saibu, native Kɔnni speaker and lawyer. The terms for emotions in Kɔnni are the following.

1. Anger is straightforwardly termed *sınyurıŋ*.

2. Surprise is straightforwardly *wuboŋkiŋ*.

3. Boredom is expressed as *wʋkpaaŋ* (the idea is "too much talking/too many issues").

4. For contempt, there are actually several phrases:

   - *dansı vuoŋ yɔɔrı* 'look person nothing' (= he does not regard anyone)
   - *Ʋ nine ka suuli vuoŋ* 'his eyes not fill person' (= he does not regard anyone)
   - *Ʋ ka daansı ye vuoŋ* 'he does not look see person' (= he does not see anyone)

5. For emphasis, Saibu suggested *Vii balı* which literally means 'say it again', which would make the speaker assume the original hearer didn't quite understand it, and she or he should repeat it more understandably. This is not a direct translation of the English term *emphasis*, but probably evokes the corresponding pronunciation.

In general, then, it appears that Kɔnni does have reasonable lexical (and hence cultural) approximation for the emotional terms or categories elicited in this study. This also makes it more likely that the phonetic similarity between the bored and contemptuous intonations do not indicate that these emotions are identical in speakers' minds.

The results of this study are compatible with those found in other languages, and specifically add to the knowledge of how tone languages are able to express paralinguistic intonation in a systematic way. For example, in German emphatic is "more of everything", with a wider pitch range and longer duration (Gibbon 1998: 91). In Swedish (Gårding 1998: 122–123), anger has a wider pitch range. In Vietnamese (Đỗ Thế Dũng, Trần Thiên Hương & Boulakia 1998; Brunelle, Hạ & Grice 2012: 402, 412–413), anger has a comparatively shorter duration, greater pitch movement, increased loudness, and higher base pitch. Of the 13 attitudinal or emotive states for Thai reported in Luksaneeyanawin (1998: 382), the ones that overlap with this study are presented in Table 6.[6]

The 13 emotions have overlapping pitch characteristics in several cases, and Luksaneeyanawin groups them into four *tunes*. For example, the group with bored also has concealed anger in it, which may possibly have a connection to the contemptuous emotion in this study.

---

[6] Pitch height being "higher and lower" in Luksaneeyanawin's terms means that pitch is higher for tones with a non-low starting point, and pitch is lower for the tones that have a low starting point. The ± symbol means that the quantity is sometimes enhanced, but not always.

Table 6: Properties of emotions in Thai (extracted from Luksaneeyanawin 1998: 382)

	pitch height	range	length	volume
bored	lower	narrower	shorter	softer
angry	higher & lower	wider	?? longer	very loud
emphatic	higher & lower	wider	longer	louder
surprise	higher	narrower	??shorter	??louder

This study, preliminary and limited as it is, nonetheless illustrates definite patterns of how Kɔnni speakers vary their speech to express different emotional states. It also shows that the paralinguistic intonational variations that indicate states such as anger, emphasis, etc. have patterns that are similar to those documented for other languages, even tonal ones such as Thai. Whether these patterns are universal or merely common will depend on the results of more studies.

Acknowledgements

Appendix

Each cell in these tables reports the average of three repetitions.

Table A: 'S/he went to market' *ʊ̀ gàwá ˈnyɔ́ŋ* (Salifu)

	L (Hz)	H (Hz)	range (H-L)	duration	compared to neutral
neutral	150	189	39	651	—
bored	135	169	34	552	lower, l-CONT, faster, quieter
angry	150	203	53	558	EXP, faster, ˜ same intensity
contemptuous	134	166	32	548	lower, l-CONT, faster, quieter
emphatic	159	212	53	704	higher, EXP, slower, louder
surprise (Q)	158	226	67	804	higher, EXP, longer

Table B: 'S/he has bought oil' *ʊ̀ dààwá kpááŋ* (Salifu)

	L (Hz)	H (Hz)	range (H-L)	duration	compared to neutral
neutral	134	144	10	700	—
bored	140	155	15	629	l-higher, l-EXP, faster
angry	140	172	32	682	l-higher, EXP, faster
contemptuous	134	149	15	692	l-EXP, l-faster
emphatic	145	172	27	717	higher, EXP, slower
surprise (Q)	152	185	33	806	higher, EXP, longer

Table C: 'S/he has cooked eggs' *ù dìgìwó gílà* (Salifu)

	L (Hz)	H (Hz)	range (H-L)	duration	compared to neutral
neutral	139	162	23	658	—
bored	139	158	19	614	l-CONT, faster
angry	146	184	38	672	l-higher, EXP, slower
contemptuous	144	165	21	669	l-higher, l-CONT, l-slower
emphatic	146	181	35	668	l-higher, EXP, l-slower
surprise (Q)	148	198	50	763	higher, EXP, longer

Table D: 'S/he has fried eggs' *ʊ̀ chʊ̀ŋwá gílà* (Salifu)

	L (Hz)	H (Hz)	range (H-L)	duration	compared to neutral
neutral	132	163	31	716	—
bored	133	156	23	683	CONT, faster
angry	130	181	51	716	EXP
contemptuous	135	159	24	706	CONT, l-faster
emphatic	142	184	42	725	higher, EXP, slower
surprise (Q)	148	189	41	779	higher, EXP, longer

Table E: 'S/he has eaten TZ' *ù dùùwó ꜛsááŋ* (Salifu)

	L (Hz)	H (Hz)	range (H-L)	duration	compared to neutral
neutral	139	173	34	697	—
bored	137	162	25	679	CONT, faster
angry	Missing data				
contemptuous	138	162	24	713	CONT, slower
emphatic	143	191	48	713	l-higher, EXP, slower
surprise Q	155	205	50	756	higher, EXP, longer

Table F: 'S/he has fetched fire' *ò chɔ̀gìsìwó ꜛbólíŋ* (Salifu)

	L (Hz)	H (Hz)	range (H-L)	duration	compared to neutral
neutral	139	191	52	801	—
bored	131	157	26	804	lower, CONT
angry	136	200	64	760	l-EXP, faster
contemptuous	132	169	37	779	l-lower, CONT, faster
emphatic	142	199	57	802	l-higher, l-EXP
surprise Q	148	196	48	909	higher, EXP, longer

References

Bolinger, Dwight (ed.). 1972. *Intonation: Selected readings*. Middlesex: Penguin Books.

Brunelle, Marc, Kiều Phương Hạ & Martina Grice. 2012. Intonation in Northern Vietnamese. *The Linguistic Review* 29. 3–36.

Cahill, Michael C. 2007. *Aspects of the morphology and phonology of Kɔnni*. Dallas: SIL International.

Cahill, Michael C. 2012. Polar question intonation in Kɔnni. In Michael R. Marlo, Nikki B. Adams, Christopher R. Green, Michelle Morrison & Tristan M. Purvis (eds.), *Selected Proceedings of the 42nd Annual Conference on African Linguistics*, 90–98. Somerville, MA: Cascadilla Proceedings Project.

Cahill, Michael C. 2016. Kɔnni intonation. In Laura J. Downing & Annie Rialland (eds.), *Intonation in African languages* (Phonetics and Phonology 24), 53–88. Berlin: DeGruyter Press.

Cruttenden, Alan. 1997. *Intonation*. 2nd edn. Cambridge: Cambridge University Press.

Đỗ Thế Dũng, Trần Thiên Hương & George Boulakia. 1998. Vietnamese. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 395–416. New York: Cambridge University Press.

Downing, Laura J. & Annie Rialland (eds.). 2016. *Intonation in African languages*. Berlin: DeGruyter Press.

Fiedler, Ines & Stefanie Jannedy. 2013. Prosody of focus marking in Ewe. *Journal of African Languages and Linguistics* 34(1). 1–46.

Gibbon, Dafydd. 1998. Intonation in German. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 78–95. New York: Cambridge University Press.

Gårding, Eva. 1998. Intonation in Swedish. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 112–130. New York: Cambridge University Press.

Green, Christopher R. 2009. Prosody and intonation in non-Bantu Niger-Congo languages: An annotated bibliography. *Electronic Journal of Africana Bibliography* 11. http://ir.uiowa.edu/ejab/vol11/iss1/1/.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Harley, Matthew. 2009. Focus constructions in Tuwuli. *Journal of West African Languages* XXXVI(1-2). 75–90.

Hirst, Daniel & Albert Di Cristo. 1998. A survey of intonation systems. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 1–44. New York: Cambridge University Press.

Hyman, Larry M. & Kemmonye C. Monaka. 2011. Tonal and non-tonal intonation in Shekgalagari. In Sonia Frota, Gorka Elordieta & Pilar Prieto (eds.), *Prosodic categories: Production, perception and comprehension*, 267–290. Dordrecht: Springer Verlag.

Ladd, D. Robert. 2008. *Intonational phonology*. 2nd edn. (Cambridge Studies in Linguistics 119). New York: Cambridge University Press.

Ladd, D. Robert, Klaus Scherer & Kim Silverman. 1986. An integrated approach to studying intonation and attitude. In Catherine Johns-Lewis (ed.), *Intonation in discourse*, 125–138. Breckenham, Kent: Croom Helm.

Lieberman, Philip. 1967. *Intonation, perception, and language* (Research monograph 38). Cambridge, MA: MIT Press.

Lomotey, Charlotte. 2014. *The prosodic realization of focus in Awutu: A preliminary analysis*. (Paper presented at the 45th Annual Conference on African Linguistics, University of Kansas, April 17-19, 2014.)

Luksaneeyanawin, Sudaporn. 1998. Intonation in Thai. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 376–394. New York: Cambridge University Press.

Rialland, Annie & Stéphane Robert. 2001. The intonational pattern of Wolof. *Linguistics* 39(5). 893–939.

Schwarz, Anne. 2009. To be or not to be? About the copula system in Buli (Gur). In Margarida Petter & Ronald Beline Mendes (eds.), *Proceedings of the Special World Congress of African Linguistics – São Paulo 2008: Exploring the African language connection in the Americas*, 263–278. São Paulo: Humanitas.

Van Bezooijen, Renée. 1984. *Characteristics and recognizability of vocal expressions of emotion.* Dordrecht: Foris.

Chapter 3

The coding of identifiability in Mooré

Amos Teo

University of Oregon

This paper looks at the morpheme /-wã/ ~ /-ã/ in Mooré (Gur, Burkina Faso), which previous analyses (e.g. Peterson 1971; Canu 1976; Nikièma 1989) call a "definite" marker. The paper aims to show that an analysis of /-wã/ ~ /-ã/ needs to consider how NPs are marked for the semantic and/or discourse-pragmatic function of referentiality and/or identifiability. By drawing on both elicited and textual data, the analysis shows that, in addition to marking identifiability, /-wã/ ~ /-ã/ is often used to mark contrastive focus. Furthermore, in some cases, /-a/ (without nasality) can still be used to mark an NP as both referential and identifiable to both the speaker and listener, but this depends on the phonological shape of the NP. The patterns suggest that one can analyze the marker /-ã/ as two morphemes: /-a/; and a separate nasal morpheme, with the later having the potential to also mark contrastive focus. However, this analysis is complicated by the phonological form of the NP.

1 Introduction

Mooré (ISO: mos), also known as Mossi, is a Gur language within the larger Niger-Congo family. It is one of the main languages of Burkina Faso, with an estimated 5 million speakers in the country (Lewis, Simons & Fennig 2013). Estimates also give around 60,000 speakers living in neighboring countries, including Côte d'Ivoire, Benin, Togo and Senegal.

Previous descriptions of Mooré (e.g. Peterson 1971; Canu 1974; 1976; Nikièma 1989) point to the existence of a "definite" morpheme /wã/ that follows the noun phrase (NP) it modifies, along with its phonologically determined allomorph /-ã/. The latter is described as occurring after nasal consonants or when replacing the vowel in a suffix that begins with a consonant (i.e. the consonant is part of the suffix). In most analyses, the treatment of /-wã/ as a "definite" marker appears to be motivated by its translation into English as 'the' or French 'le' / 'la' / 'les'. These analyses are problematic, given that the term *definite* (or *définisseur*) is not clearly defined in the descriptions cited above, along either structural or functional lines, nor are examples of /-wã/ ~ /-ã/ from actual discourse given in those descriptions.

This paper therefore aims to show that an analysis of /-wã/ ~ /-ã/ needs to consider how NPs are marked for the semantic / pragmatic notions of identifiability in Mooré;

I will reserve the term *definite* to describe a structural or formal category (e.g. as per Du Bois 1980). By drawing on both elicited and textual data, the analysis will show that /-wã/ ~ /-ã/ is often used to mark contrastive focus in addition to identifiability, while /-a/, without nasality, can be used to mark an NP as identifiable to both the speaker and listener, at least in the speaker's judgment of the listener's state of mind. This suggests that one can analyze the form /-ã/ as consisting of two morphemes: /-a/; and a separate nasal morpheme, with the latter used to mark contrastive focus. However, this analysis is complicated by the phonological form of a lexical noun with its citation noun class suffix, specifically the vowel of the noun class suffix.

To illustrate the problem, consider the following elicited examples:

(1) màm   nɛ́ɛ̀   bʊ́ʊ́ʁá.[1]
    màm   nɛ́-ɛ̀   bʊ́ʊ́-gá
    1SG.SUBJ see-AFF goat-CL12

    'I see a goat.' (AT_20140605)[2]

(2) màm   nɛ́ɛ̀   bʊ́ʊ́ʁá̰.
    màm   nɛ́-ɛ̀   bʊ́ʊ́-g-á̰
    1SG.SUBJ see-AFF goat-CL12-ã̰

    'I see the goat.' (that we talked about); or 'I see **the goat**.' (as opposed to seeing something else); or 'I **see the goat**.' (as opposed to not seeing the elephant) (AT_20140605)

(3) màm   nɛ́ɛ̀   wóbʁò.
    màm   nɛ́-ɛ̀   wób-gò
    1SG.SUBJ see-AFF elephant-CL15

    'I see an elephant.' (AT_20140605)

(4) màm   nɛ́ɛ̀   wóbʁà.
    màm   nɛ́-ɛ̀   wób-g-à
    1SG.SUBJ see-AFF elephant-CL15-*à*

    'I see the elephant (that we talked about).' (AT_20140605)

(5) màm   nɛ́ɛ̀   wóbʁà̰.
    màm   nɛ́-ɛ̀   wób-g-à̰
    1SG.SUBJ see-AFF elephant-CL15-*ã́*

    'I see **the elephant**.' (as opposed to seeing something else); or 'I **see the elephant**.' (as opposed to not seeing the elephant) (AT_20140605)

---

[1] The 1st line of the examples gives a phonetic transcription of the data, while the 2nd line gives the proposed phonemic form with morpheme boundaries. Both lines are provided since some of the proposed phonemic forms differ slightly from the phonetic forms. Furthermore, if there are any errors or disagreements, future researchers can see if an error lies in the phonetic transcription or in the phonological analysis.

[2] The source for each example is given in parentheses with the name of the file. More discussion is still needed with regards to giving access to the recorded data.

In (1)–(2), we can see a two-way distinction between the absence vs. presence of /-ã/, which corresponds to the use of the indefinite or definite article in the English translation. However, in (3)-(5) we find a three-way distinction between no marker, /-a/ and /-ã/. Here, the nasal form is not clearly required for a definite reading of the NP. It will be later shown that in discourse the lack of nasality in an example like (1) does not necessarily correspond to a non-identifiable (or indefinite) reading.

The paper is organized as follows: in §2, I define the semantic / pragmatic notion of identifiability, and its relation to the structural category of definiteness. In §3, I briefly look at some previous descriptions of "definiteness" in Mooré. In §4, I give a brief description of NP morphology. In §5, I look at the expression of non-identifiable NPs in Mooré; while in §6, I consider the expression of identifiable NPs. In §7, I consider some of the complications in analyzing how identifiability and focus are coded in Mooré. Finally, in §8, I summarize the analysis presented in this paper and propose further avenues for research.

The data for this study come from Timbwaoga Aimé Judicaël Ouermi, a male Mooré speaker in his early 20s from Ouagadougou, Burkina Faso. It is acknowledged that future research will require the participation of more speakers of the language. The data were recorded at the University of Oregon, Eugene, over a 9 month period between 2013 and 2014.

2 Identifiability and Definiteness

An important distinction to make is between the formal categories of DEFINITE / INDEFINITE and the semantic / pragmatic categories of REFERENTIALITY and IDENTIFIABILITY. According to Du Bois (1980: 280), "[a] noun phrase is *referential* when it is used to speak about an object as an object, with continuous identity over time". The "object" in question could be physical or conceptual; specifically known or unknown; a single entity or multiple ones; and it may exist in the real world or in a hypothetical world, or "universe of discourse" as per Givón (2001: 388)'s use of this term. A referential NP can function to either: (a) activate a "mental file" for a particular object; or (b) refer back to a "previously opened mental file". This is in contrast to non-referential NPs, which are not sensitive to any previous mentions in a discourse, nor are they sensitive to any semantic distinction between singular and plural (at least in English; Du Bois 1980: 210). Non-referential Ns/NPs can appear: (a) as the modifier element in a compound; (b) in predicating expressions (denoting proper inclusion in a category); and (c) as what Du Bois refers to as "conflated objects", where the object is non-individuated and "conflated" with the verb (similar to noun incorporation).

Once an NP is interpreted as referential, it can be interpreted as identifiable or non-identifiable. According to Du Bois (1980: 232), "[i]dentification ordinarily involves singling out the particular referent intended by the speaker". Du Bois (1980: 233) goes on to propose a "curiosity principle" that states, "A reference is counted as identifiable if it identifies an object close enough to satisfy the curiosity of the hearer" – though it may be more precise to think of this as what the speaker *believes* is close enough to satisfy the

curiosity of the hearer. The level of satisfaction seems to be based on the Gricean maxims of quantity and relevance – it is more common for people in everyday communication to only partially identify referents than to specify every single characteristic that would identify a referent as a unique entity.

The term *definite* is often used to describe a referent that can be identified by the listener (as in Comrie 1989: 65), or to describe a nominal expression denoting a referent that is presumed by the speaker to be identifiable by the listener (as in Lambrecht 1994: 79). The reasons for following the latter option of using the term *definite* to name a formal category (i.e. for particular forms), separate from the semantic or pragmatic concept of identifiability, are clear if we consider the use of the definite article *the* in English. Typically, *the* marks NPs that point to objects that are both referential and identifiable, as in (6). However, there are contexts in which *the* can also mark a non-referential (and generic) NP, as in (7):

(6)   I saw a small elephant get attacked by a lion. **The** elephant got pretty angry.

(7)   **The** elephant is the largest land animal on earth.

Similarly, the indefinite article *a / an* in English can mark a referential non-identifiable NP, as in (8), or a non-referential (and hence by default, non-identifiable) NP, as in (9).

(8)   I'm looking for **a** black jacket. I think I left it here this morning.

(9)   I'm looking for **a** black jacket. Do you sell any in your store?

Consequently, it is useful to reserve the terms *definite* and *indefinite* for structural / formal categories in a language, and to see how they align with the semantic / pragmatic functions of marking referentiality and / or identifiability, and perhaps even contrastive focus. As Lambrecht (1994: 79) notes, the categories of formal definite marking and identifiable information status do not always align perfectly. We shall see that this is also the case for Mooré.

3 Previous descriptions of definiteness in Mooré

Most descriptions of Mooré include a brief account of a "definite" marker /wã/. Peterson (1971: 77) states that what he calls the "definitizer" has the form /wã/, which goes at the end of relative clauses and also on nouns, as in:

(10)   Example as given in Peterson (1971: 77):
       kí wǽ [kí ẃǽ]
       'the millet'

However, according to Peterson, if the marker follows an elided vowel (or a nasal consonant), the /w/ is deleted and the vowel merges with the word, as in the following examples:

(11)   Examples as given in Peterson (1971: 77):
         bòãŋgá   bòãŋgá [bwə̀ŋgá]   béngré béngrá [béŋgɘ́rá]
         'donkey' 'the donkey'   'bean' 'the bean'

Canu (1974: 179) similarly posits the underlying form of the *définisseur* 'definitizer' as /uã⁻/, which "dans le discour rapide et l'élocution relachée … s'amalgame avec la dernièrne voyelle du nominal" ("in fast or relaxed speech … merges with the final vowel of the noun"), as in:

(12)   Examples as given in Canu (1974: 179):
         bá:gá uã̄ → /bá:gɛ́́ / [bá:ɣɛ́́]
         'le chien' ('the dog')
         uōbgò uã̄ → /uōbgɛ̀̄ / [wōbɣɛ̀̄]
         'l'éléphant' ('the elephant')
         zóm uã̄ → /zómã̄ / [zómã̄]
         'la farine' ('the flour')

Similarly, Nikièma (1989: 96) notes that: "[l]a marque du défini *wã* est réduite à *ã* après un mot terminé par une consonne" ("the definite marker *wã* is reduced to *ã* after a word ending in a consonant"); but he does not elaborate further.

Interestingly, in none of these sources is there any mention of the definitizer having the form /-a/ with no nasality. For instance, note the following example, presented earlier as (4):

(13)   màm   nɛ́ɛ̀   wóbʁà.
         màm   nɛ́-ɛ̀   wób-g-à
         1sg.subj see-aff elephant-cl15-*à*

         'I see the elephant (that we talked about).' (AT_20140605)

This leads us to the following questions that will be addressed in this paper:

- What is / are the function(s) of /-a/ on NPs? Is /-a/ a definite marker? How does it relate to the coding of referentiality and identifiability in Mooré?
- What, then, is the function of /-wã/ ~ /ã/ in Ouagadougou Mooré as represented in the speech of the consultant for this work? Is this also a definite marker? Is its function different from what has been claimed in previous descriptions of Mooré?

4 Noun phrase structure

In order to understand NP marking, we must first look briefly at the structure of the Noun Phrase in Mooré.[3] Table 1 gives a selection of noun classes and examples for each class, adopting the traditional Niger-Congo noun class numbering system. The third column provides the citation form of the noun, which is the form given when the speaker is asked

---

[3] For reasons of space, we will not look at pronouns and proper nouns in this paper.

Table 1: Mooré noun class markers and examples

Noun class	Gloss	Citation	Form with -*a*	Form with *(w)ã*[4]
1	'woman'	*pág-á*	*pág-á*	*pág-ã*
2	'women'	*pág-bá*	*pág-b-á*	*pág-b-ã*
5	'rock'	*kúg-rì*	*kúg-r-á*	*kúg-r-ã*
6	'rocks'	*kúg-à*	*kúg-à*	*kúg-ã*
12	'goat'	*bʊ́ʊ́-gá*	*bʊ́ʊ́-g-á*	*bʊ́ʊ́-g-ã*
13	'goats'	*bʊ́ʊ́-sé*	*bʊ́ʊ́-s-à*	*bʊ́ʊ́-s-ã*
14	'*sagbo*' (type of doughy food)	*ság-bó*	*ság-b-á*	*ság-b-ã*
15	'elephant'	*wób-gò*	*wób-g-à*	*wób-g-ã*
19	'bicycle'	*wèè-fó*	*wèè-f-á*	*wèè-f-ã*
21	'grave'	*yáá-dò*	*yáá-d-à*	*yáá-d-ã*
22	'alcohol'	*ráà-m*	*ráà-m-à*	*ráà-m-ã*
(no marker)	'millet'	*kí*	-	*kí-wã*
	'cart'	*ʃárétè*	*ʃárét-à*	*ʃárét-ã*

to give the Mooré equivalent of an English or French word. For labeling convenience, I use the term *citation form* to refer to such word forms, before assigning a functional label to them. The fourth column gives the form of each noun with the /-a/ suffix: in all cases, the vowel of the noun class suffix that we see in the citation form is elided. The fifth column gives the form of each noun with the other definite /-wã/ suffix, which is realized as -*ã* when added to nouns with overt noun class markers in their citation form. Monosyllabic nouns which do not have an overt class marker, e.g. /kí/ 'millet', do not take the suffix /-a/, only the suffix /-wã/.

The vowel in the noun class suffix of a noun in the citation form is typically only produced at the end of a clause / sentence, as in (14) and is elided in other contexts, as in (15) and (16).

(14)   à músá   nɛ́ àlì yáá   kàràmbíísì.
       à=músá   nɛ́ àlì yà-à   kàràmbíí-sì
       3SG.SUBJ=PN and PN COP-AFF student-CL13

       'Moussa and Ali are students.' (AT_2014-02-25_NominalPreds_Existential, Ex.3)

(15)   wóbs   rítà   tɪ̃ìsè.
       wób-s   ríd-d-à   tɪ̃ì-sè
       elephant-CL13 eat-PROG-AFF tree-CL13

       'Elephants eat trees.' (general) (AT_2014-05-21_nasality, Ex.1)

---

[4] It is unclear what the tone on the final nasalized vowel is in each word. Although the consultant does not consider there to be any difference in pitch between the -*a* and -*wã* forms, a phonetic difference in pitch has sometimes been perceived by the researcher: specifically, a slight dip in pitch on the nasalized vowel.

(16) wóbs   rítà   tĩ̀s   dár   fáằ.
     wób-s   ríd-d-à   tìì-s   dá-r   fáằ
     elephant-CL13 eat-PROG-AFF tree-CL13 day-CL5 all

     'Elephants eat trees all the time.' (general) (AT_2014-05-28)

The syntactic role of the NP does not affect whether the final vowel of the citation form (third column of Table 1) is produced or not: a comparison of (15) with (16) shows that the subject [wóbs] and the object[5] [tĩ̀s] both do not display the final vowel of their respective noun classes as long as they are not in clause-final position. However, it is possible for the final vowel to be produced in careful speech, or when listing out nouns, as in [mɔɔ́dó] in (16).[6]

(17) màm   néɛ̀   mɔɔ̀dó   lá   tììsé.
     màm   né-ɛ̀   mɔɔ̀-dó   lá   tìì-sé
     1SG.SUBJ see-AFF bush-CL21 and tree-CL13

     'I see bushes and trees.' (AT_2014-01-13_WomanDonkeyCart_Text Ex.17)

It is important to note that for some noun classes, the noun with the definite /-a/ suffix is homophonous with the noun in its citation form, i.e. nouns from noun classes 1, 2, 6 and 12. However, despite the apparent homophony between some of the nouns in citation form and their /-a/ suffixed forms (fourth column in Table 1), potential ambiguity is only an issue with such nouns in clause-final position, as with [kàràmbíígá] in (18). In general, it is possible to tell if a noun has the /-a/ suffix if it occurs in non-final position in a clause or sentence: the *-a* suffix is not elided in words like [ráwá] in (18) and [kàràmbíígà] in (19) when in non-final position. Compare these with the elided form [káràmbííg] in examples (20) and (21).

(18) ráwá yàà   káràmbíígà.
     ráw-á yà-à   kàràmbíí-gà
     man-*a* COP-AFF student-CL12

     'The man is a student.' (AT_2014-02-11_PropertyPredicates, Ex.25)

(19) à músà   ká kàràmbíígà   yè.
     à=músà   ká kàràmbíí-g-à   yè
     3SG.SUBJ=PN NEG student-CL12-*a* CFN

     'Moussa is not the student (we talked about).' (AT_2014-03-11_NominalPredicates, Ex.39)

(20) fò   yáá   kàràmbíîg   là?
     fò   yà-à   kàràmbíí-g   là?
     2SG.SUBJ COP-AFF student-CL12 Q

     'Are you a student?' (AT_2014-02-25_NominalPreds_Existential Ex.11)

---

[5] Here, I use the term *subject* to refer to both the subject of a transitive clause and the subject of an intransitive clause, and *object* to refer to the object of a transitive clause.

[6] It is possible that the final vowel is produced at the end of some kind of intonational unit, as opposed to some kind of syntactic unit.

(21) à músá   ká kàràmbíìg yè.
     à=músá   ká kàràmbíí-g yè
     3sg.subj=pn neg student-cl12 cfn
     'Moussa is not a student.' (AT_2014-02-25_NominalPreds_Existential, Ex.1)

In general, the /-a/ and /-wã/ forms appear only on the last element of the NP, where they replace the vowel of a class marker suffix. They also do not undergo vowel elision in non-final position in a clause. Deserving special mention are the demonstrative modifiers [káŋá] 'this' and [kã́ĩsá] 'these', which always end in [-a], even in non-clause-final position, as demonstrated by (22)–(24).

(22) nù(g) káŋá   yàà   bédrè.
     nù(g) kán-g-á   yà-à béd-rè
     hand dem-cl12-*a* cop-aff big-cl5
     'This hand is big.' (KB_20140224_08_Noun-phrase, Ex.2)

(23) *nù(g) káŋ yàà bédrè.
     ('This hand is big.')

(24) kòr bédà yììb kã́ĩsá   yàà   péɛ̀lsè.
     kòr béd-à yììb kã́ĩ-s-á   yà-à péɛ̀l-sè
     bag big-*a* two dem-cl13-def cop-aff white-cl13
     'These two big bags are white.' (KB_20140224_08_Noun-phrase, Ex.29)

However, in genitive and relator noun constructions, the definite /-a/ suffix is not necessarily the last element in the phrase. For example, in (25), [ròòdá] 'houses'[7] takes the /-a/ suffix, but [kwèlàms] 'doors' does not; while in (26), [fĩ̀lmà] takes the /-a/ suffix.

(25) súg   ròòdá   kwèlàms yáá   rààdó.
     sú-g   ròò-d-á   kwèlàm-s yà-à ràà-dó
     thatch-cl15 house-cl21-*a* door-cl13 cop-aff wood-cl21
     'The thatch house doors are (made of) wood.'
     (AT_2014-01-13_Women-carrying-pots-on-head_Text Ex.10-11)

(26) fĩ̀lmà síŋrè
     fĩlm-à sín-rè
     film-*a* start-cl5
     'Aat the start of the film' (Pear Story, Ln1)

With this background on the structure of Mooré nominals/NPs, we now turn to the discourse-based information structure status of nouns, and how those statuses are coded.

---

[7] An anonymous reviewer has pointed out that *roodo / rooda* is not an acceptable plural of 'house', even taking some dialectal differences into account. However, this was the form produced by our consultant, which may reflect differences due to age, geography or the fact that the speaker has been living overseas for a number of years.

5 Indefinite NPs

We first look at the distribution of NPs in citation form: recall that, as seen in (27) and (28), these forms retain the vowel of the noun class suffix when the NP occurs in the final position in a clause or phrase (cf. [tĩ̀sè] in (27)), but the final vowel is usually elided when the non-identifiable NP does not appear in final position (cf. [wóbs] and [tĩ̀s] in (28)).

(27)　wóbs　　　rítà　　　　tĩ̀sè.
　　　wób-s　　　ríd-d-à　　　tìì-sé
　　　elephant-CL13 eat-PROG-AFF tree-CL13

　　　'Elephants eat trees.' (general) (AT_2014-05-21_nasality, Ex.1)

(28)　wóbs　　　rítà　　　tĩ̀s　　dár　　fáà̰.
　　　wób-s　　　ríd-d-à　tìì-s　　dá-r　　fáà̰
　　　elephant-CL13 eat-PROG-AFF tree-CL13 day-CL5 all

　　　'Elephants eat trees all the time.' (general) (AT_2014-05-28)

In semi-elicited narrative tasks, where the speaker was asked to describe what he could see in a picture or retell a story based on a video, the first mention of a referent is usually in the citation form of the noun, e.g. [ʃárétè] 'cart' and [rààdó] 'wood' in (29); [tì bíísì] 'fruit' in .

(29)　màm　nɛ́ɛ̀　　páʁá,　　bwàŋá,　　ʃárétè lá　rààdó.
　　　màm　nɛ́-ɛ̀　　pá-gá　　bwàn-gá　ʃárétè lá　ràà-dó
　　　1SG.SUBJ see-AFF woman-CL1 donkey-CL12 cart　and wood-CL21

　　　'I see a woman, a donkey, a cart and wood.'
　　　(AT_2014-01-13_WomanDonkeyCart_Text Ex.1)

(30)　à　　yák̀d̀　　tì　bíísì.
　　　à　　yák-d̀[8]　tì　bíí-sì
　　　3SG.SUBJ pick-PROG tree baby-CL13

　　　'He was picking fruit...' (Pear Story, Ln3-a)

In these examples, the speaker is pointing out to the listener entities that have not yet been previously mentioned in the discourse. As such, these NPs denote referential, but non-identifiable entities. However, once the referents have been established in discourse, and are identifiable by the listener, subsequent mentions of the NP take the -à suffix, as in [ʃárétà] (31) and [tì bíísà] in (32):

(31)　á　　ká zó　ʃárétà　yé.
　　　á　　ká zó　ʃárét-à　yé
　　　3SG.SUBJ NEG stand cart-DEF[9] CFN

　　　'She is not on the cart.' (AT_2014-01-13_WomanDonkeyCart_Text Ex.7)

---

[8] It is not clear why the speaker did not produce the suffix /-a/ 'AFF' in this example.

(32) tì bíísà   wéndà   gwèyáàbè.
      tì bíí-s-à   wén-dà   gwèyáàbè
      tree baby-CL13-DEF resemble-PROG guava

      'The fruits look like guava.' (Pear Story, Ln3-b)

In Du Bois (1980)'s terms, it appears that the non-definite NPs are used to activate a "mental file", while the definite NPs are referring back to these "previously opened files". Given that the citation forms of nominals are used to refer to entities that the speaker may presume are non-identifiable to the listener at that point in the discourse,[10] it would be suitable to call these the 'indefinite' forms of the nominals. Conversely, we might tentatively treat the /-a/ suffix as a 'definite' marker. More evidence of this will be provided in the following section.

We can look at other examples that support an analysis of the citation forms as indefinite NPs. The NPs in (15) and (16), as well as in (33) and (34), denote non-referential and non-identifiable entities, i.e. NPs that do not refer to a particular entity in the world or universe of discourse.[11]

(33) wóbs   yàà   bédà.
      wób-s   yà-à   béd-à
      elephant-CL13 COP-AFF big-CL6

      'Elephants are big.' (in general) (AT_2014-05-21_nasality, Ex.1)

(34) wóbʁ   yàà   bédré.
      wób-g   yà-à   béd-ré
      elephant-CL12 COP-AFF big-CL5

      'The elephant is a big animal.' (Lit. 'Elephant is big.') (AT_2014-05-21_nasality, Ex.4)

Another type of construction in which we find the indefinite forms of nouns are as part of predicating expressions that either (a) mark a referent as a member of a particular category, as in (35) and (36); or (b) predicate some property about a referent, as in (37)–(40).

(35) à músá   yàà   káràmbíígà.
      à=músá   yà-à   kàràmbíí-gà
      3SG.SUBJ=PN COP-AFF student-CL12

      'Moussa is a student.' (AT_2014-02-25_NominalPreds_Existential Ex.1)

---

[9] There does not appear to be an overt locative marker in this sentence.

[10] It is acknowledged that it is difficult to make claims about the mental state of the speaker and that psycholinguistic experiments may help to clarify this statement.

[11] Although (33) and (34) make a number distinction, it is debatable whether the propositions behind these two generic statements actually have different semantic truth values.

(36) à músá   nɛ́ àlì yáá   kàràmbíísì.
     à=músá   nɛ́ àlì yà-à   kàràmbíí-sì
     3SG.SUBJ=PN and PN COP-AFF student-CL13

     'Moussa and Ali are students.' (AT_2014-02-25_NominalPreds_Existential, Ex.3)

(37) ráàmà   yáá   tóóʁó.
     ráà-m-à   yà-à   tóó-gó
     drink-CL22-DEF COP-AFF bitter-CL15

     'The drink is bitter.' (KB_20140127_04_Non-verbal-clauses, Ex.5)

(38) màm   rá yàà   bédrè.
     màm   rá yà-à   béd-rè
     1SG.SUBJ PST COP-AFF big-CL5

     'I was/had been big/fat.' (MO_20140205_Session_5, Ex.12)

(39) tónd   rá yàà   bédà.
     tónd   rá yà-à   béd-à
     1PL.SUBJ PST COP-AFF big-CL6

     'We were/had been big/fat.' (based on MO_20140205_Session_5)

(40) súg   ròòdá   kwèlàms yáá   rààdó.
     sú-g   ròò-d-á   kwèlàm-s yà-à   ràà-dó
     thatch-CL15 house-CL21-DEF door-CL13 COP-AFF wood-CL21

     'The thatch house doors are (made of) wood.'
     (AT_2014-01-13_Women-carrying-pots-on-head_Text Ex.10-11)

In (35) and (36) , [kàràmbíígà] and [kàràmbíísì] represent the category of 'student'. In (37)–(39), words like [tóóʁó] 'bitter' and [bédrè] 'big' semantically assign a quality to a referent. In (40), [rààdó] designates the material from which the referents are made, and does not refer to particular pieces of wood. Importantly, although these words denoting properties behave formally like nouns, in that they take noun class markers, they cannot take the definite suffix /-a/ in these constructions.[12]

In summary, indefinite NPs in Mooré are expressed by the citation form of the nominal and are used to refer to non-identifiable entities, i.e. entities that are presumed to be unidentifiable to the listener, as well as to describe properties of subjects in predicate position. In contrast, we shall see that definite NPs take the suffix /-a/, and denote entities that are presumably identifiable to both the speaker and listener.

6 Definite NPs and contrastive focus

As mentioned in the previous section, once a referent has been established in discourse and is presumed by the speaker to be identifiable to the listener, subsequent mentions of

---

[12] However, these non-referential predicates are still sensitive to a formal singular / plural distinction. They agree with the subject for number, as seen in (35) and (36), as well as (38) and (39).

the NP take the definite /-a/ suffix. If we look at the words for 'bicycle' (in bold) in the following examples, we can see that the first mention of [wéèfò][13] in the narrative is in the indefinite, but subsequent occurrances take the definite suffix /-a/. Similarly, the first mention of the 'rock' [kúgrì] is in the indefinite, but the next mention takes the definite suffix.

(41) bííg   hǎ zɔ́ **wéèfò**   rà pyóʁdá   tììʁá   sèèʁá.
      bíí-g   hǎ zɔ̀ wèè-fó   rà pyóg-d-á tìì-g-á   sèè-gá
      child-CL12 REL stand bicycle-CL19 PST PASS-PROG-AFF tree-CL12-DEF side-CL12

'A kid who was sitting on a bicycle passed by the tree.' (Pear Story, Ln22-23)

(42) tá   **wèèfá**   tɔ́ɔ̀r   tì   zùm kúgrì   **wèèfá**   hǎ
      tɪ=á   wèè-f-á   tɔ́ɔ̀-r   tì   zùm kúg-rì   wèè-f-á   hǎ
      CONN=3SG bicycle-CL19-DEF front-CL5 CONN sit rock-CL5 bicycle-CL19-DEF REL

      zùm kúgrà,   **wéèfà**   lwíî   mè.
      zùm kúg-r-à   wèè-f-á   lwí-ì   mè
      sit rock-CL5-DEF bicycle-CL19-DEF fall-AFF CF

'Then the front of his bicycle was on a rock. When the bicycle was on the rock, it fell down.' (Pear Story, Ln40-44)

More examples of definite forms of the noun occurring as either subject or object are given in (43)–(45).

(43) à músá   né àlì yáá   kàràmbíísà.
      à=músá   né àlì yà-à   kàràmbíí-s-à
      3SG.SUBJ=PN and PN COP-AFF student-CL13-DEF

'Moussa and Ali are the students (that the speaker and listener know about).' (AT_2014-06-04)

(44) wóbsà   yáá   bédà.
      wób-s-à   yà-à   béd-à
      elephant-CL13-DEF COP-AFF big-CL6

'The elephants (the speaker and listener know about) are big.' (AT_2014-05-21_nasality)

(45) wóbsà   ríî   tîìsà.
      wób-s-à   rí-ì   tìì-s-à
      elephant-CL13-DEF eat-AFF tree-CL13-DEF

'The elephants (the speaker and listener know about) ate the trees (the speaker and listener know about).' (AT_2014-05-28)

---

[13] The tones in [wéèfò] are the result of a tone sandhi rule, such that the underlying low tone on the first syllable of /wèèfó/ is realized as falling [wéèfò] when preceded by a word ending in low tone.

The question therefore is if /-a/ is the definite suffix, what about /-wã/ ~ /-ã/, as claimed in previous analyses? In data from elicitation[14] it appears that the use of the nasalized suffix is associated with a contrastive focus reading (indicated by underlining in examples). For example, in one reading of (46), the speaker presupposes that the listener believes some other entity to have eaten the trees. The nasalized suffix therefore highlights the argument 'the elephants' to correct this belief. However, even with the nasalized suffix, it is still possible to interpret the sentence without contrastive focus. It simply marked an entity that the speaker presumes is identifiable to the listener. Certainly, more data from naturalistic data would help to clarify this function of the nasalized suffix.

(46) wóbsã[15]   rîì   tîìsà.
     wób-s-ã   rí-ì   tìì-s-à
     elephant-CL13-*ã* eat-AFF tree-CL13-DEF

     'The elephants (the speaker and listener know about) ate the trees (the speaker and listener know about)'; or '**The elephants** (the speaker and listener know about) (not something else) ate the trees (the speaker and listener know about)' (e.g. it wasn't the mice that ate the trees). (AT_2014-05-28)

In (47), the nasalized suffix also marks contrastive focus and either has scope over the argument 'the elephant' or over the entire predicate. At present, it is unclear if a reading without focus is available for this sentence and more naturalistic data is certainly needed. Nevertheless, it should still be noted that a contrastive focus reading is not available in (48), where [wóbʁà] does not take the nasalized suffix.

(47) màm   nɛ́ɛ̀   wóbʁã.
     màm   nɛ́-ɛ̀   wób-g-ã
     1SG.SUBJ see-AFF elephant-CL15-*ã*

     'I see **the elephant**' (as opposed to seeing something else); or 'I **see the elephant**' (as opposed to not seeing the elephant). (AT_20140605)

(48) màm   nɛ́ɛ̀   wóbʁà.
     màm   nɛ́-ɛ̀   wób-g-à
     1SG.SUBJ see-AFF elephant-CL15-DEF

     'I see the elephant (that the speaker and listener talked about).' (AT_20140605)

In the examples given above, the citation form noun class suffix vowel on the nouns is not /a/, e.g. /wób-gò/ 'elephant-CL15', /wób-sè/ 'elephant-CL13'. In contrast, when the citation form noun class suffix vowel is /a/, e.g. /bʊ́ʊ́-gá/ 'goat-CL12', there is a stronger preference for the speaker to use the nasalized /-ã/ simply to indicate identifiability; compare (49) and (50). This does not necessarily place any contrast or focus on the argument.

---

[14] Although examples from semi-elicited narratives would be ideal, the analysis of such data is complicated by the presence of a connective nasal morpheme that appears on the final element of clause in running speech with the meaning 'and then'.

[15] It is unclear what the tone on this suffix is.

(49) màm   nɛ́ɛ̀   bʊ́ʊ́ʁã̌.
màm   né-ɛ̀   bʊ́ʊ́-g-ã
1SG.SUBJ see-AFF goat-CL12-*ã̌*

'I see the goat' (that we talked about); or 'I see **the goat**' (as opposed to seeing something else); or 'I **see the goat**' (as opposed to not seeing the goat).
(AT_20140605)

(50) màm   nɛ́ɛ̀   bʊ́ʊ́ʁá.
màm   né-ɛ̀   bʊ́ʊ́-gá
1SG.SUBJ see-AFF goat-CL12

'I see a goat.' (AT_20140605)

When asked about (50), the consultant stated that [bʊ́ʊ́ʁá] 'goat' in this example could only refer to a goat that has not been mentioned before; however, there is evidence from other examples – both from elicitation such as (51), and from narrative discourse such as (52) – that even such NPs without the nasal marker can be interpreted as being identifiable.

(51) à músá   yàà   káràmbíígà /   káràmbíígã.
à=músá   yà-à   kàràmbíí-g-à   kàràmbíí-g-ã
3SG.SUBJ=PN COP-AFF student-CL12-DEF student-CL12-*ã̌*

'Moussa is the student (we know about).' (nasality preferred)
(AT_2014-03-11_NominalPreds, Ex.34)

(52) páʁá   búkà   bíígá.
pág-á   búk-à   bíí-gá
woman-DEF carry-AFF child-CL12

'The woman (mentioned previously) is carrying a child.'
(AT_2014-01-13_WomanDonkeyCart_Text, Ex2)

Importantly, monosyllabic nouns that do not have a noun class marker in citation form (or, which have a zero-form class marker), e.g. /kí/ 'millet', can only be marked for identifiability with the nasalized suffix /-wã/ as in (53) (and not with /-a/, /-ã/ or /-wa/). This nasalized form may also give the possibility of a contrastive focus reading.

(53) kíwã   yáá   kʊ̌drè.
kí-wã   yà-à   kʊ́d-rè
millet-*wã* COP-AFF old-CL5

'The millet (we know about) is old'; or '**The millet** (not something else) is old.'
(AT_2014-06-04)

For NPs ending in a demonstrative modifier, the nasal /-ã/ suffix occurs on the demonstrative (i.e. on the final element of the NP), where it also marks contrastive focus, as in (54) and (55).

(54) wób   káŋɑ́   yàà   bédré.
      wób   kán-g-ɑ́   yà-à   béd-ré
      elephant DEM-CL12-*ã* COP-AFF big-CL5

      '**This elephant** is big.' (e.g. correcting someone who has said 'This mouse is big.')
      (AT_2014-05-21_nasality, Ex.30)

(55) nù(g) káŋɑ́   yàà   bédrè.
      nù(g) kán-g-ɑ́   yà-à   béd-rè
      hand DEM-CL12-*ã* COP-AFF big-CL5

      '**This hand** is the one that is big.' (not something else)
      (KB_20140224_08_Noun-phrase, Ex.2)

The data raise two interesting points. Firstly, the marking of identifiability on NPs is dependent on the phonological shape of the head, or rather final noun in the phrase, specifically the vowel (if any) in the noun class suffix. The use of nasality may be necessary in some forms to help disambiguate the definite forms from their citation forms, or perhaps when the speaker feels the need to disambiguate the two. Secondly, the formal marking of identifiability in Mooré is intimately linked to what we might call FOCUS. Using Lambrecht (1994)'s terminology, we typically find the nasalized suffixes /-wã/ ~ /-ã/ not just marking identifiability, but also ARGUMENT FOCUS where the subject or object is being contrasted with another possible referent, and sometimes even PREDICATE FOCUS, where the entire predicate is being contrasted with another one.

This then brings us to the question of whether the nasal feature of /-wã/ ~ /-ã/ can be considered a separate morpheme from /-a/. A preliminary analysis suggests that contrastive focus on indefinite NPs can be marked with a nasal feature (indicated in the second line of the examples just by a tilde) without the vowel /-a/. For example, in (56) and (57), where there is no /a/ to carry the nasalization, the nasal feature is often heard as a nasalized schwa at the end of the NP in (56); or sometimes as a full syllabic nasal consonant that is phonologically bound to the following verb in (57).

(56) wóbsə̃̀   yá bédà.
      wób-s-˜   yà béd-à
      elephant-CL13-˜ COP big-CL6

      '**Elephants** are big (in general).' (e.g. correcting someone who said 'Mice are big.')
      (AT_2014-05-21_nasality, Ex.16)

(57) yű̀ysə̃̀ ~yű̀ysn dit   sáʁbó.
      yű̀y-s-˜   ríd-d[16]   ság-bó
      mouse-CL13-˜ eat-PROG dough-CL14

      '**Mice** eat *sagbo* (a kind of doughy food).' (e.g. correcting someone who said, 'Elephants eat *sagbo*.') (AT_2014-05-28)

---

[16] It is not clear why the consultant did not produce the suffix /-a/ 'AFF' in this example. Perhaps this is the case when focus is placed on the argument preceding the verb, and may reflect the historical origin of this construction coming from a cleft construction with a relativized clause.

The presence of nasality without the /a/ vowel (i.e. the [ə̃] form) in some constructions suggests that /-ã/ could be analyzed as two separate morphemes: /-a/ plus nasalization.

In summary, although /-wã/ ~ /-ã/ are always found on identifiable NPs, the full range of data examined suggests that previous analyses of these suffix forms as the "definite" marker are not entirely accurate, at least for the speech of some Mooré speakers. Rather, the nasalized forms play a role in marking contrastive focus, and for some nouns non-nasalized /-a/ is the indicator of non-contrastive identifiability. Thus, it may be more appropriate to think of the /-a/ on NPs as the definite marker (DEF), and the nasal /˜/ as a focus marker (FOC) that gives the *potential* for some sort of contrastive focus reading. These forms represent two separate structural categories that map onto different semantic and pragmatic functions depending on the construction and the context.

7 Complications to the analysis of identifiability and focus

This study of identifiability marking of NPs is still preliminary, as more texts need to be collected and analyzed. One complication here, however, is that some non-final clauses in a chain in Mooré are marked with a nasal connective morpheme. It is therefore not always possible to tell if the nasality on [bʊ́ʊ́ʁã] in the first line of (59) marks definiteness and focus on the NP, or if it marks non-finality of the clause, or both.

(58) ráw hn̄ tákd̀   bʊ́ʊ́ʁá   píòʁà   tĩ̀ʁà   sɛ́ɛ̀ʁà.
      ráw hn̄ ták-d̀   bʊ́ʊ́-gá   píòg-à   tìì-g-á   sɛ̀ɛ̀-gá
      man REL pull-PROG goat-CL12 pass-AFF tree-CL12-DEF side-CL12[17]

      'A man who was pulling a goat passed by the tree.' (Pear Story, Ln18-19)

(59) ráwá   ták bʊ́ʊ́ʁã   píòʁ tĩ̀ʁà   sɛ́ɛ̀ʁà   lâ lóóʁè.
      ráw-á   ták bʊ́ʊ́-g-ã   píòg tìì-g-á sɛ̀ɛ̀-gá   lâ lóóg-è
      man-DEF pull goat-CL12-DEF/CONN? pass tree-CL12-DEF side-CL12 and leave-?

      'The man pulled the goat, passed by the tree and left.' (Pear Story, Ln20)

There are also examples from texts where the indefinite form of a noun is used to refer to an entity that has already been established in discourse. For example, in a story about a Rabbit and Hyena who go off in search of honey, we find in the second-to-the-last line the form [à sĩ̀dò] 'his honey', as shown in (60). In most of the preceding text, the honey has been referred to in the definite form [sɪɪdà]. Perhaps the overt definite marker is not necessary with possessed NPs, but this is something that needs to be checked.

(60) … n dík à   sĩ̀dò.
      … n=dík=à   sĩ̀-dò
      CONN=take=3SG honey-CL21

      '… and took his honey.' (DP_Rabbit & Hyena August 2013, Ln.20)

---

[17] In relator noun constructions, the relator noun comes after the head noun, but it is the head noun that takes the definite suffix /-a/.

Furthermore, the analysis has only looked at the coding of identifiability on NPs with common nouns as heads. Other kinds of NPs have been ignored, including NPs headed by borrowed nouns and NPs that refer to unique entities (entities that are referential and presumably identifiable to the listener). For example, [prezidↃ] 'the president (of Burkina Faso)' in (61), is not formally coded with the definite suffix. The addition of the nasal *-wã*, as in (62), necessarily leads to a contrastive focus interpretation of the argument.

(61) prezidↃ   wád-dà=mè.
     prezidↃ   wád-dà=mè
     president come-AFF=CF

   'The President (of Burkina Faso) is coming.' (assuming one is in Burkina Faso, so it is clear which president you are talking about) (AT_2014-06-04)

(62) prezidↃ-wã   wátàmè.
     prezidↃ-wã   wád-dà=mè
     president-FOC come-AFF=CF

   '**The President** is coming' (not anyone else). (AT_2014-06-04)

As a final point, it should be mentioned that in addition to morphological marking, arguments in Mooré can be brought into focus by syntactic means. For example, word order can be manipulated to bring an argument to the front of a clause, as in (63).

(63) kàràmbíígà   yáá   máám.
     kàràmbíí-g-à   yà-à   máám
     student-CL12-DEF COP-AFF 1SG.O

   'The student (that we were talking about) is me.' (AT_2014-03-04_NominalPreds, Ex.15)

In addition, there is a construction containing the copula *la*, which marks the subject for greater emphasis: [SUBJECT *la* PREDICATE]. This is illustrated in (64) and (65). This is a more structurally marked construction, especially in (64) where the object form of the pronoun appears preverbally, as one would normally expect it to appear post-verbally. In (65), the alternate *kàràmbíígã* is preferred, though the form without the final nasalization is also possible.

(64) máám lá   kàràmbíígà.
     máám lá   kàràmbíí-g-à
     1SG.O COP student-CL12-DEF

   '**I** am the (one) student' (not anyone else). (AT_2014-03-04_NominalPreds, Ex.16)

(65) à=músá   lá kàràmbíígà /   kàràmbíígã.
     à=músá   lá kàràmbíí-g-à   kàràmbíí-g-a-˜
     3SG.SUBJ=PN COP student-CL12-DEF student-CL12-DEF-FOC

   '**Moussa** is the (one) student' (not anyone else). (AT_2014-03-11_NominalPreds, Ex.31)

8 Summary and further research

It is hoped that this new analysis of /-a/ as a definite marker in Mooré will serve as a useful update to previous analyses (e.g. Peterson 1971; Canu 1974) that have only considered the nasal forms /-wã/ ~ /-ã/. The paper has also argued that /-ã/ could be treated as bimorphemic, with the nasal element marking (potentially contrastive) focus. However, it is more difficult to morphologically analyze /-wã/, which is used with monosyllabic nouns like /kí/ 'millet'. Most importantly, it has shown that the three forms /-a/, /-ã/ and /-wã/ may share similar semantic / pragmatic functions of coding identifiability, but this depends on the underlying phonological form of the noun class suffix.

The link between identifiability and the marking of contrastive focus is particularly interesting, though perhaps not too surprising given that the marking of definiteness is also associated with speaker and listener attention. In Dagbani, a related Gur language, Olawsky (1999: 40) reports two "definite articles": *la* and *maa*, with the latter described as a "strong" article that adds "more emphasis" than the former.[18] In general, the languages of West Africa look like a fertile place to examine the interaction between morphosyntax and information structure. Previous studies of these languages include Fiedler & Schwarz (2005) that looks at the "non-focal" part of sentences[19] in five Kwa and Gur languages, as well as Schwarz (2010) on argument and predicate focus in four Gur languages; neither of these include data on Mooré.

Future studies could also investigate how syntactic focus-marking strategies interact with the use of the nasal focus morpheme. For instance, Kabore (1985: 445-450) looks at a number of constructions that pertain to the marking of contrast and emphasis, including what he writes as a construction that uses /sĩ/ which seems to corresponds to a /hã/ subordinator / relativizer in the data collected here. Indeed, further studies of identifiability and focus in Mooré will need to consider relativization strategies.

The study of information structure calls for a move away from elicited data and demands much more work in the analysis of texts collected from a wide range of speakers across different genres. More importantly, it suggests the need for greater experimental work that considers the role of speaker and listener attention and which provides ways of describing cognitive categories *independent* of the linguistic correlates that linguists traditionally use as evidence for such cognitive categories. It would seem that we are just barely beginning to scratch the surface when it comes to understanding the notions of *referentiality*, *identifiability* and *contrastive focus* and how they might be coded by linguistic expressions.

Acknowledgments

---

[18] Kropp Dakubu (1991), cited in Olawsky (1999: 30), suggests that meaning of a particle *la* in Dagaare and Gurune appears to be related to definiteness and new information or focus.

[19] This appears to be similar to Lambrecht (1994)'s notions of "predicate focus" and "sentence focus".

Abbreviations

1,2,3	1st, 2nd, 3rd person	FOC	argument focus
AFF	affirmative (on the first verb in a clause chain / verb in a single clause)	FUT	future tense
		NEG	negative
		O	object form (of pronoun)
CF	affirmative clause-final marker	OBJ	pronominal object suffix
CFN	negative clause-final marker	PL	plural
CL	noun class marker	PROG	progressive
CONN	connective	PST	past tense
COP	copula	REL	relativizer
DEF	definite	Q	question
DEM	demonstrative	SUBJ	subject form (of pronoun)
EMPH	emphatic pronoun	SG	singular

References

Canu, Gaston. 1974. *Description synchronique de la langue mòòré.* Lille: Université de Lille III. (Doctoral dissertation.)

Canu, Gaston. 1976. *La langue mò: Rē: Dialecte de Ouagadougou, Haute-Volta: Description synchronique* (Langues et Civilisations à Tradition Orale 16). Paris: Société d'Études Linguistiques et Anthropologiques de France.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology, 2nd edn.* Chicago: University of Chicago Press.

Du Bois, John W. 1980. Beyond definiteness: The trace of identity in discourse. In Wallace L. Chafe (ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, 203–274. Norwood, NJ: Ablex.

Fiedler, Ines & Anne Schwarz. 2005. Out-of-focus encoding in Gur and Kwa. In Shin Ishihara, Michaela Schmitz & Anne Schwarz (eds.), *Approaches and findings in oral, written and gestural language* (Interdisciplinary Studies on Information Structure 3), 111–142. Potsdam: Universitätsverlag Potsdam.

Givón, Talmy. 2001. *Syntax: An introduction, 2 vols.* Amsterdam/Philadelphia: John Benjamins.

Kabore, Raphael. 1985. *Essay d'analyse de la langue mʋʋre (parler de Wàogdgò).* Paris: Université de Paris 7, Département de Recherches Linguistiques, Laboratoire de Linguistique Formelle.

Kropp Dakubu, Mary Esther. 1991. A note on the particle LA in Northwestern Oti-Volta. Accra: Linguistics Association of Ghana. Ms.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig. 2013. *Ethnologue: Languages of the world*. 17th edn. Dallas, TX: SIL International. http://www.ethnologue.com, accessed 2014-03-14.

Nikièma, Norbert. 1989. Du non cumul des fonctions classificatoires et quantificatives dans les suffixes nominaux du moore. *Journal of West African Languages* 19(2). 93–124.

Olawsky, Knut J. 1999. *Aspects of Dagbani grammar: With special emphasis on phonology and morphology*. Munich: LINCOM Europa.

Peterson, Thomas H. 1971. *Mooré structure: A generative analysis of the tonal system and aspects of the syntax*. Los Angeles: UCLA. (Doctoral dissertation.)

Schwarz, Anne. 2010. Verb-and-predication focus markers in Gur. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure. A documentation of its diversity across Africa*, 287–314. Amsterdam: John Benjamins.

Chapter 4

The syntactic status of objects in Mooré ditransitive constructions

Sara Pacchiarotti

University of Oregon

This paper offers a structural description of the overt and covert properties of objects in Mooré (Gur, Burkina Faso) ditransitive constructions, along with a list of verbs which can be classified as ditransitive in this language on the basis of specific syntactic properties. The paper evaluates whether object relations in Mooré ditransitive constructions can be characterized according to proposals about the typology of object relations present in the literature, i.e. primary vs. secondary object type language or symmetrical vs. asymmetrical object type language.

1 Introduction

Mooré (Mòòré) [mos] is a Gur language spoken in Burkina Faso by approximately 6 million people (Lewis, Simons & Fennig 2016).[1] Ditransitive constructions in Mooré resemble what have been called DOUBLE OBJECT constructions (Dryer 1986; 2007; Goldberg 1995). There are several proposals for typological classification of the object systems of languages which allow two objects in a construction.

Dryer (1986) argues that while many languages employ object grammatical relations which can be best described as DIRECT OBJECT (DO) and INDIRECT OBJECT (IO), other languages use the grammatical relations of PRIMARY OBJECT (PO) versus SECONDARY OBJECT (SO). In the latter case, the PO has morphosyntactic properties similar to those of a Direct Object (DO) in a monotransitive clause. The SO, on the other hand, does not display the same object properties as the PO. In dealing with double-object constructions in Bantu languages, Bresnan & Moshi (1990) introduce the concepts of SYMMETRICAL versus ASYMMETRICAL object type languages. In symmetrical object languages, the THEME

---

[1] The data for this paper comes from elicitation at the University of Oregon with Timbwaoga Aimé Judicaël Ouermi, a 25 year-old male native Mooré speaker born and raised in Ouagadougou, Burkina Faso and residing in Eugene, Oregon since 2009.

object of a causative or ditransitive (possibly created by means of an applicative) verb retains its object properties in the presence of another object. In such a language, both objects *simultaneously* display the same object properties. In an asymmetrical object language, on the other hand, only one of the objects displays the full array of object properties.

The aims of this paper are twofold. First, it offers a structural description of the overt and covert properties of objects in Mooré ditransitive constructions, along with a list of verbs which can be classified as ditransitive in this language on the basis of specific syntactic properties displayed when the verb occurs in a ditransitive construction. This description relies heavily on the parameters proposed by Malchukov, Haspelmath & Comrie (2010) in their cross-linguistic study of ditransitive constructions,[2] but it also takes into account object properties proposed by Hyman & Duranti (1982). Second, the paper evaluates whether object relations in Mooré ditransitive constructions can be characterized according to major proposals about the typology of object relations present in the literature. The former goal is a contribution to the extant grammatical descriptions of the language (Alexandre 1953; Canu 1974; Peterson 1971; Kouraogo 1976; Kabore 1985; *inter alia*) which only marginally deal with ditransitive constructions (see Canu 1974; Kabore 1985). The latter goal represents a contribution to the typological literature on ditransitive constructions cross-linguistically and on the grammatical relation of object in general. To my knowledge, no attempt has been made so far to describe object relations of Gur languages from a typological perspective.

The paper is organized as follows: §2 sets out relevant definitions and terminology; §3 presents a list of verbs which can be defined syntactically as ditransitives in Mooré; §4 presents overt properties of objects in Mooré ditransitive constructions; §5 deals with covert properties of objects in Mooré ditransitive constructions; §6 determines whether, based on overt and covert properties, the grammatical relation of object in Mooré ditransitive constructions fits any of the proposals present in the literature; §7 concludes the paper.

2 Definitions and terminology

In Mooré, a ᴅɪᴛʀᴀɴsɪᴛɪᴠᴇ ᴄᴏɴsᴛʀᴜᴄᴛɪᴏɴ contains the following structural components: a subject, a verb and two objects (see Olawsky 1999 for the same structure in Dagbani). The two objects which follow a ditransitive verb are not morphologically or analytically

---

[2] In this paper, I use the term *ditransitive constructions* in the sense of Malchukov, Haspelmath & Comrie (2010). According to these authors, ditransitive constructions are defined semantically as constructions in which a verb denotes the transfer of an entity (T) from an ᴀɢᴇɴᴛ (A) to a ʀᴇᴄɪᴘɪᴇɴᴛ (R) (Malchukov, Haspelmath & Comrie 2010: 1). Malchukov, Haspelmath & Comrie (2010), unlike Goldberg (1995), include under the term *ditransitive construction* different types of alignments of the two object arguments (i.e. indirective, secundative and neutral). Therefore, the D(irect) O(bject)-I(ndirect) O(bject) alignment type (i.e. indirective) is also an instance of a ditransitive construction under this definition.

marked, that is, they do not show anything like object case-marking morphology and they are not introduced by any adposition or relator noun, as in (1).[3]

(1) màm tóól-à   pág-ã̀   rú-kã̀.
     1SG.I send-AFF woman-CL1.IDN pot-CL12.IDN
     'I sent the pot to the woman.'

In (1), the semantic role of 'woman' is that of Recipient (R). The semantic role of 'pot' is that of Theme (T). According to Kittilä (2005: 274), Recipients can be defined as "animate entities that receive something concrete transferred to their sphere of control or domain of possession". The Theme is the item being transferred to the animate receiver. Throughout the paper, I will use O_R to refer to the syntactic object mapped onto the semantic role of Recipient and O_T to refer to the syntactic object mapped onto the semantic role of Theme.

Note that when O_T is inanimate and O_R is animate/human, as in (1), O_T cannot precede O_R, as shown by the ungrammaticality of (2).

(2) * màm tóól-à   rú-kã̀   pág-ã̀.
     1SG.I send-AFF pot-CL1.IDN woman-CL12.IDN
     ('I sent the pot to the woman.')

In Mooré, a DITRANSITIVE VERB is defined as displaying the following syntactic features: (i) it can or must appear in a construction followed by two morphologically and analytically unmarked objects and (ii) *both* of its objects can be expressed by means of optional bound pronominal marking in the verb, although not simultaneously. Objects can be optionally indexed in the verb regardless of their animacy value (i.e. inanimate, animate, and human objects can all be indexed).

In Mooré, the grammatical relation of object (O), in both monotransitive and ditransitive clauses, displays the following overt syntactic properties: (i) a lexical or free pronominal expression of it occurs immediately after the verb or after another object; (ii) such a phrase is not introduced by prepositions or relator nouns; (iii) such a phrase does not bear any case marking;[4] and (iv) such a phrase can be optionally indexed in the verb. Verb indexation of an object and lexical/free pronominal expression of a co-referential object can never co-occur in the same clause. For reasons of space, these properties will be illustrated only for ditransitive clauses. Additional properties, especially covert

---

[3] The Mooré orthography used in this paper is grounded in a phonemic representation using IPA characters, except [j] which is represented here as <y>. The vowel system is extremely complex and how many phonological vowels exist and whether or not there is some type of harmony (ATR or otherwise) are debated topics (see Peterson 1971; Canu 1974; Rennison 1990; Nikièma 2000; Calamai & Bertinetto 2005; among others). For the purposes of this paper, vowel graphemes and their phonetic realizations are as follows: <i> [i], <ɪ> [i̧], <e> [e, ɪ], <ɛ> [ɛ], <a> [a, ə], <ɔ> [ɔ], <o> [o], <ʊ> [ʊ] and <u> [u]. Tone: high=v́, low=v̀ (though low is not written in the language name). The tone system is not fully understood at the present time, but see Peterson (1971).

[4] Case marking on nouns is absent in Mooré, except for the putative locative case endings -ẽ̀ and -ṍwã̀ as in *zák-ẽ̀, zák-ṍwã̀* 'in the house'. The difference between the two forms is obscure at the present time.

ones, will be discussed throughout the paper and summarized in Table 4. Some of the morphosyntactic details necessarily relied on in §3 will be dealt with in more depth in §4.

3 Ditransitive verbs in Mooré

There seems to be a restricted number of ditransitive verbs in Mooré (Table 1). A DITRANSITIVE VERB in this language is one which displays the following syntactic features: (i) it can appear in a construction followed by two morphologically unmarked NPs and (ii) either of its objects can be optionally indexed in the verb, although not simultaneously.[5]

Table 1: Mooré ditransitive verbs

Verb	Meaning	Verb	Meaning
tóólè	'send'	*kɔ́*	'give'
pénè	'lend/borrow'	*bɔ́sè*	'ask for'
wínígì	'show'	*tɔ́gsè*	'tell'
kóósè	'sell'	*rílgè*	'feed'
yǔngì	'make drink'	*kéllè*	'leave'
zámsè	'teach'	*lóbgè*	'throw (to)'

Some ditransitive verbs present special lexical (selectional) restrictions. For instance, the verb *kéllè* 'leave' can only be used when the THEME is inanimate (artifacts or dead animals but not living animals or humans). A couple of ditransitive verbs are synchronically segmentable morphological causatives: this is the case of *rílgè* 'feed' derived by means of the causative morpheme *-g* from *rí* 'eat', and *yǔngì* 'make drink' derived from *yǔ* 'drink'. The verb *wínígì* 'show' has the phonological form of a causative verb but presuming this was its origin, the verb has completed a process of lexicalization and the 'source' verb root from which it was derived is no longer synchronically present in the language.

Several ditransitive verbs can occur in constructions with different argument structures and a concomitant change in their semantic meaning.[6] This is illustrated with 'throw' in (3) and (4):

---

[5] Several semantically ditransitive verb concepts such as 'bring (someone or something somewhere)' have been omitted for two reasons. First, some of these verbal concepts are expressed in multi-verb constructions. Second, some of these verbal concepts are found in constructions which include a SPATIAL GOAL or LOCATION, such as 'to the house', 'to the market', or 'to Ouagadougou'. As will be discussed in §4.4, locative NPs in Mooré are marked by a putative locative case marker and they cannot be optionally indexed in the verb.

[6] Other examples include the verb *zámsè*, which can mean 'learn', 'teach' or 'dream' depending on the argument structure it occurs with.

(3) à   lóbg-à   máám kúg-r-ȁ̀.
3SG.I throw-AFF 1SG.II stone-CL5-IDN

'She threw the stone to me (I am trying to catch it).'

(4) à   lóbg-à   máám nɛ́   kúg-rì.
3SG.I throw-AFF 1SG.II with stone-CL5

'She threw the stone at me (she wants to hit me with it).'

The ditransitive argument structure in (3) is benefactive, whereas (4) is malefactive. The only difference between the two is that 'stone' in (4) is introduced as an oblique by means of the instrumental/comitative preposition 'with', whereas in (3) it appears as a core argument.

Besides their occurrence in a ditransitive construction, the major syntactic criterion for claiming that a verb is ditransitive in Mooré is the possibility of optionally indexing either of its objects on the verb, albeit not simultaneously. Many verbs which formally look like good candidates for the ditransitive label were excluded due to their failure at the indexation test. This can be illustrated with verbs such as *wí* 'bring water', *rá* 'buy' and *mwɛ́ɛ̂* 'build'. For instance, 'build' can appear in a ditransitive construction (5) or in the *ń kɔ́* 'and give' construction (6), used to express an NP with the semantic role of BENEFACTIVE.[7]

(5) à   Músá mwɛ́-ɛ̂   máám zá-kà.
3SG.I M.   build-AFF 1SG.II house-CL12

'Musa built a house for me' (lit: 'Musa built me a house.').

(6) à   Músá mwɛ́-ɛ̂   zá-kà   ń   kɔ́   máám.
3SG.I M.   build-AFF house-CL12 PVC give 1SG.II

'Musa built a house for me.'

Although 'build' can appear in a ditransitive construction such as (5), this verb has restrictions on the verbal indexation of one of its objects. Specifically, the optional indexation of O_R on 'build' is possible (7), but the indexation of O_T is never possible (8).

(7) à   Músá mwɛ́-m-là   zá-kà.
3SG.I M.   build-1SG.III-AFF house-CL12

'Musa built me a house.'

(8) *à   Músá mwɛ́-ɛ̂-là   máám.
3SG.I M.   build-3SG.III-AFF 1SG.II

('Musa built it for me.')

---

[7] The *ń kɔ́* construction illustrates grammaticalization of the verb *kɔ́* 'give' into a benefactive marker.

With the ditransitive R plus T meaning, the O_T 'house' can be successfully indexed only if the verb is followed by the main clause final marker -*mɛ̀*[8] and the Recipient/Bene-factive is expressed by means of the *ń kɔ́* construction (9):

(9) à   mwɛ̃́-ɛ̀-là-mɛ̀   ń   kɔ́ máám.
     3sg.i build-3sg.iii-aff-phb pvc give 1sg.ii

     'He built it for me.'

The verbs *wí* 'bring water' and *rá* 'build' display the same restrictions illustrated for 'build' in (8): the indexation of O_T is impossible if they are used in a ditransitive frame and possible only in a construction such as (9). The indexation test proposed here differentiates ditransitive verbs (i.e. those listed in Table 1) from ditransitive clauses (i.e. as in (5)), showing the relevance of morphosyntactic transitivity at both levels for this language.

4 Overt properties of ditransitive constructions

This section illustrates the alignment system of Mooré, including object alignment in the sense of Malchukov, Haspelmath & Comrie (2010) and overt properties of ditransitive constructions such as word order, constituency and clause-level person agreement or indexing.

4.1 Alignment in transitive and ditransitive clauses

Mooré displays a nominative-accusative alignment system in most areas of its grammar and therefore an emic category of Subject (S/A). Word order is rigid both in intransitive (SV) and transitive clauses (AVO). There appear to be three sets of pronouns: one for the S/A category (set I which has both long and short form pronouns) and two for the O category (sets II[9] and III). The O argument of a monotransitive verb and O_R and O_T of a ditransitive verb can be optionally indexed in the verb by the set III forms (see §4.4 for details of their use). When an optional bound pronominal form occurs, if the polarity of the clause is affirmative, the set III pronoun must co-occur with the affirmative marker -*là*, which is an allomorph of -*à*.[10]

---

[8] The obligatory presence of -*mɛ̀* before the *ń kɔ́* construction is syntactic evidence that the pre-verbal conjunction *ń* actually introduces a second clause (at least historically) and that 'me' is (or was) not simply an oblique argument. If the latter were the case, -*mɛ̀* would not be present because -*mɛ̀* only appears when no core or oblique arguments or adverbs follow a given verb within the same (main) clause.

[9] The dubious status of set II plural forms as syntactically accusative forms will be discussed in §4.3.

[10] In the present paper I gloss -*là* as 'affirmative' (aff), and treat it as an allomorph of -*à*. The affirmative marker -*à* obligatorily appears on the verb if, and only if, the verb is in an affirmative declarative main clause. The allomorph -*là* occurs only when a set III object pronoun occurs in the verb. This same analysis was proposed by Manessy (1963). The appropriate glossing of this morpheme is, however, controversial. Alexandre (1953: 96) calls it a 'marker of indicative mood'; Peterson (1971: 112) a 'complement marker'; Canu (1974) a 'marker of realis mood'; Kabore (1985) a 'marker of modality'; and Nikièma (2003) a mark of the 'effectiveness and declaration of the process'.

The pronoun sets are shown in Table 2. It should be noted that several authors have diverging opinions about the characterization of Mooré pronouns. Though there are some differences, the account offered here is largely the same as Kouraogo's (1976: 53), who claims that set I is used as a nominative set, whereas sets II and III are accusative sets. In contrast, Canu (1974) and Kabore (1985) argue that both sets I and II can express any syntactic role and choice lies mainly in the emphatic use of the pronoun.[11]

Table 2: Mooré pronoun sets

| | SET I (S/A) | | SET II (O) | SET III (O) |
	long	short	long	short
1sg	*màm*	*m* [m̩]	*máám*	*-m-là*
2sg	*fò*	*f* [ˀf]	*fóó*	*-f-là*
3sg	*à*	*à* [à]	*yéndá/yé̌*	*-à-là*
1pl	*tónd*	*d* [ˀd]	*tónd(ò)*	*-d-là*
2pl	*yǎmb*	*b* [ˀb]	*yǎmb(à)*	*-í̌-là*
3pl	*bámb*	*b* [ˀb]	*bámb(á)*	*-b-là*

The following examples show that the same set I 3sg pronoun codes the S argument (10) and the A argument (11). The O argument, however, is coded by a set II 3sg pronoun (11).

(10) à   kí-ì-mè.
      3sg.ɪ die-ᴀꜰꜰ-ᴘʜʙ
      'She died.'

(11) à   kɔ́-ɔ̀   yéndà.
      3sg.ɪ kill-ᴀꜰꜰ 3sg.ɪɪ
      'She killed him.'

Notice in (10) and (11) that the affirmative marker *-à* has allomorphs which are copies of the vowel of the verb root when the verb root is CV, as in *kí* 'die' and *kɔ́* 'kill'. This same behavior can be observed in (12) and (13) with the verb *kɔ̃́* 'give'. The verb phrase boundary marker *-mè*, present in (10), occurs only when the clause is declarative, its polarity affirmative and no NP functioning as a core or oblique argument, nor adverb, follows a given verb.

In ditransitive constructions, Mooré displays a neutral-object alignment system, as defined by Malchukov, Haspelmath & Comrie (2010), in terms of pronoun form: the O of a monotransitive verb (11), and O$_R$ (12), and O$_T$ (13) of a ditransitive verb are expressed formally by the same set of pronouns.[12]

---

[11] Specifically, Kabore (1985: 220 and ff.) uses the term *valeur d'insistance* to differentiate sets I and II. Set II would be used, according to him, when the pronoun is focused or emphasized.

[12] Examples (12) and (13) are to be understood as pronominal counterparts of (1), 'I gave the pot to the woman'.

(12) màm kɔ́-ɔ̀   yɛ́ndà rú-kà̃.
      1SG.I give-AFF 3SG.II pot-CL12.IDN

      'I gave the pot to her.'

(13) màm kɔ́-ɔ̀   pág-à̃   yɛ́ndà.
      1SG.I give-AFF woman-CL1.IDN 3SG.II

      'I gave it to the woman.'

Other Gur languages such as Dagaare (Ghana) display this type of object alignment system (Bodomo 1997), as do a number of non-Gur African languages (cf. Ackerman, Malouf & Moore to appear, on the Kordofanian language Moro).[13]

4.2 Word order of objects in relation to animacy

In Mooré the two lexical objects of a ditransitive construction are unmarked: they bear no case marking or adpositional marking indicating their syntactic role within the clause. Humanness and animacy are, however, determinant factors affecting constituent order in ditransitive constructions.[14] If one of the two objects is [+human] or [+animate] and the other is [-animate], the order is [+human/animate] followed by [-animate] (see also Canu 1974: 394). This pattern can be illustrated with verbs such as 'lend' (14) and 'feed' (15).

(14) màm pɛ́ŋ-à   pág-à̃   wéé-f-à.
      1SG.I lend-AFF woman-CL1.IDN bike-CL19-IDN

      'I lent the bike to the woman.'

(15) màm rílg-à   bɔ̌ɔ̌-s-á   nàngù-rí.
      1SG.I feed-AFF goat-CL13-IDN peanut-CL5

      'I fed peanuts to the goats.'

However, the hierarchy-based ordering rule can be violated by pragmatic factors such as the placement of emphasis on one of the two objects (Kabore 1985: 375).

When O_R and O_T are both [+human], both [+animate], or one [+human] and the other [+animate], the order is variable and the construction may display ambiguity as to which of the two objects is semantically the RECIPIENT and which is semantically the THEME if no further specification is added.[15] Thus, in (16)-(18) the semantic role associated to each object could be either R or T, independently of order. Therefore, in this language a [+human] object does not seem to outrank a [+animate] object with respect to word order.

---

[13] I am thankful to an anonymous reviewer for suggesting this reference to me.

[14] In the following discussion I use [+animate] to refer to a non-human animate participant, and [+human] to refer to a human animate participant.

[15] An anonymous reviewer asked: (i) whether an inanimate O_T can precede the O_R when O_T receives emphasis/focus; (ii) whether O_T can precede O_R when O_T has contrastive focus. The answer to the first question is yes (see Kabore 1985: 375 for examples). At the present time I do not have enough data to answer the second question.

(16) màm wíníg-à   bíí-gɛ̀   pág-ɛ̀.
     1SG.I show-AFF child-CL12.IDN woman-CL1.IDN

     'I showed the boy to the woman.' or 'I showed the woman to the boy.'

(17) màm wíníg-à   báá-gɛ̀   bʊ́ʊ́-s-á.
     1SG.I show-AFF dog-CL12.IDN goat-CL13-IDN

     'I showed the goats to the dog.' or 'I showed the dog to the goats.'

(18) màm wíníg-à   wób-g-à   bii-gɛ̀.
     1SG.I show-AFF elephant-CL15-IDN child-CL12.IDN

     'I showed the elephant to the child.' or 'I showed the child to the elephant.'

In sum, the two objects are not morphosyntactically differentiated based on semantic role, though linear order of O_R and O_T is constrained by animacy.

4.3 Split in the coding of O_R and O_T as independent pronouns

Either O_R or O_T, or both simultaneously, can be expressed as independent pronouns. Objects expressed as independent pronouns show the same behavior as full-fledged NPs with respect to animacy-governed word order (see §4.2).

There appears to be a split in Mooré between the set of independent pronouns used for 1SG, 2SG, 3SG and the set used for 1PL, 2PL and 3PL. In the singular persons, the independent pronoun for the first linear object always comes from set II (19).

(19) Set II singular forms coding the first linear object when followed by another NP
     A   V O_R [SET II] O_T
     bámb kɔ́-ɔ̀   máám/fóó/yɛ́ndà péén-d   myʉ́-g-à.
     3PL.I give-AFF 1SG/2SG/3SG   scarf-CL5 red-CL12-IDN

     'They gave {me, you, him/her} a red scarf.´

For 1PL, 2PL and 3PL, a pronoun form from set I is used *if* another NP, including one functioning as O_T, an oblique, or an adverb follows (20). Using a set II pronoun form for the first linear object in (20) would result in ungrammaticality.

(20) Set I plural forms coding the first linear object when followed by another NP
     A   V O_R [SET I] O_T
     à   kɔ́-ɔ̀   tɔ́nd/yɛ́mb/bámb péén-d   myʉ́-g-à.
     3SG.I give-AFF us, you-all, them scarf-CL5 red-CL12-IDN

     'He gave {us, you, all, them} a red scarf.´

If there is no NP following the first linear object after the verb, the plural persons are expressed by set II pronouns. This split in pronoun-form selection appears to be phonological rather than syntactic in nature (see Peterson 1971 for a phonological account). This casts doubts on the nature of set II plural forms as a set of specifically case-marked

'accusative' pronouns: they do not appear strictly in 'object' position, but rather the set II plural forms only occur in final position within the clause. Further, while the singular persons of set I and set II differ in vowel length and tone, the plural persons of the two sets differ only in the absence vs. presence of a final vowel, respectively.

4.4 Optional indexation of O_R and O_T in relation to animacy

In Mooré, only core syntactic object arguments can be optionally indexed[16] in the verb by the set III clitic pronouns. Obliques (21), NPs followed by relator nouns like *zúgù* (which in (23) indicates something like 'off' or 'from' the surface of), and locatives (25) cannot be indexed in the verb, as shown by the ungrammatical examples in (22), (24) and (26).

(21) à Músá wɛ́-ɛ̀ à Ríhnátà nɛ́ kúg-rì.
3SG.I M. hit-AFF 3SG.I R. with stone-CL5

'Musa hit Rihnata with a stone.'[17]

(22) * à Músá wɛ́-ɛ̀-là à Ríhnátà.
3SG.I M. hit-3SG.III-AFF 3SG.I R.

('Musa hit Rihnata with it'.)[18]

(23) à yɛ́ɛ́s-à zíí-m-à fɯ́-gà̀ zú-gù.
3SG.I wipe-AFF blood-CL22-IDN cloth-CL12.IDN head-CL15

'She swiped the blood off the cloth.'

(24) * à yɛ́ɛ́s-à-là zíí-m-à.
3SG.I wipe-3SG.III-AFF blood-CL22-IDN

('She wiped the blood off it.')

(25) màm tɔ́m-à bíí-gà̀ zá-k-ɛ̀.
1SG.I work-AFF child-CL12.IDN house-CL12-LOC

'I sent the child home.'[19]

(26) * màm tɔ́m-à-là bíí-gà̀.
1SG.I work-3SG.III-AFF child-CL12.IDN

('I sent the child to it.')

---

[16] Throughout the paper, *optional indexation* should be understood as a synonym of *optional bound pronominal agreement*.

[17] Proper names are usually preceded by the 3SG pronoun *à* as shown in (21). This *à* is not a case marker; it occurs before any proper name in all syntactic roles (S, A, O).

[18] The set III 3SG object pronoun *-à* indexed in the verb in §4.1 undergoes the same phonological process observed for the affirmative marker *-à* in §4.1. When the shape of an immediately-preceding verb root is CV, the set III 3SG pronoun *-à* takes the form of a copy of the last vowel of the verb root, as in (22) where *-à* > *-ɛ̀* because the verb root is CV and the last vowel is *ɛ̀*. If the contiguous verb root is not CV, then the 3SG bound pronoun surfaces as *-à(-là)* as in (24) and (26). (Manuel Otero, p.c.).

[19] In Mooré, the verb *tɔ́mè* can mean 'work' but also 'send someone'.

It was stated in §4.1 that the set III bound pronouns optionally index in the verb the O of a monotransitive clause, and O_R and O_T of a ditransitive clause. Thus in terms of pronoun forms, Mooré shows a neutral object alignment system in optional bound pronominal object indexation.

In ditransitive constructions, either O_R or O_T can be *optionally* indexed in the verb, but never both at the same time. By *optionally* I mean that if either O_R or O_T are indexed in the verb, a lexical NP co-referential to the object that is indexed cannot occur in the clause. If one of the two objects is indexed, the other needs to be expressed either as a NP or as an independent pronoun.

There seem to be in Mooré no restrictions on optional bound pronominal indexation depending on a hierarchy of animacy or person (such as 1SG>2SG>3SG), plurality or definiteness. The combinations that have been tested are listed in Table 3.

Table 3: Optional indexation of O_R and O_T depending on animacy

	O_R	O_T	Tested verbs	Indexation of O_R and O_T
I	+human	+human	'show', 'give'	YES
II	+human	+animate	'show', 'send', 'lend'	YES
III	+human	-animate	'send', 'give', 'tell', 'teach', 'ask for', 'show', 'leave', etc.	YES
IV	+animate	+animate	'show', 'feed'	YES (but see below)
V	+animate	+human	'show'	YES
VI	+animate	-animate	'feed', 'show', 'give'	YES

For reasons of space, only combination IV is illustrated with the verb *wíníg*ì 'show' in (27)–(29). Whenever both objects refer to animate or human entities, the interpretation of their semantic role is inherently ambiguous. This is true also in the case of optional indexation.

(27)   à   Músá wíníg-à   báá-gã̀   bʊ̌ʊ̌-sé.
      3SG.I M.   show-AFF dog-CL12.IDN goat-CL13

      'Musa showed goats to the dog.'

(28)   à   Músá wíníg-b-là   báá-gã̀.
      3SG.I M.   show-3PL.III-AFF dog-CL12.IDN

      'Musa showed them to the dog.'

(29)   à   Músá wíníg-à-là   bʊ̌ʊ̌-sé.
      3SG.I M.   show-3PL.III-AFF goat-CL13

      'Musa showed goats to it (i.e. to the dog).'

Some verbs seem to require the indexed object to be semantically a Recipient, as is the case with *rílgè* 'feed'.[20] In (31) the indexed object corresponds to the Recipient of (30), 'lion'. The ungrammaticality of (32) shows that with the verb 'feed' the indexed object cannot refer to the Theme (i.e. 'sheep' in (30)); rather, a Recipient interpretation of the indexed object is forced, even if it is semantically awkward.

(30) à   Músá rílg-à   bɔ̀yí-gɑ̃̀   píí-s-á.
      3sg.i M.   feed-aff lion-cl12.idn sheep-cl13-idn

      'Musa fed the sheep (plural) to the lion.'

(31) à   Músá rílg-à-là   píí-s-á.
      3sg.i M.   feed-3sg.iii-aff sheep-cl13-idn

      'Musa fed the sheep to it (to the lion).'

(32) à   Músá rílg-b-là   bɔ̀yí-gɑ̃̀.
      3sg.i M.   feed-3pl.iii-aff lion-cl12.idn

      ? 'Musa fed the lion to them (to the sheep)' not *'Musa fed them (the sheep) to the lion.'

Examples (30)-(32) are the only instance I have found in which O_R and O_T appear to be treated differently in the language. Further research is needed to fully understand the dynamics of these 'exceptions'.

4.5 Constituency

Adverbs of time such as 'yesterday' cannot go between V and O_R or between O_R and O_T. Thus, acceptable versions of (33) and (34) feature the temporal Adverb at the beginning or at the end of the clause, but never in a position which disrupts the sequence [V O_R O_T].

(33) * màm tóól-à   záàmɛ̀   pág-ɑ̃̀   rú-kɑ̃̀.
        1sg.i send-aff yesterday woman-cl1.idn pot-cl12.idn

      ('I gave yesterday the pot to the woman.')

(34) * màm tóól-à   pág-ɑ̃̀   záàmɛ̀   rú-kɑ̃̀.
        1sg.i send-aff woman-cl1.idn yesterday pot-cl12.idn

      ('I gave to the woman yesterday the pot.')

Identical behavior is observed with other adverbs, such as *bàlá* 'only' and *yàsá* 'again'. This indicates that the verb and its two objects form one indivisible constituent.

---

[20] The verb *rílgé* 'feed' is clearly a derived causative from the verb *rí* 'eat', by means of the causative suffix *-g*. Conceivably 'show' (see (30)) is more lexicalized than 'feed' as a ditransitive verb; perhaps there might be some restrictions on indexation of productively-derived Causees as with 'feed'.

5 Covert properties of ditransitive constructions

This section deals with covert properties of ditransitive constructions, such as controller or target of co-reference, relativization, reflexivization, etc. Passivization is excluded because in Mooré, a functional equivalent to a passive is at most an impersonal construction in which whoever realizes an action is introduced as generic 'people' and the clause construction remains transitive. Quantifier floating is irrelevant because quantifiers in Mooré can appear only immediately after the noun they modify.

5.1 Relativization

In Mooré, both O_R (35) and O_T (36) can be relativized by means of the same strategy:

(35) pág   lánníngà á   Músá hɔ́   tóól rú-kằ.
     woman REL   3SG.I M.   SUBRD send pot-CL12.IDN
     'the woman to whom Musa sent the pot'

(36) rú-k   lánníngà á   Músá hɔ́   tóól pág-ằ.
     pot-CL12 REL   3SG.I M.   SUBRD send woman-CL1.IDN
     'the pot which Musa sent to the woman'

In both cases, the head noun is followed by the invariable relativizer *lánníngà*. The relativizer is followed by the subject of the relative clause, a 'subordinate' clause marker, a verb root without aspectual markers and the other object (i.e. the one that is not being relativized on). Essentially, relativization of objects occurs by means of a gap strategy.

5.2 Control of co-reference of possessive NPs

O_R has the ability to control co-reference of a following possessor in the same way as the subject argument does.

(37) à   Músá kɔ́-ɔ̃   à   Ouérmí lígí-d-à   *à*   zá-k-ɔ́wằ.
     3SG.I M.   give-AFF 3SG.I O.   money-CL21-IDN 3SG.I house-CL12-LOC
     'Musa$_i$ gave the money to Ouermi$_j$ in his$_{ij}$ house.'

When both objects are [+human], both can control co-reference of a possessor. This means that O_T shows the same property as O_R for the control of possessors. In (38), 'his house' can be the house of the subject 'Musa', O_R 'Ouermi' or O_T 'the child'.

(38) à   Músá wíníg-à   à   Ouérmí bíí-gằ   à   zá-k-ɔ́wằ.
     3SG.I M.   show-AFF 3SG.I O.   child-CL12.IDN 3SG.I house-CL12-LOC
     'Musa$_i$ showed the child$_j$ to Ouermi$_y$ in his$_{ijy}$ house.'

5.3 Control of co-reference under coordination and subordination

In Mooré, when a main clause has a 3SG subject and a 3SG object and a following clause has a 3SG subject, different linking elements disambiguate the possible co-reference of arguments between the two clauses. If the subject of the linearly first clause is co-referent with the subject of the second clause, the coordinator *là* is used.

(39) à   Músá mók-à   à   Ríhnátà lá=à   lóóg-è.
      3SG.I M.   kiss-AFF 3SG.I R.   and=3SG.I leave-CFV

'Musa$_i$ kissed Rihnata$_j$ and he$_i$/*she$_j$ left.' or 'Musa$_i$ kissed Rihnata$_j$ and he$_k$/she$_k$ left.'

If the object of the linearly first clause is co-referential with the subject of the following clause, the complementizer *tí* is used:

(40) à   Músá mók-à   à   Ríhnátà t=à   lóóg-è.
      3SG.I M.   kiss-AFF 3SG.I R.   COMPL=3SG.I leave-CFV

'Musa$_i$ kissed Rihnata$_j$ and *he$_i$/she$_j$ left' or 'Musa$_i$ kissed Rihnata$_j$ and he$_k$/she$_k$ left.'

When the linearly first clause is ditransitive, the same pattern described above is observed. Importantly, when *tí* is used, either O$_R$ or O$_T$ can be co-referential with the subject of the following clause:

(41) à   Músá wíníg-à   à   Ouérmí bíí-gằ   t=à   lóóg-è.
      3SG.I M.   show-AFF   3SG.I O.   child-CL12.IDN COMPL=3SG.I leave-CFV

'Musa$_i$ showed the child$_j$ to Ouermi$_y$ and *he$_i$/he$_{jk}$ left.'

Co-reference tests were also explored under the condition of subordination (42). When the subordinate clause is ditransitive and both objects are [+human], *either* object can control co-reference of the subject in the main clause (along with the subject of the subordinate clause).

(42) à   Músá hằ   wíníg à   Ouérmí bíí-gằ   à   lóóg-à-mè.
      3SG.I M.   SUBRD show   3SG.I O.   child-CL12.IDN 3SG.I leave-AFF-PHB

'When/since Musa$_i$ showed the child$_j$ to Ouermi$_y$ he$_{ijk}$ left.'

Thus, in control of co-reference under coordination and subordination, as happens with control of co-reference of a possessor, O$_R$ and O$_T$ display the same behavioral properties.

5.4 Reflexivization and reciprocalization

In Mooré both O$_R$ (43) and O$_T$ (44) can be the target of reciprocalization:

(43) à   Ouérmí nɛ́ Músá wíníg-à   tááb   báá-gã̀.
      3SG.I O.   with M.   show-AFF each.other dog-CL12.IDN
      '[Musa and Ouermi]ᵢ showed the dog to each otherᵢ.'

(44) à   Ouérmí nɛ́ Músá wíníg-à   tááb   à   Alí.
      3SG.I O.   with M.   show-AFF each.other 3SG.I A.
      '[Musa and Ouermi]ᵢ showed each otherᵢ to Ali.'

In both cases, the reciprocal NP must immediately follow the verb, regardless of its semantic role.

Both O_R and O_T can also be targets of reflexivization. In this case too, the reflexive NP must immediately follow the verb.

(45) pág-ã̀   kɔ́-ɔ̀   à   mɛ́ng máng-rè.
      woman-CL1.IDN give-AFF 3SG.I self mango-CL5
      'The woman gave a mango to herself.'

(46) pág-ã̀   kɔ́-ɔ̀   à   mɛ́ng rà̀ò-á.
      woman-CL1.IDN give-AFF 3SG.I self man-CL1
      'The woman gave herself to the man.'

5.5 Information questions

In Mooré, both O_R and O_T are fronted to the beginning of the clause when an information question is built upon them.

(47) bwɛ̌ lá   Músá kɔ́ kám-bà?
      thing which M.   give children-CL1
      'What did Musa give to the children?'

(48) án   lá   Músá kɔ́ péén-dã̀?
      person which M.   give scarf-CL5.IDN
      'To whom did Musa give the scarf?'

6 O_R and O_T in Mooré ditransitive constructions within a typological perspective

This section offers a summary of the overt and covert properties of O_R and O_T in Mooré ditransitive constructions as discussed in §4 and §5, and determines whether double-object constructions in this language can be classified in terms of a PRIMARY versus SECONDARY OBJECT (Dryer 1986) type system, or SYMMETRICAL versus ASYMMETRICAL object type system (Bresnan & Moshi 1990).

Mooré does not appear to be a PRIMARY/SECONDARY OBJECT language in the sense of Dryer (1986) because both objects of a ditransitive construction display the same properties as the object of a monotransitive construction. According to Bresnan & Moshi (1990: 147), in symmetrical object languages, both objects display PRIMARY OBJECT properties (in the sense of Dryer 1986); the properties they test for selected Bantu languages include passivizability, object agreement, adjacency to the verb, unspecified object deletion, retention of object marking on verbs in presence of applicatives, reflexivization and reciprocalization. The authors state that in a true symmetrical object language, "different arguments can *simultaneously* have primary object properties" (Bresnan & Moshi 1990: 153, emphasis in the original).

Many of the object properties of Bantu languages, such as passivization and object case marking, are absent in Mooré, and the concept of 'simultaneity' could pose some difficulties in the case of Mooré just because of the morphosyntactic nature of the language itself. However, by and large this paper has shown that the two objects of a ditransitive construction in Mooré are indistinguishable and share the same value for all the object properties that have been investigated (Table 4).

7 Conclusions

Mooré shows a neutral object alignment system in independent pronouns and in optional bound clitic pronouns in the verb. There appears to be a split in the selection of independent pronouns for plural persons if another NP or adverb follows. This, however, should probably be considered a phonological rather than a syntactic split. Word order of objects in ditransitive clauses is strictly [+human/+animate] followed by [-animate]. Semantic role assignment is variable when both objects are [+human/animate]. In bound pronominal agreement, both O_R and O_T can optionally be expressed by means of bound pronouns in the verb, but never at the same time. As for constituency, both O_R and O_T form a constituent with the preceding verb.

As for covert properties, O_R and O_T show the same relativization strategy, both can be the target of reflexivization and reciprocalization, both can control co-reference of a following possessor and co-reference under coordination/subordination. In information questions, the question word referring to both objects is fronted at the beginning of the clause.

Based on this evidence, Mooré appears to be a symmetrical object language type. So far, symmetrical object systems have been attested, among others, in Bantu (Bresnan & Moshi 1990), Austronesian (Donohue 1996), Yagua, an isolate from Peru (Payne & Payne 1989), Western Australian (Dench 1995) and Totonac-Tepehua (McKay & Trechsel 2008) languages. However, Bresnan & Moshi's (1990) condition of simultaneity as pivotal to defining the symmetrical object language type is not always applicable in Mooré simply due to morphosyntactic restrictions in its grammar (e.g. only one object can be marked on the verb at a time). This suggests room for further typological research on subtypes of symmetrical object systems. It should also be noted in this respect that some of the covert properties discussed here for Mooré are not listed among the common tests for

Table 4: Object properties in Mooré ditransitive constructions

OBJECT PROPERTY	O_R	O_T	Notes
	OVERT		
CASE MARKING	–	–	Neither object displays case or adpositional marking.
WORD ORDER	✓	✓	Either O_R and O_T can appear immediately after the verb when both are [+human/animate].
OPTIONAL BOUND PRONOMINAL AGREEMENT	✓	✓	No significant restrictions found based on animacy, number, definiteness or hierarchy of person. O_R and O_T cannot both be indexed at the same time.
CONSTITUENCY	✓	✓	Both appear to be a single constituent with the verb.
	COVERT		
CONTROL OF CO-REFERENCE OF A POSSESSOR	✓	✓	Both O_R and O_T can control co-reference.
CONTROL OF CO-REFERENCE UNDER SUBORDINATION	✓	✓	Both O_R and O_T can control co-reference.
RELATIVIZATION	✓	✓	The same strategy is used for O_R and O_T.
REFLEXIVIZATION	✓	✓	Both O_R and O_T can be the target.
RECIPROCALIZATION	✓	✓	Both O_R and O_T can be the target.

investigating types of object systems in the literature on other languages. This raises the theoretical question of whether overt and covert properties should be assigned the same amount of 'weight' or 'relevance' in establishing subtypes of systems. The analysis advanced here for Mooré also invites future comparative work on object properties in other Gur languages. This would improve the understanding of the grammatical relations in this language family and add to the typological understanding of symmetrical vs asymmetrical object systems.

Acknowledgments

Abbreviations

1,2,3	1st, 2nd, 3rd person	CL	noun class marker
I	set I pronoun	LOC	locative case marker
II	set II pronoun	IDN	identifiable
III	set III pronoun	PHB	(verb) phrase boundary
AFF	affirmative marker	PL	plural
CFV	citation form vowel	PVC	pre-verbal conjunction
		SG	singular

References

Ackerman, Farrell, Robert Malouf & John Moore. to appear. Symmetrical objects in Moro: An HPSG account. *Journal of Linguistics*.

Alexandre, Gustave. 1953. *La langue mōré. (Mémoires de l'Institut Français d'Afrique Noire)*. Dakar: IFAN.

Bodomo, Adams. 1997. *The structure of Dagaare*. Stanford: CSLI Publications.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21. 147–185.

Calamai, Silvia & Pier Marco Bertinetto. 2005. Indagine preliminare sul vocalismo orale della lingua mòoreé (Gur, Niger Congo). *Quaderni del Laboratorio di Linguistica* 5. http://linguistica.sns.it/QLL/QLL04_05/Calamai_Bertinetto.pdf, accessed 2015-04-15.

Canu, Gaston. 1974. *Description synchronique de la langue mòòré*. Lille: Université de Lille III. (Doctoral dissertation.)

Dench, Alan. 1995. *Martuthunira: A language of the Pilbara region of Western Australia* (Pacific Linguistics, Series C 125). Canberra: Australian National University.

Donohue, Mark. 1996. Bajau: A symmetrical Austronesian language. *Language* 72. 782–793.

Dryer, Matthew. 1986. Primary objects, secondary objects, and antidative. *Language* 62. 808–845.

Dryer, Matthew. 2007. Clause types. In Timothy Shopen (ed.), *Language typology and syntactic description*, 224–275. Cambridge: Cambridge University Press.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Hyman, Larry M. & Alessandro Duranti. 1982. On the object relation in Bantu. *Syntax and Semantics* 15. 217–239.

Kabore, Raphael. 1985. *Essay d'analyse de la langue m ʋʋre (parler de Wàogdgò)*. Paris: Université de Paris 7, Département de Recherches Linguistiques, Laboratoire de Linguistique Formelle.

Kittilä, Seppo. 2005. Recipient-prominence vs. Beneficiary prominence. *Linguistic Typology* 9(2). 269–297.

Kouraogo, Pierre. 1976. *Serial verb constructions and some aspects of Moore grammar: Towards a case grammar solution for serialization*. Doctoral dissertation. Tampa: University of South Florida.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the world*. 19th edn. Dallas, TX: SIL International.

Malchukov, Andrej, Martin Haspelmath & Bernard Comrie (eds.). 2010. *Studies in ditransitive constructions: A comparative handbook*. Berlin: Mouton de Gruyter.

Manessy, Gabriel. 1963. Les particules affirmatives post-verbales dans le groupe voltaïque. *Bulletin de l'Institute Fundamental d'Afrique Noire* 25(1). 108–124.

McKay, Carolin J. & Frank R. Trechsel. 2008. Symmetrical objects in Misantla Totonac. *International Journal of American Linguistics* 74(2). 227–255.

Nikièma, Norbert. 2000. Gradin et harmonies vocaliques en mooré. *Afrika und Übersee* 83. 209–251.

Nikièma, Norbert. 2003. Les constructions à verbes seriels en moore. In Mary Esther Kropp Dakubu & E. Kweku Osam (eds.), *Studies in the languages of the volta basin*, vol. 1, 111–132. Proceedings of the Annual Colloquium of the Legon-Trondheim Linguistics Project. Legon: University of Ghana, Department of Linguistics.

Olawsky, Knut J. 1999. *Aspects of Dagbani grammar: With special emphasis on phonology and morphology*. Munich: LINCOM Europa.

Payne, Doris L. & Thomas E. Payne. 1989. Yagua. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, 249–474. Berlin: Mouton de Gruyter.

Peterson, Thomas H. 1971. *Mooré structure: A generative analysis of the tonal system and aspects of the syntax*. Los Angeles: UCLA. (Doctoral dissertation.)

Rennison, John. 1990. On the elements of phonological representation. *Folia Linguistica* 24(3-4). 175–244.

Chapter 5

All-in-one and one-for-all: Thetic structures in Buli grammar and discourse

Anne Schwarz

Universidad Regional Amazónica IKIAM

This paper describes the form and function of the thetic statement and its categorical counterpart in the north Ghanaian Gur language Buli. Instead of applying cross-linguistically well-known means to encode thetic statements, such as intonation and word order, a connective particle is positioned at the beginning of the verb phrase and followed by a dependent verb form or a noun in predicative function. This thetic construction may contain a prosodic break between the strictly preverbal subject constituent and the predicate. Interestingly, the structural features of thetic statements are also encountered in various embedded clause types and give way to different stacking configurations. In terms of function, it is argued that the internal lack of information structure renders the thetic statement compatible with a number of different pragmatic situations and thus contributed to its wide distribution in Buli grammar.

1 Introduction

The thetic/categorical opposition is a useful information-structural distinction in some languages, in particular in those that have developed special grammatical features for pragmatically marked and unmarked topics. However, despite its information-structural potential, the thetic/categorical distinction is only rarely considered in the functional-typological literature concerning African languages (but cf. Güldemann 1996; 2010; 2012; Fiedler et al. 2010: 250-253; Schwarz 2010a; beyond Africa, see in particular Kuroda 1972; Sasse 1987; 1995; Lambrecht 1987; 2000; Ulrich 1988).

The thetic/categorical distinction in linguistics was borrowed from philosophy in the 19th century when Brentano and his student Marty proposed a fundamental dichotomy between two different basic types of logical statements (Sasse 1987: 535). Categorical statements are those with a classical bipartite subject-predicate structure in the Aristotelian sense. These statements involve two successive acts: naming an entity and then

making a statement about it. The thetic statement, on the other hand, is internally unstructured and expresses an event, a state or a situation without the concept of a topic that is commented about. Kuroda (1972) was one of the scholars who applied this statement classification in linguistics in order to describe the distribution of certain Japanese morphemes. Ulrich (1988: 388) states that a distinction between categorical and thetic constructions is universal, but that these constructions are expressed in different ways by different languages. Since the recognition of the thetic/categorical opposition provides new insights into the intricate distribution of certain sentence construction types in Buli, the distinction between statements that are mentally construed as unstructured (thetic), versus structured into topic and comment (categorical), has also been adopted here (Table 1, cf. also Ulrich 1988).

Table 1: The thetic/categorical distinction

	Categorical	Thetic
Philosophical level	gives a statement about an argument, predicative linkage	presents a fact as an inseparable unit
Pragmatic level	topic-comment structure	unstructured (without the notion of a topic to be commented on)

Transferred to the pragmatic level, the bipartite structure of categorical utterances consists of a topic and of a comment that predicates something about it. In the thetic statement, on the other hand, the propositional information is conceptualized as a single unstructured package. Unlike categorical statements, thetic statements have no internal information structure and have been described as monomial predications where "the entire situation, including all of its participants, is asserted as a unitary whole" (Sasse 1995: 4f).

The term TOPIC is used here to refer to "what the utterance is about," thinking of the topic as a textual location or storage address for subsequent information. Reinhart (1982) uses a file card metaphor according to which the topic corresponds to the title of a file card to be followed by the comment. Functioning as a storage address, it is a characteristic feature of topics to be established early in the discourse to be then commented on. Accordingly, the subject often has the (unmarked) topic role. In coherent discourse, topics tend to display certain continuity and thus often represent information already shared between the interlocutors. Languages may therefore provide particular (pronominal) morphology or allow even zero forms for encoding such topic continuity. Accordingly, it is possible to encounter categorical statements that have a conceptual topic about which the speaker comments, even though the topic itself is not mentioned in the sentence.

In intonation languages, such as English, the formal distinction between thetic and categorical statements often relies on accentuation. This is exemplified in (1)-(2). In the unstructured thetic utterance there is only one accent (marked in small capital letters) which is placed on the subject. In the pragmatically bipartite categorical utterance, on the contrary, both subject and predicate are accented.

(1) English thetic utterance (unstructured):
The CHICKEN's burnt.
As a reply to 'What the hell is this stench?' (Sasse 1987: 529)

(2) English categorical utterance (bipartite topic-comment structure):
The CHICKEN's BURNT.
In the context of 'Where is the roast chicken I expected, why do I have a hamburger on my plate?' (Sasse 1987: 529)

Another well-known linguistic means for the expression of the thetic/categorical distinction is word order change. In Modern Greek, for instance, thetic statements display subject inversion. The non-topical subject in (3) follows the verb rather than preceding it; the latter is, however, the case in the categorical statement in (4).

(3) Modern Greek thetic utterance:
Xtipise to   tilefono.
rang   ART phone
'The PHONE rang.' (thetic) (Sasse 1987: 536)

(4) Modern Greek categorical utterance:
To   tilefono xtipise.
ART phone   rang
'The PHONE RANG.' (categorical) (Sasse 1987: 536)

In terms of discourse functions, it is obvious that the categorical organization of the utterance particularly serves incremental information transfer, as the conceptual presence of a topic provides a cognitively useful dedicated storage address for any new information. Categorical statements seem to represent the dominant pattern in several communicative situations in coherent discourse, and are therefore overall pragmatically less marked than the unstructured thetic organization. Thetic statements, on the contrary, can be used to deliberately deviate from discourse coherence, for instance, by presenting an unexpected event in episode-initial position. Cross-linguistic studies (Sasse 1995) have also confirmed that certain lexical classes of predicates are typically not used to comment on topics across languages. Weather verbs are particularly well-known in this respect. Cross-linguistically we can therefore expect categorical statements to represent the basic communicative model in most coherent discourse sections, while thetic statements might be more appropriate at scene-initial points and with particular predicate types. Following Güldemann (2010: 86), we will suggest that when the speaker has the choice and applies the thetic rather than the categorical configuration, she deliberately

uses this "grammatical device to CANCEL the sentence-internal information structure that is induced by the morphosyntax of an unmarked sentence."

The remainder of the paper describes the basics of the thetic/categorical distinction in Buli. §2 provides a brief introduction to the language. §3 deals with information structure in Buli and first outlines the language-specific distinction between the two configurations (§3.1), followed by more detailed structural accounts of categorical (§3.2) and thetic statements (§3.3). §3.4 deals with the pragmatic range of situations for the application of thetic statements and §3.5 discusses the embedding of corresponding structures. A short summary is provided in §4.

2 The language

Buli belongs to the Gur language family and is spoken by approximately 150,000 people (Lewis, Simons & Fennig 2015) in northern Ghana. There are three dialectal zones, two of them with further minor internal variations. The data used for illustration here all stem from the central zone and were recorded with speakers from Sandema and Wiaga between 1996-2009. Within the Gur family, Buli is classified as a member of the Oti-Volta branch, in which it forms, together with its closest relative and south-western neighbour Kɔnni, the Buli/Kɔnni subgroup (Manessy 1979; Naden 1989). Buli shares several linguistic features with other Gur languages, but also reveals various peculiarities. Like most Gur languages, it is a tone language and has a three-tone contrast (High, Mid, Low). Deviating from several relatives, though not from its sister language Kɔnni (Cahill 2007), Buli verbs have no lexical tone and the grammatical relevance of tone is fairly high. Both lexical and grammatical tones are active beyond their primary tone bearing unit (TBU): High tones extend their domain in certain environments, and in most dialects, Low tones spread rightward into High TBU's, but spare Mid ones (Schwarz 2003; 2007). Grammatical affixes and clitics are attested both with and without inherent tone. Tone alignment patterns are complex and depend on lexical, grammatical, and sometimes syntactic variables. In addition to transparent tone patterns which only change in their surface realization, there are also some nouns and verb forms which display an "unstable" rising tone pattern, realized as LM on bimoraic segments in phrase-final position. However, the final rise often gets completely lost, unlike the High component in other tone patterns which remains traceable even under tone spreading conditions. This and other phenomena point to intonational influences in grammatical and lexical tone.

Buli is a head-marking language with a morphological type that is mildly agglutinating, predominantly with suffixes in the nominal domain and with mostly proclitics in agreement. The basic word order is SV(O)(OTHER). The verb system is aspectually organized, with the perfective representing the unmarked aspect for dynamic verbs. Serial verb constructions are common and fulfil voice as well as particular tense-aspect-mood-specifications. The nominal domain is characterized by a fully operational and mostly overt noun class system. Nouns display suffixes and agreement operates primarily phrase-externally employing proclitics or prefixes. Two nominal forms with different suffix sets exist. In singular, only the definite forms have fully overt noun class suffixes,

while the indefinite forms often display just suffix traces (see Schwarz 2007 for more details on the noun class morphology).

3 Information structure in Buli

Buli grammar closely interacts with information-structural categories. Under certain conditions, for instance, it seems that the morphological marking of a focal constituent within the sentence is grammatically required, i.e. its lack would yield an ungrammatical sentence. This situation and the dominance of morphosyntactic means for information-structural encoding is reminiscent of other African languages for which grammaticalized focus systems are reported (Hyman & Watters 1984; Robert 2000; Wolff 2005).

With respect to the information-structurally controlled morphosyntactic variation in Buli, we observe major encoding differences in terms of different sentence constructions. Long-lasting studies concerning this variation, taking data into account from both natural discourse and more controlled elicitation tasks (as those by Skopeteas et al. 2006 and Schwarz 2007), suggest that it is not just sentence-internal information-structural categories such as topic and focus that are systematically morphosyntactically encoded, but rather the distinction between thetic and categorical statements in order to structure and organize discourse above the simple clause-level.

3.1 The thetic/categorical distinction in Buli

In Buli, the propositional content of an utterance can be expressed in at least two formal ways: either in a simple categorical form as in (5), or in a thetic statement as in (6). Note that the different subject forms in the sentences used here for illustration – a definite subject noun in the categorical statement and an indefinite subject noun in the thetic one – are not mandatory but reflect some typical features to be discussed further below.[1]

(5) Categorical (simple):
    Nípōōwá   ŋɔ̀b kà túé.
    nípōk=wá   ŋɔ̀b ká túe
    woman=DEF1 eat KA bean.NC6
    'The woman ate beans.' [BL (1a-trans) 2004: 28.01]

---

[1] The representation of examples contain a first data line that follows common orthographic practice to some degree, but also marks surface tone. Most examples contain a second data line that provides morpheme segmentation and underlying tones (wherever these differ from orthographic practice) and uses IPA symbols with those consonants that could otherwise be misinterpreted. Unpublished data is from my fieldwork. The examples stem from audio-recorded narratives, conversations and contextualized elicitations and translations. The source provided in square brackets at the end of an example free translation represents the name and number of my Toolbox record under which they are stored.

(6) Thetic:

Nípōk   lē   ŋɔ̀b tùē.
nípōk   lē   ŋɔ̀b túe
woman.NC1 CON eat bean.NC6

'A woman ate beans.' [BL (2) 2005: 454-2]

Both sentences display SVO order, but they differ in two principled ways: In simple (e.g., monoclausal) categorical statements, the verb is subject to regular inflection including tonal subject agreement where applicable. In the affirmative perfective, the inflection includes tonal differences on the verb depending on whether the subject is a speech act participant, or not (Schwarz 2007). Furthermore, in simple categorical statements in the affirmative, there is usually a postverbal particle *ká*. This particle marks the left edge of a postverbal constituent, either one immediately after the verb or one more distantly following the verb. In (5), *ká* precedes the object noun and is subject to Low-tone spreading that extends from the grammatical Low tone of the immediately preceding verb.

In the thetic statement on the other hand, the subject is followed by a predicate-initial connective particle *lē* (allomorph *nē*) after which the content verb follows. Perfective content verbs do not display tonal subject agreement but are characterized by the unstable rising tone. Since the verb in (6) is followed by an object noun, it remains without trace of the final rise. The Low tone of the verb form (*ŋɔ̀b*) spreads right onto a following High TBU, just as would happen from any regular Low tone verb. Note also that in the thetic statement, the postverbal particle *ká* cannot occur.

The formal dichotomy between simple categorical and thetic statements is schematically outlined in Table 2. The properties of both configurations are discussed in more detail in §3.2–§3.3.

Table 2: Basic categorical/thetic distinction for a Buli SVO sentence in the affirmative perfective

Categorical (simple):	S (**lē*)	V (subject agreement by tone, if applicable)	*ká*	nominal O
Thetic:	S *lē*	V (no subject agreement; unstable rising tone)	(**ká*)	nominal O

3.2 Simple categorical encoding

The unmarked sentence construction in Buli usually correlates with a categorical reading in which the subject has the unmarked topic role about which the predicate (with its further arguments and adjuncts) comments. In the absence of subject nouns or disjunctive

emphatic pronouns, Buli pronouns bind to the verb and display noun class agreement where applicable. In addition, there is also an invariable vowel *à=* that optionally pro-cliticizes to conjunctions, connective (clausal) morphemes and to verbs. This vowel, a prosodic linker glossed as '&', is neither grammatically nor pragmatically required and occurs only after prosodic breaks. Hence we find this optional element clause-initially (7a) and phrase-initially (7b), with clausal or phrase conjunctions (7a), or directly attaching to the verb (7b). The predicate-initial occurrence is most common when the preceding subject is of particular phonological weight and therefore more likely to occur with its own intonation contour. This results in a prosodic break between the noun and the following verb and the appearance of the prosodic linker. In (7a), the prosodically conditioned vowel initiates a verbless topic-establishing question. The speaker mentions two referents about which she requests further information. In the response (7b), the complex noun phrase with the topical subject is repeated and the following predicate begins with the prosodic linker.

(7)   a.   Question:
          Àgē   nípòkbìlsáŋá   bàyɛwā   mɛɛ?
          à=gē   nípòk-bíl-sa-ŋá   bà=jɛ-wá   mɛ-:
          &=CNJ   woman-small.type-NC13-DEF6   NC2=two-DEF1 also-Q

          'And what about the two girls?' [BL (1a-trans) 2004: 306]

      b.   Answer:
          Nípòkbìlsáŋá   bàyɛwā   à pà   kā flòwèrs.
          nípòk-bíl-sa-ŋá   bà=jɛ-wá   à=pà   ká flowers
          woman-small.type-NC13-DEF6   NC2=two-DEF1 &=take   KA flowers

          'The two girls picked up flowers.' (Schwarz 2009a: 271)

With given and continuous discourse topics (especially in monologues), it is less common to repeat topical subject nouns in every sentence. Topical subjects that are not new, unexpected, contrastive, or otherwise particularly worth mentioning are typically simply encoded by a proclitic subject pronoun,[2] e.g. *ɔ=* in (8) bound to the sentence-initial verb or the first verb in a verb series.

(8)   Question: 'What did the woman eat?'
      Answer:
      Wà ŋɔ̀b kā mùmàŋā.
      ɔ̀=ŋɔ̀b   ká mùm-a-ŋá
      NC1=eat KA rice-NC6-DEF6

      'She ate the rice.' [BL (3) 2005: 561-1]

Particle *ká* typically precedes the immediate postverbal constituent which tends to also be the sentence-final constituent. It targets the object of the verb or a lexeme that

---

[2] This is especially true for third person subjects, even though there are also proclitic forms for speech act participants in singular and plural.

provides circumstantial information in the postverbal domain by appearing at the left edge of this constituent and setting it literally apart from the preceding information. There seems to be a preference for marking the sentence-final constituent, but if there is more than one, the particle remains flexible with respect to which postverbal constituent it selects. In the case of multiple postverbal constituents, the particle is thus reminiscent of an argument focus marker. On the other hand, if the lexical verb or the truth value of the sentence is in focus, the particle *ká* still appears in postverbal position and not before the verb that represents (part of) the focal information. In such predicate-centered focus conditions, the particle combines with an extension (*-mā*) which emphatically confirms the state of affair (Schwarz 2010b). Example (9) illustrates the use of particle *ká* with the extension, here following an intransitive stative verb.

(9) Jāāmú   wɔ̄ŋā kámā   pāā.
    dʒāa-mú   wɔ̄ŋa ká-mā   pāa
    thing-DEF14 be.long KA-CONF much
    'This story is really long!' [Jein 43]

Categorical encoding in Buli also applies to nominal predicates (cf. also Schwarz 2009b). The predicate in (10) consists of the nominal compound *mí-wɔ̀ŋ* 'long rope' which comments on a topical subject noun. The categorical configuration is expressed by the particle *ká* marking the left edge of the constituent in predicative function.

(10) Mííkádɛ́   ká mí-wɔ̀ŋ.
     míi-ká-dɛ́   ká mí-wɔ̀ŋ
     rope-DEF12-PROX KA rope-long.NC12
     'This rope is long.' (Schwarz 2009a: 267)

Some aspects of the categorical encoding with nominal predicates are remarkable. In the absence of a verb, the invariable particle resembles a copula element that signals the predicative function of the nominal it precedes. Unlike verbs, however, the particle *ká* does not occur with a prosodic linker (*à=*). Example (11) illustrates that even after a dedicated pause (indicated by the comma), the linker does not occur.

(11) Nípōōmá   mééná, ká wà lèēbà.
     nípōk-má   méena ká wà=lée-ba
     woman-DEF2 all   KA NC1=daughter-NC2
     'All the women are his daughters.' (Schwarz 2009a: 268)

Unlike categorical statements with verbal predicates, categorical configurations with nominal predicates in the function of the comment do not even require a bound subject pronoun to express the topic within the same clause. The topics of the categorical statements in (12a)-(12b) are only contextually retrievable; within the statements they are not expressed.

(12) a. Ká mí.
          ka 1s.ᴅ
          'It's me.' (e.g., reply to 'Who's that?' after knocking) (Schwarz 2009a: 269)

      b. Ká mí-wɔŋ.
          ᴋᴀ rope-long.ɴᴄ12
          'It [a rope] is long.' [den noff4: 10]

Note that the configuration described here operates within a clause. Buli has another categorical configuration with a comment that is expressed by an entire clause rather than the verb phrase. The clausal comment is particularly suitable in situations in which the sentence-initial topic does not have subject function (13). Such a complex categorical statement contains the clausal conjunction *tè* at the beginning of the comment and prohibits the use of particle *ká* in postverbal position. For reasons of space, the complex categorical construction is not further discussed here.

(13) nídɔāwá tè   nípōōwá   nàk.
      nídɔa-wá tè   nípōk-wá   nàg
      man-ᴅᴇꜰ1 ᴄɴᴊ woman-ᴅᴇꜰ1 hit

      'The man, he was hit by a woman.' (lit. 'The man, the woman hit him.') [BL (1a-trans) 2004: 330]

3.3 Thetic encoding

The thetic configuration is marked by a predicate-initial marker *lē* (allomorph *nē*) and a dependent verb form. In the perfective, the verb (including suffixes) is characterized by the unstable rising tone, rather than the inflectional verb tone pattern with tonal subject agreement as present in categorical statements. Example (14) illustrates this with a perfective predicate to which an object pronoun binds.

(14) Sírí   lē   dòm wā.
      síri   lē   dòm-wā
      bee.ɴᴄ5 ᴄᴏɴ bite-ɴᴄ1

      'A bee stung her.' (Reply to: 'What happened?') [BL (3) 2005: 444-2]

The thetic statement in (15) contains an imperfective predicate (stem *vī* 'follow'). The imperfective verb is encoded like the dependent, non-initial verb in a series, as evidenced by the High-toned imperfective marker *á* which elsewhere precedes dependent verbs. The imperfective marker here binds to the preceding connective particle *lē* and the resulting form is often subject to considerable and irregular surface tone changes (Schwarz 2007).

(15) Bīāk   làà   vī   bíík,   nídɔ̀àbìlī.
     bīak   lē-á   vī   bíik   nídɔ̀a-bíli
     dog.NC12 CON-IPF follow child.NC12 man-small.type.NC5
     'A dog is chasing a child – a boy.' (Reply to: 'What is going on?') [BL (3) 2005: 440-1]

Parallel to verbs and clausal connectives, the predicate-initial marker *lē* can also host the prosodic linking vowel *à=* (16).

(16) Gbáŋ   àlē   dɔ̀à tébùlkù   zúk.
     gbáŋ   à=lē   dɔ̀a tébùl-kú   zúk
     book.NC12 &=CON lie table-DEF15 top.NC15
     'There is a book on the table.' (Schwarz 2009a: 271)

Thetic configurations also extend into the domain of nominal predicates (17). The particle *lē* here directly precedes the predicative nominal, complementing the constructions with particle *ká* encoding categorical nominal predicates. Note that while nominal predicates in categorical statements (where *ká* precedes the nominal predicate) are only occasionally accompanied by a topical subject expression, the non-topical subject of thetic statements is always expressed. In categorical nominal predication, the semantic reading depends on the nature and encoding of the predicative noun. Available readings are PROPER INCLUSION where "a specific entity is asserted to be among the class of items specified in the nominal predicate", or EQUATION where the entity referred to by the subject is "identical to the entity specified in the predicate nominal" (Payne 1997: 114). Thetic nominal predication, on the contrary, always serves identification. Accordingly, the thetic configuration is particularly common when identifying an individual by name, as illustrated in (17).

(17) ǹwàlá   yúé   lē   Àkànchīēm-áɲāmí.
     ǹ-wá-lá   júe   lē   A.
     DEM-NC1-DET name.NC6 CON A.
     'The other one's name is Akanchiem-anyami.' [BL (1b-sess) 2004: 222]

There is also a subtype of thetic nominal predication which uses a demonstrative identifier for identification and serves the presentation of entities "out of the blue" (18)-(19). The subject can thus be introduced into the discourse without anchoring it via situational deixis or other contextual information. This is achieved by way of a nominal dummy predicate following the connective particle *lē*. The dummy predicate, a syllabic nasal to which the determiner *lá* suffixes, functions like a demonstrative identifier that provides no lexical information but rather refers back to the initial subject noun.

(18) Nààwā   lē   ǹnā.
     nàa-wá   lē   ǹ-lá
     chief-DEF1 CON DEM-DET
     'This is the chief.' (Schwarz 2007: 88)

(19) Ǹ̀ bìsáŋá   wàɲī   àlē   ǹnā.
      m̀=bí-sa-ŋá   wà=ɲī   à=lē   ǹ-lá
      1s=child-NC12-DEF6 NC1=one &=CON DEM-DET

      'This is one of my children.' (Schwarz 2007: 88)

It is noteworthy that in the absence of a content verb in nominal predication, the connective particle *lē* resembles a copula verb. By hosting the prosodic linker (and even proclitic subject pronouns in embedded sentences, see §3.5) it is also more similar to other verbs (and some connectives) than is its postverbal counterpart *ká* of the categorical construction. Would the particle *lē* therefore be better analyzed as a defective copula verb that is also used in clefts for the thetic encoding? Schwarz (2009b) argues that synchronically neither *lē* nor *ká* are copulas, even though there could be a historical relation between the connective particle *lē* (*nē*) and a widely attested old Niger-Congo copula that surfaces as *ni* or in a somewhat similar form in various languages. Interestingly, the only preposition in Buli, *lè* (allomorph *nè*) 'with, and', differs only in tone from the predicate-initial connective particle found in thetic statements (cf. also Schwarz 2010a). Hence there is a tonal and a syntactic differentiation between two linking lexemes having the segmental structure *le (ne)* (Table 3).

Table 3: Comparison of linking morphemes with segmental structure *le*

lè (*nè*) 'with, and'	preposition (NP/argument-initial) in comitative and coordinative function	[*wá **lè** nààwā*] … 'he **and** the chief …'
lē (*nē*) (equative 'be')	connective particle (VP/predicate-initial) in thetic statements	*wá **lē** nààwā* 'he **is** the chief'

Both elements share connecting/coordinating semantics and both have prepositional properties, but they differ with respect to their tones, their syntactic domains and consequently also their predicative force. Despite these differences, it is very likely that they are historically related.

So far we have seen that there is a systematic formal distinction between categorical and thetic statements in Buli, which applies to sentences with verbal and nominal predicates alike. Simple categorical statements contain the particle *ká* in affirmation, the position of which is restricted to the postverbal domain (if there is a verbal predicate) and seems to be able to single out focal constituents to some degree. Subject and verb are in such a close relationship that the verb even displays subject agreement tone in the perfective. Thetic statements, on the other hand, contain the connective predicate-initial particle *lē* and never the postverbal particle *ká*. In the perfective, there is no tonal subject agreement at the verb, and the verb in the imperfective resembles the dependent, non-

initial verb in a series. The verbal predicate of the thetic statement is obviously less finite than that of the categorical statement, even though it still functions as predicate. Subject and the dependent verb form are markedly concatenated by means of the particle *lē*.

In the following, I outline the major pragmatic and syntactic conditions under which this marked subject-verb concatenation is very common in Buli.

3.4 The pragmatics of thetic statements

Thetic statements have no internal information structure. Semantically and pragmatically, the thetic configuration responds to discourse conditions in which a sentence topic is (a) either not available or (b) not wanted for discourse-structuring purposes. For example, the thetic encoding is used whenever the speaker wants to deliberately interrupt the coherence of the discourse and set the following paragraph apart from the previous text. This is most typically the case when new scenes, major participants, and unexpected events are presented.

We find the thetic encoding very often with indefinite subjects, while such subjects are rather rare in simple categorical statements. The thetic encoding with indefinite subjects is characteristic for existential clauses, such as in (20). Here the existence or presence of the entities referred to by the subject noun is asserted with the help of a posture verb.

(20) Yíé   ŋàyɛ̀   lē   zà.
     jíe   ŋà=jɛ̀   lē   zà
     house.NC6 NC6=two CON stand
     'There are two houses.' [BL (1b-sess) 2004: 218]

The high frequency of indefinite subjects in thetic statements is a result of the scene-setting function of the thetic configuration, among others, with respect to the presentation of major discourse participants. Example (21) illustrates the introduction of the two protagonists at the beginning of a story, immediately after the time frame has been set.

(21) núr   nè   wà dɔ̄ā   àlē   jàm bōrɔ́.
     núr   nè   wà=dɔ̄a   à=lē   dʒàm bō-rɔ́
     person.NC1 with NC1=friend.NC1 &=CON PAST exist-LOC
     'there was a man and his friend.' [Azuima 2]

The scene-setting function of thetic statements often also consists of reporting particular events. Such reports are typically about sudden changes and include a predicate in the perfective (22); but if the situation persists, imperfective encoding may also be encountered (23).

(22) Nááb   lē   bìàg   bìīk.
     náab   lē   bìag   bíik
     cow.NC14 CON give.birth child.NC12
     'A calf has been born.' [Heiss 392 PA]

(23) M̀ bìīk   làā   yūāk.
      m̀=bíík   lē-á   jūag
      1s=child.NC12 CON-IPF become.sick
      'My child got sick (and still is).' [BL (2) 2005: 0450-2]

Examples (21)-(23) show that the structural thetic properties in Buli remain the same – the predicate-initial connective particle *lē* and a less finite, dependent verb form – whether the thetic statement states the existence of an entity (21) or that of an event (22)-(23). The two subtypes were identified by Sasse (1987: 526f.) as "entity-central" and "event-central" thetic statements and are only in some languages formally differentiated. Pragmatically, both subtypes lack any internal structuring, but semantically, they focus on different ontological categories.

The formal properties of thetic statements can also be regularly observed in the responses to questions of the 'What happened?' type, in particular if the reply does not include impersonal pronouns but nouns in subject function (24). While some scholars interpret such responses as sentence focus occurrences (using Wh-questions as focus diagnostics; cf. also Fiedler et al. 2010), in line with Sasse (1987: 572f.) and Güldemann (2010: 86) we can readily identify them as event-central thetic statements.

(24) Gàlásì   àlē   lò tēŋ   à mɔ̀b.
      galasi   à=lē   lò tēŋ   à=mɔ̀b
      glass.NC1 &=CON fall ground.NC12 &=break
      'A glass has fallen down and got broken.' (reply to: 'What happened?') [BL (3) 2005: 0448-1]

The formal properties of thetic statements are also regularly used when the focus of assertion is on the subject constituent, as in (25). The utterance represents the direct response to an information question about the identity of the agent (which demands subject function). The utterance can be interpreted as an entity-central thetic statement which states the identity of the subject referent in respect to the rest of the proposition. Regardless of the particular subtype, it can be concluded that the thetic encoding is chosen in order to revoke the categorical interpretation of the simple clause; since otherwise the subject would be regarded the unmarked topic and the focus would be restricted to information pertaining to the comment.

(25) (Ká) Mary àlē   ŋɔ̀bī.
      ká   Mary à=lē   ŋɔ̀bī
      KA   M.   &=CON eat.ASS
      'Mary ate them.' (reply to: 'Who ate the beans?') (Fiedler et al. 2010: 246)

As indicated by the particle in parenthesis in (25) and by (26), thetic statements of both subtypes can also begin with the sentence-initial particle *ká*.

(26) Ká bòlīm   nē chɔ̀gsì yèrí   á dɛ̄.
      ká bòlīm   lē chɔ̀gsi jéri   á dɛ̄
      KA fire.NC14 CON catch house.NC5 IPF eat

      'A house has caught fire.' (reply to: 'What happened?') [BL (1b-sess) 2004: 414.15]

The presence of the particle *ká* is grammatically optional and seems to be pragmatically determined. From the data at hand it appears that the sentence-initial particle has scope over the whole thetic construction and is overall more frequent in thetic statements that occur in dialogues and respond to immediate verbal contexts. I therefore hypothesize that the particle *ká* both supports discourse cohesion and is part of a comment that may also include an embedded thetic statement (27). The topic of this comment is not expressed in the sentence itself, but is provided in the previous discourse context.[3]

(27) [Categorical [TOPIC Ø-expression] [COMMENT *ká* [Thetic ... *lē* ...]]]

3.5 Embedding with thetic encoding

One of the striking features of thetic encoding in Buli is the frequency with which the formal properties occur in various grammatical functions. In the following, I present some common cases of embedded clauses that are based on the same morphosyntax as thetic statements.

One of the two relative clause strategies in Buli, the head-internal relative clause which relativizes subjects and sometimes the objects of the matrix clause (cf. also Hiraiwa 2003), employs the structural means used for thetic statements. The head-internal relative clause in (28) refers to the object the interlocutor sees, and her perception represents the object of the matrix clause, a request to provide a thorough description. In general, relativized head nouns in Buli are marked by an indefinite noun class pronoun or, in case of a headless clause as in (28), often by an indefinite pronoun of noun class 5. The predicate of the head-internal relative clause always begins with the connective particle *lē*. In addition to the basic thetic encoding, relative clauses as well as other embedded clauses typically end with the clause-final determiner *lá,* as is the case in the following example.

(28) Māgsī [fi làā   ɲā dīī lá].
      māgsi fi=lē-á   ɲā dīī lá
      describe 2s=CON-IPF see IND5 DET

      'Describe what you see!' [BL (1b-sess) 2004: 106]

Other embedded clause types regularly using the thetic encoding plus the clause-final enclitic determiner *lá* are sentence-initial clauses with adverbial functions as well as what are called TAIL-HEAD LINKAGE clauses. The latter term describes a "discourse pattern which consists in repeating, at the beginning of a new sentence, the main verb of the

---

[3] The etymology of the particle *ká* is not clear, but it is possible that it originated from a proform of noun class 12 (anaphoric pronoun *kà=*).

preceding sentence for discourse cohesion" (Guillaume 2011: 109). The linkage clause in (29) repeats propositional content of the immediately preceding sentence.

(29) '… and he caned me.'

Wà nē   mìì mì lá,   mí yàà   kàláá kūm …
wà=lē   mìì mí=lá   mí jàa   kàli-á kūm
NC1=CON cane 1S.D=DET 1S.D TENSE sit-IPF cry

'He caned me/when he had caned me, and I was then sitting crying …'
[Karichiwade 14-15]

Note that the connective particle *lē* can host proclitic subject pronouns (29)-(29), as if it were a (stative[4]) verb. Bound subject pronouns at the connective particle are not uncommon in adverbial and linkage clauses with thetic structures and they can also be found in head-internal relative clauses, the object of which is relativized. Bound subject pronouns are absent, however, from entity-central thetic statements (or subject focus occurrences), where nominal or disjunctive emphatic subject pronouns prevail for semantic/pragmatic reasons.

Summarizing a particular discourse topic for the interlocutor is another common situation where thetic structures occur. Speakers often present the title of their monologue at the beginning or end of their speech, as illustrated by the stacked thetic construction in (30). The initial embedded thetic structure provides the content description in form of a subject clause. The matrix clause presents the title to the listeners by means of a matrix thetic structure that contains the demonstrative identifier (*ǹnā*). The speaker summarizes the topic of her speech by means of a headless relative clause 'how we celebrate our Fiok festival in Buluk'. Instead of a lexical head, it contains an indefinite pronoun in noun class 5 (*dīī*) which contributes to the circumstantial manner interpretation. The predicate in the relative clause includes the connective particle *lē*, here with the prosodic linker (à=). The verb *dɛ* 'eat', here with the sense 'celebrate', comes in a dependent verb form marked for the imperfective aspect by preceding morpheme (*á*). The embedded clause ends with the determiner *=lá* and represents the clausal subject in a thetic matrix clause. The connective particle *lē* of the matrix clause is followed by the demonstrative identifier and fulfils the presentational function.

(30) [Tàmā bùlìdɛ̄   ànàā   dɛ̄ tì fìɔ̀kū   dīī lá]   àlē   ǹnā.
      tàmá bùli-dɛ   à=lē-á   dɛ̄ tì=fìɔ-kú   dīi=lá   à-lē   ǹ-lá
      1P.D Bulsa-here &=CON-IPF eat 1P=harvest-DEF17 IND5=DET &-CON DEM-DET

'This is how our fiok festival is celebrated in Buluk.' [Fiok 1]

Another example with a clausal subject and the demonstrative identifier as nominal predicate is given in (31). The speaker announces her future career aspirations as the topic of her upcoming talk. In contrast to (30), she here also adds the sentence-initial particle *ká* which explicitly connects her presentation to my earlier request for her speech.

---

[4] The tonal properties of the connective particle with the subject proclitic differ from those with dynamic verbs, but resemble those with the irregular stative verb class.

(31) Ká [mí nàà   yāā mí chúmdīīpō   bōká   à ɲɛ̄ dīī lá]
     ká mí lē-á   jāa mí chúm   dīi-pō bō-ká   à=ɲɛ̄ dīi=lá
     KA 1S.D CON-IPF want 1S.D tomorrow IND5-in exist-DEF12 &=do IND5=DET
     àlē   ǹnā.
     à=lē   ǹ-lá
     &=CON DEM-DET

     'This is how I want my future career to be.' [Chumdiipo 21]

These common embeddings of thetic structures are schematically summarized in (32). Scheme A in (32) reflects the thetic encoding in sentence-initial dependent clauses which have adverbial or discourse linkage function, as illustrated in (29). Scheme B refers to the stacking of two thetic constructions, as illustrated by the presentation of the discourse title in (30). Scheme C is an extension of the stacked thetic structures. The embedding thetic structure gets itself embedded in a categorical statement, as indicated by the sentence-initial particle *ká* and illustrated in (31). A simpler version of the scheme in C with just one thetic structure embedded in a categorical statement is given in (27) above.

(32) A. [Thetic … *lē* …]ADVERBIAL/LINKAGE + Matrix clause
     B. [Thetic [Thetic …*lē*…] …*lē* …]
     C. [Categorical [TOPIC Ø] [COMMENT *ká* [Thetic [Thetic …*lē*…] …*lē* …]]

4 Summary

This paper has described the structural and pragmatic features of the thetic statement and its (simple) categorical counterpart in Buli. Buli grammar applies a connective particle with prepositional properties for the encoding of thetic statements. The connective particle (*lē*) differs only tonally from the comitative and the NP coordinating preposition. In the thetic statement, it is followed either by a predicative noun or by a dependent verb form, as evidenced by the tone of the perfective verb and the tone of the preverbal imperfective marker, respectively. While the connective particle *lē* literally *connects* the predicate to a non-topical subject, both elements are often prosodically separated. The optional presence of the predicate-initial prosodic linker, which appears verb-phrase initially after a pause, indicates that subject and predicate are not necessarily within the same intonational phrase in thetic statements in Buli.

Thetic statements are used to present entities and report events for various discourse structuring reasons. Thetic encoding is also used in response to information questions concerning the referent of a subject constituent or an event. The function of thetic encoding consists in such cases in the cancelation of the prototypical categorical interpretation in favor of a single information chunk. The application of a simple concatenating morphosyntactic device to this effect, which can be easily embedded, has further contributed to the success of thetic encoding in Buli grammar and the apparent polyfunctionality of thetic structures.

Acknowledgements

Abbreviations

ASS	assertive	IPF	imperfective
CNJ	conjunction	LOC	locative
CON	connective	NC	noun class
CONF	confirmation	P	plural
D	disjunctive pronoun	PROX	proximal
DEF	definite	Q	question marker
DEM	demonstrative	S	singular
DET	determiner	TBU	tone bearing unit
IND	indefinite pronoun	&	prosodic linker

References

Cahill, Michael C. 2007. *Aspects of the morphology and phonology of Kɔnni*. Dallas: SIL International.

Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz & Malte Zimmermann. 2010. Subject focus in West African languages. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives*, 234–257. Oxford: Oxford University Press.

Guillaume, Antoine. 2011. Subordinate clauses, switch-reference, and tail-head linkage in Cavineña narratives. In Rik van Gijn, Katharina Haude & Pieter Muysken (eds.), *Subordination in native South American languages*, 109–139. Amsterdam: John Benjamins.

Güldemann, Tom. 1996. *Verbalmorphologie und Nebenprädikationen im Bantu – Eine Studie zur funktional Motivierten Genese eines konjugationalen Subsystems*. Bochum: Universitätsverlag Dr. N. Brockmeyer.

Güldemann, Tom. 2010. The relation between focus and theticity in the Tuu family. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure – a documentation of its diversity across Africa*, 69–93. Amsterdam: John Benjamins.

Güldemann, Tom. 2012. Thetic speaker-instantiating quotative indexes as cross-linguistic type. In Isabelle Buchstaller & Ingrid Van Alphen (eds.), *Quotation: Cross-linguistic and cross-disciplinary perspectives*, 117–142. Amsterdam: John Benjamins.

Hiraiwa, Ken. 2003. Relativization in Buli. In George Akanlig-Pare & Michael Kenstowicz (eds.), *Studies in Buli grammar* (Working Papers on Endangered and Less Familiar Languages 4), 45–84. Cambridge, MA: MIT.

Hyman, Larry M. & John R. Watters. 1984. Auxiliary focus. *Studies in African Linguistics, Supplement* 15(3). 233–273.

Kuroda, Shige-Yuki. 1972. The categorical and the thetic judgement: Evidence from Japanese syntax. *Foundations of Language* 9(2). 153–185.

Lambrecht, Knud. 1987. Sentence focus, information structure, and the thetic-categorical distinction. In Jon Aske, Natasha Beery, Laura A. Michaelis & Hana Filip (eds.), *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society – General Session and Parasession on Grammar and Cognition*, 366–382. Berkeley, California: Berkeley Linguistics Society.

Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language* 24(3). 611–682.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the world.* 18th edn. Dallas, TX: SIL International. http://www.ethnologue.com.

Manessy, Gabriel. 1979. *Contribution à la classification généalogique des langues voltaiques.* Paris: Centre National de la Recherche Scientifique; SELAF.

Naden, Anthony J. 1989. Gur. In John Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages,* 140–168. New York: University Press of America.

Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists.* Cambridge: Cambridge University Press.

Reinhart, Tanya. 1982. Pragmatics and linguistics. An analysis of sentence topics. *Philosophica* 27. 53–94.

Robert, Stéphane. 2000. Le verbe wolof ou la grammaticalisation du focus. In Bernard Caron (ed.), *Topicalisation et focalisation dans les langues africaines*, 229–267. Louvain – Paris: Peeters.

Sasse, Hans-Jürgen. 1987. The thetic/categorial distinction revisited. *Linguistics* 25(3). 511–580.

Sasse, Hans-Jürgen. 1995. 'Theticity' and VS order: A case study. In Yaron Matras & Hans-Jürgen Sasse (eds.), *Verb-subject order and theticity in European languages*, 3–31. Berlin: De Gruyter.

Schwarz, Anne. 2003. Low tone spreading in Buli. *Cahiers Voltaïques / Gur Papers* 6. 121–130.

Schwarz, Anne. 2007. *Aspekte der Morphosyntax und Tonologie im Buli – Mit Schwerpunkt auf dem Buli von Wiaga.* Berlin: Humboldt-Universität zu Berlin. (Doctoral dissertation.) www:urn:nbn:de:kobv:11-10085362.

Schwarz, Anne. 2009a. To be or not to be? About the copula system in Buli (Gur). In Margarida Petter & Ronald Beline Mendes (eds.), *Proceedings of the Special World Congress*

of African Linguistics – São Paulo 2008: Exploring the African language connection in the Americas*, 263–278. São Paulo: Humanitas.

Schwarz, Anne. 2009b. Tonal focus reflections in Buli and some Gur relatives. *Lingua* 119. 950–972.

Schwarz, Anne. 2010a. Discourse principles in grammar: The thetic/categorical dichotomy. In *Proceedings of the Inaugural Tropics of the Imagination Conference, James Cook University Cairns, 2-3 November 2009* (Etropic 9). https://journals.jcu.edu.au/etropic/article/view/3425/3363.

Schwarz, Anne. 2010b. Verb-and-predication focus markers in Gur. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure. A documentation of its diversity across Africa*, 287–314. Amsterdam: John Benjamins.

Skopeteas, Stavros, Ines Fiedler, Samantha Hellmuth, Anne Schwarz, Ruben Stoel, Gisbert Fanselow, Caroline Féry & Manfred Krifka. 2006. *Questionnaire on information structure (QUIS)* (Interdisciplinary Studies on Information Structure 4). Potsdam: University of Potsdam.

Ulrich, Miorita. 1988. Thetisch vs. Kategorisch und Informationsstruktur. In John Albrecht, Harald Thun & Jens Lüdtke (eds.), *Energeia und Ergon: Sprachliche Variation – Sprachgeschichte – Sprachtypologie*, 387–399. Tübingen: Narr.

Wolff, H. Ekkehard. 2005. Intrinsic focus and focus control in two varieties of Hausa. In F. K. Erhard Voeltz (ed.), *Studies in African linguistic typology*, 397–415. Amsterdam: John Benjamins.

Subpart I-C

Kwa

Chapter 6

Beyond resumptives and expletives in Akan

Sampson Korsah

Universität Leipzig

In Akan, resumption is obligatory for extraction from a subject position. Accordingly, there is usually agreement between the resumptive pronoun (RP) and its referent constituent. However, data from the Asante-Twi dialect suggest that it is possible to have a non-agreeing pronoun in the extraction site. But this is only possible for the highest subject position. In this paper I show that what occupies the subject position in the non-agreeing situation is an instance of the insertion of a default pronoun. Gaps in subject positions are not allowed at all in Akan. In order to avoid violating this constraint in the case of the highest subject, the language has two options; it either uses a resumptive pronoun, or a default pronoun.

1 Introduction

It is a standardly held view in Generative syntax traditions that follow the Chomskyan approach, e.g. Chomsky (1981), that various constituents in a given syntactic structure may be extracted from argument positions (A-positions) to non-argument positions (Ā-positions) for various information structure purposes. For instance, the internal argument of a verb may be fronted to head a relative clause, as in (1). I will refer to such extractions as Ā-operations.

(1)   [The woman]$_i$ who ~~the woman~~$_i$ bought the car is rich.

   Ā-operations in languages may exhibit various kinds of reflexes (Georgi 2014). In some languages, e.g. English, the extraction site shows no phonetic signs of such operations. In such cases, there are gaps, see e.g. Salzmann (2011). But there are languages which do not permit gaps; rather they require resumptive pronouns (RP) in the extraction site. For languages that allow RPs, some allow gaps only in certain positions (see Klein 2014). Unlike languages like English however, Akan (Kwa, Niger-Congo) sometimes does not permit gaps in an extraction site. As (2) shows, the RP is always obligatory.[1]

---

[1] The data presented here is based on the intuitions of three native speakers. Akan has both high and low tones. But I will mark only high tones.

Note that an RP agrees with its referent NP in terms of number, person, animacy, and case features. For instance, *Amma* in (2a) is a third person singular animate NP which has been extracted from a subject position, and thus has nominative case. *Kofi* (2b) has similar properties except that, because it is extracted from an object position, it has accusative case.

(2) a. Amma$_i$ na   *Ø$_i$/ɔ$_i$-hú-u   Kofi.
         A.   FOC 3SG.NOM.ANIM-see-PST K.

         'AMA [and not, say, John] saw Kofi.'

     b. Kofi$_i$ na   Amma hú-u   *Ø$_i$/nó$_i$ nó.
         K.   FOC A.   see-PST 3SG.ACC.ANIM CD

         'Ama saw KOFI [and not, say, John].'

This paper focuses on resumption in Akan, particularly resumption that affects constituents in subject position in the Asante-Twi dialect, as in (2). (Thus, use of the term *Akan* here is restricted to this dialect only.) Where necessary, the facts of the other dialects will be pointed out. The interesting thing about subject resumption in Akan is that sometimes the expected agreement between an extracted NP and its RP does not seem to obtain. The pronominal form that is used in this non-agreeing configuration is *ε-*. In (3), although the subject NP is animate, its RP does not necessarily agree with it in terms of number and animacy. Yet, (3) is perfectly grammatical (but see Marfo 2005).

(3) Amma$_i$ na   ε$_i$-hú-u   Kofi.
     A.   FOC 3.NOM.-see-PST K.

     'AMA [and not, say, John] saw Kofi.'

This paper has two main aims regarding this phenomenon. First, it examines the status of the non-agreeing *ε-* in relation to other homophonous pro-forms in Asante-Twi. I present evidence to show whether it is a regular RP, an expletive, or a default pronoun. Second, the paper investigates the constraints on the distribution of the non-agreeing *ε-* vis-à-vis the other competing forms.

Regarding the former aim, I will argue that the non-agreeing *ε-* is a special kind of pronoun; though it is referential, it has certain underspecified pronominal features. This makes it the default pronoun, and it is inserted in contexts where the (agreeing) RP is not possible due to certain constraints. Regarding the latter aim, I will assume, following Korsah & Murphy (2015), and contrary to Saah (1994; 2010), that Ā-operations in Akan are movement-based, and that in the particular instances of Ā-operations involving subjects, two factors affect the nature of what can occupy the extraction site. If the extraction is done from an embedded position, then the extracted constituent can only reach its final Ā position via the intermediate subject position of the clause. In such a situation, only an (agreeing) RP is permitted at the extraction site. But when the extraction is done from a non-embedded, i.e. the highest, subject position then Asante-Twi has the option of either skipping the subject position, or going through it. When the former option is adopted, then the default pronoun *ε-* is inserted. This insertion is necessary in order to

repair a violation of a constraint that places a ban on gaps in subject position in Akan syntax.

The remainder of this paper is structured as follows: §2 gives a general overview of resumption in Akan, and how it can be analyzed. §3 shows instances of agreement mismatches, and discusses the kinds of features that the relevant pronouns have. §4 deals with the constraints on the distribution of the non-agreeing pronoun. §5 is the conclusion.

2 Resumption in Akan

In Akan, it is possible to extract NPs from various argument and non-argument positions (Marfo & Bodomo 2005; Saah 2010) for various Ā-operations. In the present discussion, I will focus on extraction from only subject and object positions. As in many languages (Keenan & Comrie 1977; Klein 2014), the resumptive pronouns correspond to the personal pronoun paradigm (see Table 1).[2]

Table 1: Pronouns in Akan

NUMB	PERS	NOM	ACC
SG	1	*me-*	*me/-m*
	2	*wo-/i*	*wo/-w*
	3 (ANIM)	*ɔ-/no*	*no/-n*
	3 (INANIM)	*ɛ-no/ ɔ-no*	*(no/-n)*
PL	1	*yɛ-*	*yɛn/hɛn*
	2	*mo-/hom-*	*mo/hom*
	3 (ANIM)	*wɔ-*	*wɔn/hɔn*
	3 (INANIM)	*ɛ-(no)/ɔ-(no)*	*wɔn/hɔn*

In this section, I discuss three issues: the nature of subject and object resumption, how they may be explained, and how to deal with a constraint that I will refer to as the NO-SUBJECT-GAP (NSG) constraint.

2.1 Subject and object resumption

As indicated earlier, Akan allows (and sometimes requires) both subject and object resumption. Resumption is obligatory for all extracted subjects , whether the extracted constituent is animate (4a), or inanimate (4b). Similar restrictions have been reported for languages like Hebrew (Rizzi & Shlonsky 2007). For instance, while it is fine to extract *ha-sulxan* as the object of *raca* (5a), extracting *ya'ale* which is in subject position renders the construction ungrammatical (5b).

---

[2] Where there are options, the forms to the right of the slash are used in the Fante dialect.

(4) a. Ɔbááᵢ áa *Øᵢ/ɔᵢ-wáré-e Kofi nó fi Aburi.
       woman REL 3SG.NOM-marry-PST K. CD be.from A.

    'The woman who married Kofi is from Aburi.' (Saah 2010: 92)

  b. [Kŕataá nó]ᵢ áa *Øᵢ/ɛᵢ-dá pónó nó só nó yɛ́ fɛ.
       paper DEF REL 3SG.NOM.INANIM-lie table DEF top CD be nice

    'The paper that is on the table is nice.'

(5) *Hebrew* (Rizzi & Shlonsky 2007: 120-121)

  a. kaniti et ha-šulxanᵢ še oto xana amra še dalya ma'amina še
     (I).bought ACC the-table that him Hannah said that Dalya believes that
     Kobi raca Øᵢ.
     Kobi wanted

    'I bought the table that Hannah said that Dalya believes that Kobi wanted.'

  b. *kaniti et ha-sulxanᵢ še hu xana amra še dalya ta'ana še Øᵢ
     (I).bought ACC the-table that he Hannah said that Dalya claimed that
     ya'ale harbe kesef.
     will.cost a.lot money

    'I bought the table that Hannah said that Dalya claimed that will cost a lot of money.'

The agreement requirements between the RP and the extracted NP in Akan include person, number, animacy, and case specifications. (I defer a discussion of how Akan handles the ban on gaps in the extraction sites of subjects to §2.3.)

For object extraction, resumption is obligatory for only animate NPs. For instance, the RP *nó* in (6) is obligatory. For inanimate object NPs, however, the overt realization of an RP appears to be optional; gaps are sometimes permitted, as comparison of (7a) and (7b) shows. This distribution of gaps and RPs with regards to extracted inanimate object NPs may be attributable to independent properties of either the Akan pronoun system, or the verbs involved (see e.g. Chinebuah 1976 and Larson 2005 properties in other Kwa languages).

(6) Me-huu ɔbááᵢ áa Kofi wáre-e nóᵢ/*Øᵢ nó.
    1SG-see woman REL K. marry-PST 3SG.ACC CD

    'I saw the woman whom Kofi married.' (Saah 2010: 92)

(7) a. [Kŕataá nó]ᵢ áa Kofi hú-ui *nóᵢ/Øᵢ nó da pónó nó só.
       paper DEF REL K. see-PST 3SG.ACC CD lie table DEF top

    'The paper that Kofi saw is on the table.'

  b. [Kŕataá nó] áa Kofi té-e nóᵢ/*Øᵢ nó da pónó nó só.
       paper DEF REL K. tear-PST 3SG.ACC CD lie table DEF top

    'The paper that Kofi tore is on the table.'

The data that we have seen so far show that resumption may result from extraction within a single clause. But unlike languages like Tsez, as reported by Polinsky & Potsdam (2001), resumption in Akan is not clause-bound; it is possible to extract NPs from deeply-embedded contexts, across several clauses. In (8), for instance the extracted NP *ɔbáá* spans three clauses i.e. three CPs.

(8) Me-hu-u [$_{NP}$ ɔbáá$_i$ áa Ama pɛ́ [$_{CP}$ sɛ óbíárá té [$_{CP}$ sɛ Kofi
    1SG-see-PST woman REL A. like COMP everybody hear COMP K.
    á-ká [$_{CP}$ sɛ ɔ-bɛ́-waré nó$_i$ nó]]]].
    PERF-say COMP 3SG.NOM-FUT-marry 3SG.ACC CD

    'I saw the woman whom Ama wants everybody to hear that Kofi has said he will marry.' (lit: 'I saw the woman whom Ama wants everybody to hear that Kofi has said that he will marry her.')

2.2 Analysis of Akan resumption

Resumption in Akan has been traditionally analyzed as involving base-generation (see Saah 1992; 1994; 2010). This view suggests that resumption is not due to movement but rather the result of binding. However, following Korsah & Murphy (2015), I will assume that resumption in Akan involves movement.

The main argument for the base-generation approach to resumption is that resumption is possible in contexts which have been argued to be syntactic islands in languages like English. For instance, movement of *who* (9b), out of the relative clause for the purposes of question formation is illicit. This is because relative clauses are syntactic islands (Ross 1967). Furthermore, movement out of such syntactic configurations results in ungrammaticality cross-linguistically.

(9) a. John met the lady that owns the publishing company.
   b. * Who$_i$ did John meet Ø$_i$ that owns the publishing company?

Given the above, displacing *Sikaní* in (10) is predicted to be illicit if it involves movement. But, it is completely acceptable in Akan (11). It follows, then, that whatever process results in such NPs ending up in Ā positions in Akan could not involve movement. Under this assumption, the NP in the Ā position is assumed to be directly merged at its surface position. Its surface relationship with the RP is established via semantic binding, as sketched in (12). A similar analysis has been proposed by McCloskey (2011) for Irish.

(10) Me-**nim** baabi aa Sikaní nó fí.
     1SG-know where REL rich.man DEF come.from

     'I know where the rich man comes from.'

(11) Sikaní$_i$ áa me-**ním** baabi aa ɔ$_i$-fí nó á-da.
     rich.man REL 1SG-know where REL 3SG.NOM-come.from CD PERF-sleep

     'The rich man who I know where he comes from is asleep.'

(12) Base-generation
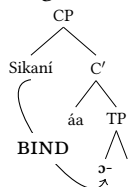
(13) Movement
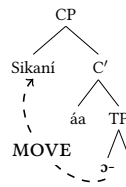
Contrary to the base-generation approach, the movement approach would proceed as sketched in (13). Here, the target NP is literally extracted from one position to the final landing site (in the direction of the dashed arrow). (13) is a simplified illustration, but when the extraction is done from a much lower position in the structure, the movement path includes all available intermediate landing sites. This idea is succinctly expressed by the notion of Successive Cyclic Movement proposed by Chomsky (1977). Going by the movement approach, the explanation for the agreement relationship between an RP and its antecedent NP in Ā contexts is that a resumptive pronoun is the most economical way of realizing the features of the copy of an extracted NP, see e.g. Nunes (2004).

One core issue to be addressed if one is to account for resumption in Akan in terms of movement is the empirical justification for movement in the language. Empirical evidence in Akan comes from tones. Korsah & Murphy (2015) show that there are movement reflexes on the stem of every verb across which movement has taken place. This is registered in the tonal changes. For instance, when we compare the tone of the verb stem *nim* in (10), in which the target NP is in-situ, with the same verb in (11), where the NP has moved, we observe this tonal difference. The claim here is that the tonal change from low to high on the verb *nim* indicates that *Sikaní* has undergone movement. Thus, for Akan, we need not rely solely on island effects when talking about Ā-operations; evidence may be found in phonological reflexes of such operations. The interested reader is referred to Korsah & Murphy (2015) for the details. But as far as the present discussion is concerned, this will be the theoretical assumption for all Ā-operations.

2.3 Dealing with the *no-subject-gap* constraint

We showed in §2.1 that Akan permits no gaps in subject positions. This property seems to fit into a larger picture which suggests that some languages tend to disprefer extraction from subject positions. However, sometimes it is imperative for languages to deploy strategies that force extraction from such positions, and thus a violation of this constraint becomes inevitable. When this is the case, then languages tend to adopt one of the following strategies as a repair mechanism.

(14) Repair strategies for NSG (see Rizzi & Shlonsky 2007)

   a. Base-generation or movement, resulting in a resumptive pronoun

   b. Skipping subject position, and filling it with an expletive pronoun

Regarding (14a), I follow Shlonsky (1992) and propose that Akan uses RPs as a repair strategy against violating the NSG. I will demonstrate later that the language also has the option of using what looks like a version of the EXPLETIVE INSERTION strategy in (14b).

2.4 Summary

So far we have made the following observations and arguments about Ā-operations, and resumption in Akan. First, when the extraction is from a subject position, it is obligatory to overtly realize an RP. But from object positions, this may be optional. Second, the most crucial property about the resumption process for the present discussion is the requirement that the extracted NP and its RP must agree in terms of number, person, animacy, and case features. Third, Ā-operations in Akan involve movement. Fourth, Akan uses an RP as a repair strategy for what would otherwise be a violation of the NSG constraint.

3 Subject agreement mismatches

The agreement requirement between a moved NP and its RP is particularly crucial for subjects, given the NSG constraint. However, I show in this section that a non-agreeing pronominal element ε- may fill the extraction site of the highest subject position. I will claim that this is a default pronoun; it is less specified than the inanimate ε- form in Table 1, and it is also not an expletive pronoun.

3.1 Non-agreeing subject pronoun

In many contexts (particularly of spoken Asante-Twi), sometimes the expected agreement between an extracted NP and its RP does not obtain. Consider (15)-(17), where I have glossed the otherwise RP as 'default' (DFLT). I will argue later on that it is a DEFAULT pronoun; see also McCracken (2013). The expected RP for *ɔbáá* in (15) is *-ɔ*, i.e. 3SG.SUBJ.ANIM, and that of *Nnípá dú nó* in (16) is *wɔ-*, i.e. 3PL.SUBJ.INANIM. Yet speakers of Asante-Twi have absolutely no problem with parsing (15)-(16).

(15) Ɔbáá$_i$ áa ε_i-wáré-e Kofi nó fi A.
woman REL DFLT-marry-PST K. CD from A.
'The woman who married Kofi is from Aburi.'

(16) [N-nípá dú nó]$_i$ áa ε_i-hyíá-έ nó ma-a hɔ́ sɔ-ɔe.
PL-person ten DEF REL DFLT-meet-PST CD cause-PST there light-PST.
'The ten people who met made the place exciting.'

With (17), it is not obvious whether the ε- is the non-agreeing form since it is homophonous with the expected RP; see Table 1. Thus syncretism cannot be ruled out in these contexts. I do not intend to address this issue in this paper.

(17) a. [Kŕataá nó   ]$_i$ áa   ɛ$_i$-yérá-aɛ   nó níé.
         paper DEF   REL DFLT/3.SG.INANIM-lost-PST CD this

      'This is the paper which got missing.'

    b. [N-kŕataá **nó**]$_i$ áa   ɛ$_i$-yérá-aɛ   nó níé.
         PL-paper DEF   REL DFLT/3.PL.INANIM-lost-PST CD this

      'These are the papers which got missing.'

The agreement mismatch observed above is not restricted to only relative clause constructions. As (18) and (19) show, it also obtains in focus constructions and ex-situ content questions. Note that for constructions like (19b), some speakers prefer the non-agreeing pronoun over the agreeing form.

(18) a. [Ɔbáá   yí]$_i$ na ɔ$_i$-/ɛ$_i$-bɛ́-wáré   Kofi.
         woman DEM   FOC 3SG.ANIM.NOM/DFLT-FUT-marry K.

      'This woman (as opposed to some other woman) will marry Kofi.'

    b. [Nnípá **dú** pɛ́]$_i$ na wɔ$_i$-/ɛ$_i$-hyíá-ɛ́.
         people ten only   FOC 3PL.ANIM.NOM/DFLT-meet-PST

      'Only ten people (as opposed to more people) met.'

(19) a. Hwán$_i$ na ɔ$_i$-/ɛ$_i$-wáré-e   Kofi?
         Who   FOC 3SG.ANIM.NOM/EXPL-marry-PST K.

      'Who married Kofi?'

    b. Hwán-mó$_i$ na   ?wɔ$_i$-/ɛ$_i$-hyíá-ɛ.
         Who-PL   FOC 3PL.ANIM.NOM/DFLT-meet-PST

      'Who met?'

The agreement mismatch reported here obtains only in the Asante-Twi dialect. For instance, (20b) shows that the equivalent of the non-agreeing *ɛ-* in similar syntactic contexts is illicit in the Fante dialect of Akan.[3]

(20) Fante dialect

    a. [Nyípá kór pɛ́]$_i$ na ɔ́$_i$-bá-í.
         person one only   FOC 3SG-come-PST

      'Only one person (as opposed to more people) came.'

    b. [N-nyípá dú pɛ́]$_i$ na wɔ́$_i$/*ɔ́$_i$-hyíá-í.
         PL-person ten only   FOC 3PL/DFLT-meet-PST

      'Only ten people (as opposed to more people) met.'

---

[3] Note that Fante uses the same form, i.e. *ɔ-*, for both 3SG.NOM.+/-ANIM, and EXPL(=non-referential pronoun).

3.2 Distribution of non-agreeing pronoun *ɛ-*

Interestingly, the non-agreeing *ɛ-* is restricted to only matrix subject positions; subjects of embedded clauses do not seem to allow this non-agreeing *ɛ-* (21a). Also, this *ɛ-* cannot refer to extracted objects, irrespective of their agreement features and level of embedding. Thus *ɛ-* cannot be co-indexed with *ɔbáá* in (21b).

(21) a. [N-nípá   dú nó]$_i$ áa me-ním   sɛ   wɔ$_i$-/*ɛ$_i$-hyíá-ɛ́   nó nié.
        PL-person ten DEF   REL 1SG-know that 3PL/DFLT-meet-PST CD DEM
        'These are the ten people who I know met.'

   b. Ɔbáá$_i$ áa Kofi wáré-e   nó$_i$/*ɛ$_i$ nó fi   Aburi.
      woman REL K.   marry-PST 3SG.ACC/DFLT CD be.from A.
      'The woman whom Kofi married is from Aburi.'

3.3 Summary

I have shown that in the Asante-Twi dialect of Akan, there is optional agreement between an extracted matrix subject NP and what replaces it at the extraction site. This is possible in all Ā-operations. The form *ɛ-* is always used in the contexts where this optionality is allowed. These observations raise at least two fundamental questions: First, what is the nature the pro-form *ɛ-* that is allowed in these non-agreeing configurations, and how different is it from other homophonous forms in the language? Second, why is it only the highest subject position that allows this agreement optionality? I address these issues in the next section.

4 Accounting for *ɛ-* pro-forms

In this section I propose an account for the non-agreeing pronoun and its distribution. As far as I can see, there are at least two ways to deal with this. One way is to assume that *ɛ-* is a default pronoun that is inserted whenever movement skips a subject position. The alternative approach is to assume that the regular RP *ɛ-* loses some of its pronominal features in the context of the matrix subject. In the present discussion, I will argue for the plausibility of the former. I will treat the non-agreeing *ɛ-* as a DEFAULT (DFLT) pronoun that is inserted whenever movement skips the (matrix) subject position. I show that this *ɛ-* has certain features that make it less specified as a personal pronoun, and which also make it different from similar pro-forms in the language. I argue further that this special feature makes it the default choice in instances where a repair is needed for what otherwise would be a violation of the NSG constraint.

4.1 Three types of *ɛ-* in Akan

Based on the anaphoric properties of the *ɛ-* pro-forms that we have seen so far, I distinguish between three kinds of *ɛ-* pro-forms in Asante-Twi: the AGREEING form(s), the DEFAULT form, and the EXPLETIVE form. I summarize the properties of these homophonous

forms in Table 2. In order to account for their distinctions, I assume that every personal pronoun in Akan needs to specify a value for at least one of the following properties **in order to be anaphoric or referential**:

  i. PERSON: Participant (i.e. 1st/2nd person) or Non-participant (i.e. 3rd person).
  ii. NUMBER: Singular (Sg) or Plural (Pl)
  iii. ANIMACY: Animate or Inanimate (Inan)

Table 2: Features of ε- pro-forms

	PRO-FORM	PERSON	NUMBER	ANIMACY
AGREEING	ε	NPart	Sg	Inan
AGREEING	ε	NPart	Pl	Inan
DEFAULT	**ε**	**NPart**	-	-
EXPLETIVE	ε	-	-	-

Given Table 2, all agreeing ε- forms have specified values for all the pronominal properties. I treat these as one type of ε-, i.e. they all fully agree with their antecedent NPs, as in (22), and (4b).

(22) [N-kŕataá nó]$_i$ áa ε_i-dá   pónó nó só nó yɛ́ fɛ.
      PL-paper DEF   REL 3PL.NOM.INAN-lie table DEF top CD be nice

      'The papers that are on the table are nice.'

The default ε- differs from the agreeing pro-forms in the sense that it has no specified values for number and animacy. But since it is specified for at least third person, it is still referential, as in (23a). We know this because when a native speaker is presented with (23b), independent of (23a), the ε- in (23b) helps the speaker to select the pronoun that matches the coindexed gap in terms of person features, as (23c) shows. Thus, ε- here simply refers to third person. In this regard, the use of ε- in (23a) is comparable to the use of *their* in (24a). At least some native speakers of English use *their* in such variable binding contexts (although some speakers would use (24b) instead). Here, although *their* and *your driver* do not totally agree, there still exists a binding relationship between them. I emphasize the striking similarity between the person features of both *their* and the default ε- in Akan; both are third person.

(23) a. [Amma$_i$ (ne Yaa)]$_i$ na ε_i-wáré-e   Kofi.
         A.   CONJ Y.   FOC DFLT-marry-PST K.

         'AMA (and YAA) married Kofi.'
      b. $-_i$ na ε_i-wáré-e   Kofi.
           FOC DFLT-marry-PST K.

         '— married Kofi.'

c. **Ɔno**$_i$/**Wɔ́n**$_i$/*wó$_i$/*mé$_i$ na ɛ$_i$-wáré-e   Kofi.
3sg/pl/2sg/1sg   foc dflt-marry-pst K.

's/he/they married Kofi.'

(24) a. If there are any changes to the service, [your driver]$_i$ will do **their**$_i$ best to inform you.[4]

b. If there are any changes to the service, [your driver]$_i$ will do **his**$_i$ or **her**$_i$ best to inform you.

The expletive ɛ- in Table 2 has no specified feature for any of the three properties. It is only similar to the other two in terms of form. This may be due to the fact that all three ɛ- forms get the same case feature, i.e. nominative, by virtue of where in the syntax they are permitted, i.e. subject position. This is the only feature that it shares with the agreeing ɛ-, and the default ɛ. But certainly this property is not an inherent property of this pronoun.

The expletive ɛ- is like expletive pronouns in other languages, e.g. (25) where the subject pronoun is absolutely non-referential. As (26) shows, similar constructions exist in Akan. Here, ɛ- has a purely formal function, and neither agrees nor refers to any NP.

(25) It is raining.

(26) Expletive ɛ-
Ɛ-wɔ   sɛ́   obíáá$_i$ túmí kyéréɛ n$_i$-ádwén.
expl-be that everyone able show 3sg.poss-mind

'It ought to be the case that everybody is able to express their opinion.'

Based on the above discussions, I conclude that Asante-Twi has at least three kinds of ɛ- pro-forms: the agreeing one, the default one, and the expletive one. This distinction is based on their pronominal properties. Contexts that require a less specific but referential ɛ- favor the default form over the expletive form.

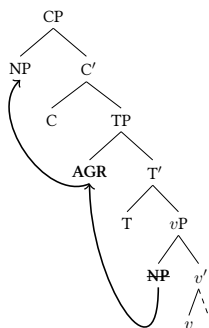## 4.2 On the distribution of agreeing and default ɛ-

I showed early on that Akan uses resumptive pronouns as a repair strategy for the NSG. Given that RPs are the result of movement (see §2.3), I deduce that all instances of RPs in subject positions in Akan are derived via movement, as represented in (27). I propose further in this section that apart from this RP strategy, Akan also sometimes inserts the default ɛ- as a repair mechanism for a potential violation of the NSG constraint. This latter strategy is, however, relativized to only the matrix subject position in the syntactic structure. I propose that the matrix subject position is a privileged position in Akan. This is supported by the fact that there is a back-up strategy to repair any instance of an NSG violation in case resumption fails.

---

[4] This is an extract from an audio clip played on some National Express coaches in Britain.

We recall from §2.2 that a language may also skip a subject position altogether in order to avoid an extraction that would result in a violation of the NSG constraint. Given the assumption that the use of a RP in the extraction site results from movement only, I deduce that any instance of the non-agreeing pronoun in the extraction site cannot be attributed to movement, because movement yields only (agreeing) resumptive pronouns. I claim therefore that such non-agreeing cases involve the skipping of the subject position. In order to avoid a violation of the NSG constraint, the default ε- pronoun is inserted. This strategy is available only for the matrix subjects. Crucially, this process is ordered, as shown in (28).

(27)   Movement

(28)   Deriving default ε-
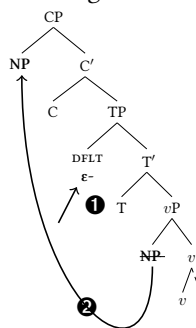
5 Summary and conclusion

The distribution of (agreeing) resumptive pronouns and the (non-agreeing) default pronoun in Akan can be summarized as presented in Table 3.

Table 3: Distribution of RPs in Akan

Strategy	Matrix Subj	Embedded Subj	Obj
Gap	✗	✗	(✓)
RP	(✓)	✓	✓
DFLT insertion	(✓)	✗	✗

With regard to the two main issues that this paper set out to address, I have argued that, based on the relevant pronominal properties, three types of ε- pro-forms can be distinguished in Asante-Twi: the agreeing, the default, and the expletive forms. I have claimed that the non-agreeing pronoun is analyzable as a default pronoun that is inserted in the matrix subject position whenever resumption fails. Thus Akan, like Yoruba (Ades-

ola 2010), has two strategies for repairing a potential violation of the NO-SUBJECT-GAP-like constraint; it uses either an RP, or a default pronoun. In §4, I suggested a potential alternative explanation for the agreement mismatches with regard to the distribution of these agreeing and non-agreeing subject pronouns in Akan. Further discussion of this will be insightful for how SPELL-OUT works in Generative syntax. But this is an issue that future research on this phenomenon will attempt to address.

Abbreviations

1,2,3	first, second, third person	(IN)ANIM	(in)animate
ACC	accusative	NOM	nominative
CD	clausal determiner	NPart	non-participant
COMP	complementizer	Obj	object
CONJ	conjunction	PERF	perfective
DEF	definite	PL	plural
DEM	demonstrative	POSS	possessive
DFLT	default	PST	past
FOC	focus	REL	relativizer
FUT	future	SG	singular
HAB	habitual	Subj	subject

Acknowledgments

References

Adesola, Oluyese. 2010. The non-agreeing subject resumptive pronoun in Yoruba. In Enoch O. Aboh & James Essegbey (eds.), *Topics in Kwa syntax* (Studies in Natural Language and Linguistic Theory 78), 65–89. London: Springer.

Chinebuah, Isaac K. 1976. Pronominalization in some Kwa languages. *Journal of West African Languages* 11(1-2). 57–72.

Chomsky, Noam. 1977. On wh-movement. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Georgi, Doreen. 2014. *Opaque interactions of Merge and Agree. On the nature and order of elementary operations*. Leipzig: Universität Leipzig. (Doctoral dissertation.)

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 3. 63–99.

Klein, Timo. 2014. Derving a typology of resumption. In Anke Assmann, Sebastian Bank, Doreen Georgi, Timo Klein, Philipp Weisser & Eva Zimmermann (eds.), *Topics at InFL* (Linguistische Arbeits Berichte 92), 233–249. Leipzig: Institüt für Linguistik, Universität Leipzig.

Korsah, Sampson & Andrew Murphy. 2015. High tone insertion in Akan as a reflex of successive-cyclic movement. (Paper presented at the Workshop on Morphosyntactic Triggers of Tone. University of Leipzig.)

Larson, Martha. 2005. *The empty object construction and related phenomena*. Ithaca, NY: Cornell University. (Doctoral dissertation.)

Marfo, Charles. 2005. Akan focus and topic constructions and the phonology-syntax interface. *Journal of West African Languages* 32(1). 45–59.

Marfo, Charles & Adams Bodomo. 2005. Information structure in Akan question-word fronting and focus constructions. *Studies in African Linguistics* 34(2). 179–208.

McCloskey, James. 2011. Resumptive pronouns, ā-binding, and levels of representation in Irish. In Alian Rouveret (ed.), *Resumptive pronouns at the interfaces* (Language Faculty and Beyond: Internal and external Variation in Linguistics 5), 65–119. Amsterdam: John Benjamins.

McCracken, Chelsea. 2013. Relative pronouns in Asante-Twi. *Rice Working Papers in Linguistics* 4. 1–28.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.

Polinsky, Maria & Eric Potsdam. 2001. Long-Distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19. 583–646.

Rizzi, Luigi & Ur Shlonsky. 2007. Strategies of subject extraction. In Hans Martin Gärtner & Uli Sauerland (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the view from Syntax-Semantics*, 115–160. Berlin: Mouton de Gruyter.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation.)

Saah, Kofi K. 1992. Null object construction in Akan. *MIT Working Papers in Linguistics* 17. 219–244.

Saah, Kofi K. 1994. *Studies in Akan syntax, acquisition and sentence processing*. Ottawa: University of Ottawa. (Doctoral dissertation.)

Saah, Kofi K. 2010. Relative clauses in Akan. In Enoch O. Aboh & James Essegbey (eds.), *Topics in Kwa syntax* (Studies in Natural Languages and Linguistic Theory 78), 91–107. London: Springer.

Salzmann, Martin. 2011. Silent resumptives in Zurich German possessor relativization. In Melani Wratil & Peter Gallmann (eds.), *Null pronouns*, 141–221. Berlin: De Gruyter Mouton.

Shlonsky, Ur. 1992. Resumptive pronouns as a last resort. *Linguistic Inquiry* 23. 443–468.

Chapter 7

Valency changing processes in Akan

E. Kweku Osam
University of Ghana

Valency has been considered as both a semantic and syntactic notion. Semantically it is used to refer to the participants in an event; as a syntactic notion it is used to indicate the number of arguments in a construction. In Akan, a Kwa language spoken in Ghana, we can identify various transitivity classes of verbs: verbs that are strictly intransitive; those that are strictly transitive; and verbs that are used ditransitively. Apart from these, there are verbs that can be used both intransitively and transitively. Critical to the study of the notion of valency in Akan is the fact that there is clear evidence for grammatical relations in the language. As is the case in many languages, Akan possesses morphosyntactic means through which the valency of verbs can be adjusted. The application of these morphosyntactic processes reduces or increases the valency of verbs. This paper examines these processes in Akan. The critical valency-reducing processes in Akan are reflexivization, reciprocals, anticausative/inchoative constructions, impersonal constructions, object omission constructions, and unspecified object constructions. Valency-increasing processes include causativization and agentivization through serialization.

1 Introduction

The linguistics literature is replete with studies on valency – those that are theoretical, cross-linguistic, and others that focus on the study of valency in specific languages. The interest in understanding the notion of valency has resulted in a number of volumes dedicated to the subject. Notable among these are Comrie & Polinsky (1993), Dixon & Aikhenvald (2000a), Malchukov & Comrie (In press). The papers in Malchukov & Comrie (In press) came out of the Leipzig Valency Classes Project and the Conference on Valency Classes in the World's Languages held in 2011. Worthy of mention is also Nichols, Peterson & Barnes (2004) that put forward the idea of the basic valency orientation of languages. As in the case of the Leipzig Project, a number of scholars adopting the methodology of Nichols and associates, have examined the basic valency orientation of various languages (see, for example, Narogg 2009; Plank & Lahiri 2009; van Gelderen 2011; Luraghi 2012).

The goal of this paper is to examine the valency of Akan verbs and explore the morphosyntactic processes that apply to modify the valency of verbs. It is expected that

the paper will help expand further our understanding of the behavior of verbs in Akan, building on what has been reported in Osam (2008b).

This paper is based on the Akan Verbs Database project which was implemented in the Department of Linguistics, University of Ghana, between 2009 and 2011, and for which I was the Principal Investigator. The project was funded by the University of Ghana Research Fund. By the end of the project, a database of over 3,500 verb stems and other verb forms had been created.

The data come from various sources, including published narratives, translated works and, in some cases, constructed examples based on my knowledge as a native speaker.

The paper is structured as follows: §2 provides an overview of the verbal morphology, focusing on the tense-aspects distinctions of the language and clause structure/grammatical relations. This section also covers some aspects of serial verb constructions in the language since they have a bearing on valency adjustment in the language. §3 deals with Akan verb classes. In §4 the morphosyntactic processes that reduce the valency of verbs are discussed; §5 and §6 focus on the processes that increase valency. §7 is the summary and conclusion.

The label Akan is used to refer to a group of closely related dialects spoken in Ghana, and partially in the south eastern parts of Cote d´Ivoire. It belongs to the Kwa sub-family of Niger-Congo. The name also refers to the people who speak the language. The dialects of Akan include: Agona, Ahafo, Akuapem, Akwamu, Akyem, Asante, Assin, Bono, Denkyira, Fante, Kwahu, and Wassa. Generally, except Fante, all the other dialects tend to be classified as Twi in terms of Akan dialectology. Current speaker population is estimated at nearly ten million.

2 Some morphosyntactic features of Akan

In this section I provide an overview of tense/aspect and clause structure in the language. I should point out that Akan is a two tone language and one of the outstanding phonological features of the language is the presence of Tongue Root Harmony in all the dialects and rounding harmony in the Fante dialect. The discussion throughout the paper will draw examples from the Fante (Fa) dialect and in some cases from the Asante (As) and Akuapem (Ak) dialects as well. Examples drawn from Fante will generally not be identified as such; those from Asante and Akuapem will be indicated as As and Ak, respectively, in parentheses generally at the end of a translation.

2.1 Verbal affixes

In Akan verbal constructions are generally formed with the verb and its person, tense, aspect, mood, and polarity affixes. There are also verbal prefixes to mark motion towards or away from a deictic centre (Table 1).

The pre-verbal affixes in Akan include the tense/aspect markers. As I have argued elsewhere (Osam 1994a; 2008a), Akan can be said to have a Future Tense and the follow-

Table 1: Subject Prefixes

1SG	*me-*
2SG	*wo-, i-* (Fa)
3SG	*ɔ-*
3SG (inanimate)	*ɛ-* (in Ak, As and some Fa subdialects)
1PL	*yɛ-*
2PL	*mo-/wɔ-* (Fa)
3PL	*wɔ-/yɛ-*

ing aspectual forms: Completive (COMPL), Perfect (PERF), Progressive (PROG), Habitual (HAB), Continuative (CONT), and Consecutive (CONSEC).

The Future tense is coded by the prefix *bɛ-*. The realization of the vowel is determined by vowel harmony, especially in the Fante dialect (see Dolphyne 1988).

(1) a. Íyí bɛ́-mà kɔ́ŕ-yɛ́ à-bà hóḿ ńtámú.
   this FUT-make one-be CONS-come 2PL.POSS middle

   'This will bring unity among you.' (Krampah 1970: 79)

   b. Dàákyé yè-bé-nyá bì á-ká á-kyèrɛ́ à-fófóró.
   future 1PL.SUBJ-FUT-get some CONS-say CONS-show PL-new

   'In future we will have something to tell others.' (Adi 1973: 34; Ak)

What I consider the Completive is what in the general literature on Akan is referred to as the Past tense. However, I have shown in various places (Osam 1994a; 2004; 2008b) that this form is more of an aspect than tense. The Completive is a suffix in the affirmative; but a prefix in the negative.

(2) a. Né ǹ-sèẁ-nóḿ bèènú nyìná yɛ́-ɛ̀ èdzìbàń.
   3SG.POSS PL-in.law-COL two all make-COMPL food

   'Each of his two in-laws cooked.' (Krampah 1970: 57)

   b. Né ǹ-sèẁ-nóḿ bèènú nyìná à-ǹ-yɛ́ èdzìbàń.
   3SG.POSS PL-in.law-COL two all COMPL-NEG-make food

   'His two in-laws did not cook.'

The Perfect aspect is realized by the prefix *a-*, and generally agrees with the vowel of the verb root in ATR harmony.

(3) a. Bànyíń nó á-tɔ̀ àsàásé.
   man DEF PERF-buy land

   'The man has bought a piece of land.'

b. Ǹ-bòfrá nó á-bà   fíé.
   PL-child DEF PERF-come home

   'The children have come home.'

The Progressive is a prefix, *re-*.

(4) a. Mà`àmé nó rè-hwè-hwɛ́   né   bá   nó.
        woman DEF PROG-REDU-look 3SG.POSS child DEF

        'The woman is looking for her child.'

     b. Àbòfrá nó rù-tù-tú   ǹ-wúrá nó.
        child   DEF PROG-REDU-uproot PL-weed DEF

        'The child is removing the weeds.'

Akan also has a Habitual aspect which is realized by tone (see Dolphyne 1988 for more discussion on this).

(5) a. Òkùàfó nó kɔ̀   hàbáń mù ànɔ̀pá   bíárá.
        farmer DEF go;HAB farm   in morning every

        'The farmer goes to the farm every morning.'

     b. Akosua tɔ̀ǹ   ǹdzɛ́ḿbá wɔ̀ gúá-mú.
        Akosua sell;HAB things   in market-in

        'Akosua sells in the market.'

Two of the aspects, the Continuative and Consecutive, are treated as derived aspects. The Continuative is used for stative verbs where the Progressive is used for dynamic verbs. The Consecutive aspect applies to non-initial verbs in a serial construction in which the initial verb is in either the Progressive aspect or the Future tense. Generally, the Consecutive is realized as a low tone *à-*.

(6) a. Àbòfrá nó tsè   dùá nó ásé.
        child   DEF sit/CONT tree DEF under

        'The child is sitting under the tree.'

     b. Kwesi gyìnà   fíé   nó ényíḿ.
        Kwesi stand/CONT house DEF front

        'Kwesi is standing in front of the house.'

(7) a. Bànyíń nó bɛ́-sáẃ   à-kyèrɛ́   dɔ́ḿ nó.
        man   DEF FUT-dance CONS-show crowd DEF

        'The man will dance for the crowd.'

     b. Kofi rè-náńtsèẁ à-kɔ̀   skúùl.
        Kofi PROG-walk CONS-go school

        'Kofi is walking to school.'

There are verbal constructions in Akan which sometimes have two prefixes, *bɛ-* and *kɔ-*. These prefixes arise from the verbs for 'come' and 'go', respectively. I have referred to these as Motional prefixes (Osam 2002). They reflect the movement towards or away from a deictic centre where an event takes place. Movement towards the deictic centre is marked by the 'come' verb; and away from the centre is marked by the 'go' verb.

(8) a. Ama bɔ̀-sɔ́-ɔ̀   gyá wɔ̀ dáń   nó ékyìr̀.
         Ama come-light-COMPL fire at building DEF back

         'Ama came and lit a fire behind the building.'

     b. Esi kɛ̀-fá-à   èkùtú nó bá-à   fíé.
         Esi go-take-COMPL orange DEF come-COMPL home

         'Esi went and brought the oranges home.'

Even though it is too early to say that Akan has a prospective aspect, the language has the means to express prospective meaning. This is done through a combination of the Progressive and the 'come' Motional prefix.

(9) a. Ɔ̀màǹpànyíń rè-bɔ́-sɔ́ŕ.
         president   PROG-come-stand

         'The president is about to rise.'

     b. Hyɛ́ń   nó rè-bé-gyíná.
         vehicle DEF PROG-come-stop

         'The vehicle is about to stop.'

The language makes a two way distinction in terms of mood – the indicative and the Imperative. The Imperative has two manifestations; what we may call Imperative proper and the Optative. Whereas the Imperative has no segmental representation, the Optative is realized through the use of a homorganic nasal with a high tone, *ń-*.

(10) a. Gyàè!
         stop/IMP

         'Stop it.'

     b. Yɛ́-ń-kɔ́.
         1PL-OPT-go

         'Let's go.'

The language uses a homorganic nasal prefix to express negation.

(11) a. Mì-ǹ-nyím̀   èdwúmá â ɔ́-yɛ́.
         1SG.SUBJ-NEG-know work   REL 3SG.SUBJ-do

         'I don't know what work she does.'

b. Màámé nó à-ǹ-ká àséḿ nó à-ǹ-kyèré né ḿ-bá
woman DEF COMPL-NEG-say matter DEF COMPL-NEG-show 3SG.POSS PL-child

nó.
DEF

'The woman didn't tell her children about the case.'

2.2 Verbal reduplication

Many verbs in Akan are subject to the morphological process of reduplication. Generally, verbal reduplication is required where the event is iterative and/or where either the Subject or Direct Object NPs or both are plural.

(12) a. Áféí ǹ-nàm̀fò yí kyìǹ-kyíń-ń à-dɔ́fó kàkrá â wó-níḿ
now PL-friend these REDU-roam-COMPL PL-lover few REL 3PL.SUBJ-know

wɔ́ń nó sò krá-krà-à wɔ̀ǹ.
3PL.OBJ DCM on REDU-bid.farewell-COMPL 3PL.OBJ

'Now these two friends visited some of their friends to bid them farewell.'
(Adi 1973: 41; Ak)

b. Né nà nà n-égyà é-wú-wú.
3SG.POSS mother and 3SG.POSS-father PERF-REDU-die

'Both his mother and father are dead.' (Martin 1936: 24)

2.3 Akan clause structure

The syntax of Akan distinguishes grammatical relations, as has been discussed in various studies (Osam 1994a; 1996; 1997; 2000; 2004). The language has Subject and Direct Object with Nominative Accusative marking that is realized through word order. Word Order in the language is strictly SVO, with very little variation.

2.3.1 Transitive constructions

The prototypical Akan transitive clause has A (as Subject) and O (as Direct Object) as core arguments. The A argument precedes the O; and each can be realized as full NP or as a pronominal element.

(13) a. Akosua kyè-è àkókɔ́ nó.
Akosua catch-COMPL chicken DEF

'Akosua caught the chicken.' (As)

b. Ò-kùà-fó nó á-dáá-dàà ɔ̀-bɔ̀-fó nó.
SG-farm-IDM DEF PERF-REDU-deceive SG-hit-IDM DEF

'The farmer has deceived the hunter.'

If the O argument is inanimate, it is not represented by a pronominal form unless some other clausal element comes after it.

(14) a. Ò-wíà-à   náḿ nó.
         3sg.subj-steal-compl fish def
         'He stole the fish.'

      b. Ò-wíà-è   Ø.
         3sg.subj-steal-compl
         'He stole it.'

      c. Ò-wíà-à   nò   ńdédà.
         3sg.subj-steal- compl 3sg.obj yesterday
         'He stole it yesterday.'

A transitive clause could also have optional Oblique elements expressed by locational (both spatial and temporal) phrases. It could also be a postpositional phrase. Generally, obliques would occur in clause-initial or clause-final positions.

(15) a. Màámé nó sí-ì   dáń ḿfédà.
         woman def build-compl house last.year
         'The woman built a house last year.'

      b. Ḿfédà   màámé nó síì   dáń.
         last.year woman def build-compl house
         'Last year the woman built a house.'

2.3.2 Intransitive constructions

The single argument of an intransitive clause, the S argument, always precedes the predicate.

(16) a. Èdzìbàń nó á-bèǹ.
         food   def perf-be.cooked
         'The food is cooked.'

      b. Ò-sɔ́-fó   nó wú-ù   ńdédà.
         sg-pray-idm def die-compl yesterday
         'The priest died yesterday.'

2.3.3 Ditransitive constructions

Akan has ditransitive constructions in which there are three core arguments, agent, benefactive and theme. In ditransitive constructions, the NP in the immediate post-verbal position is grammatically the Direct Object and semantically the benefactive. The entity that is transferred, the theme, occurs after the benefactive NP. I have referred to this as the *Asymmetrical Object* (Osam 2000).

(17) a. Màámé nó má-à   m̀-bòfrá nó èdzìbáń.
         woman def give-compl pl-child def food
         'The woman gave the children food.'

b. Pàpá nó kyέ-ὲ   hɔ̀ǹ   sìká.
man DEF gift-COMPL 3PL.OBJ money

'The man gave them money.'

2.3.4 Serial verb constructions

One feature of Akan syntax crucial to valency discussions is serial verb constructions (SVC). Akan serialization has been studied extensively (see, for example, Lord 1973; Schachter 1974; Essilfie 1977; Forson 1990; Osam 1994a,b; 1997; 2004; 2014; Agyeman 2002; Hellan, Beermann & Andenes 2003; Kambon 2012). Without going into the details of Akan SVCs, it is important to identify some salient features.

Subject marking: the subject may be a NP or a pronominal form that occurs on only the initial verb.

(18) a. Yaakwa n-ábákáŋ   gyínà-ὲ   yέ-ὲ   kyèámé   bɔ́-ɔ̀
Yaakwa 3SG.POSS-first.born stand-COMPL be-COMPL spokesperson tell-COMPL
fíé   àmàǹdzὲέ kyérὲ-ὲ   à-hɔ́hó   nó.
home news   show-COMPL PL-visitor DEF

'Yaakwa's eldest son stood as the spokesperson and briefed the visitors.' (Krampah 1970: 83)

b. Wɔ̀-twé-ὲ   nó gyìná-à   ǹkyéŋ́.
3PL.SUBJ-pull-COMPL DEF stand-COMPL aside

'They pulled him aside.' (Adi 1973: 31; Ak)

Generally, there is uniformity in tense/aspect coding, as shown in the examples above. However, mixed tense/aspect is possible in some SVCs.

Negation Marking: across all dialects of Akan, each verb in the series takes the negation prefix when the sentence is negated, as in (19).

(19) Màààmé nó à-ǹ-tɔ́   èdzìbàǹ à-m̀-má   né   ḿ-bá   nó.
woman DEF COMPL-NEG-buy food   COMPL-NEG-give 3.SG.POSS PL-child DEF

'The woman did not buy food for her children.'

In some serial constructions, the initial verb is *de/dze*. This is a form with reduced verbal properties. The *de* is used in the Twi dialects and *dze* is used in Fante.

(20) a. Wɔ̀-dè   àsέŋ́ nó kɔ́-ɔ̀   àhèm̀fíé.
3PL.SUBJ-take matter DEF go-COMPL palace

'They took the case to the palace.' (Ak)

b. Ɔ̀-dzè   n-áhɔ̀ɔ̀dzéŋ́   nyìná yέ-ὲ   èdwúmá nyá-à   sìká.
3SG.SUBJ-take 3SG.POSS-strength all   do-COMPL work   get-COMPL money

'She worked very hard and made money.'

The de serialization is also relevant in the expression of agentive arguments in the case of verbs of spatial location. This will be further discussed in §5.2.

3 Akan verb classes

Based on their argument structure, various transitivity classes of verbs can be identified in Akan. Some are strictly intransitive (§3.1), some are strictly transitive (§3.2), some are ditransitive (§3.3), and some have varying expressions of arguments. Verbs of the last type are addressed in §4 and §5.

3.1 Strictly intransitive/monovalent

Monovalent verbs occur with a single argument; the S argument only (Table 2).

Table 2: Monovalent verbs

béń	'be cooked'	*bèr̀*	'be ripe, be fair in complexion'
dwó	'cool'	*fé*	'become soft/smooth'
hóń	'swell'	*hwéḿ*	'blow one's nose'
hwèǹtsì	'sneeze'	*péŕ*	'struggle, roll around during sleep'
prɔ̀	'rot'	*húŕ*	'boil'

(21) a. Èdzìbàń nó   á-bèǹ.
          food   DEF PERF-cook
          'The food is cooked.'

     b. Àǹkàá   nó á-prɔ́.
          oranges DEF PERF-rot
          'The oranges are rotten.'

3.2 Strictly transitive/bivalent

There are verbs that require two arguments, A and O (Table 3).

Table 3: Bivalent verbs

bòr̀	'beat'	*dààdàà*	'deceive'	*hyìrà*	'bless'
ká	'bite'	*kú*	'kill'	*kyé*	'catch'
nyéń	'rear'	*pàà*	'curse'	*sànè*	'infect'
sìè	'bury'	*táń*	'hate'		

(22) a. Ɔ̀-bɔ̀-fóɔ́   nó kú-ù   ɔ̀sónó.
          SG-hit-IDM DEF kill-COMPL elephant
          'The hunter killed an elephant.'

b. Abam síè-è né màámé.
Abam bury-COMPL 3SG.POSS mother

'Abam buried his mother.'

3.3 Ditransitive/trivalent

There are verbs that are ditransitive or trivalent. These require three core arguments (Table 4).

Table 4: Trivalent verbs

kyέ	'to gift, give as a gift'	*gyé*	'charge'
má	'give'	*kyèrè*	'teach, show'

(23) a. Ńtsí ɔ̀-má-à nò à-hέń èbìásá.
so   3SG.SUBJ-give-COMPL 3SG.OBJ PL-boat three

'So she gave him three boats.' (Martin 1936: 8)

b. Esi kyέ-ὲ pàpá nó sìká.
Esi gift-COMPL man DEF money

'Esi gave the man money.'

4 Decreasing valence

There are morphosyntactic processes that reduce verb valence. In various languages, morphological derivations are utilized in reducing the valence of a verb. But as has been pointed out in the literature, verbs can also manifest a change in the valence structure without the application of any morphological processes: "Alternations in a verb's valency pattern are not necessarily the result of a morphological derivational process. Verbs or whole classes of verbs may have alternate valency patterns without any change in their formal makeup …" (Haspelmath & Müller-Bardey 2004: 1131).

4.1 Reflexivization

Reflexivization in Akan is marked by the use of a possessive (POSS) pronoun and the morpheme *ho* 'self'.

(24) a. Kofi é-kù Yaw.
Kofi PERF-kill Yaw

'Kofi has killed Yaw.'

b. Kofi é-kù   nó   hó.
   Kofi PERF-kill 3SG.POSS self

   'Kofi has killed himself.'

In reflexivization, the notion of coreference is crucial. Kemmer (1993: 44) notes that "Coreference … means that two participants in a single event frame designate the same entity in the described situation." It requires that the A and O arguments have the same referent. Reflexive constructions in Akan involving bivalent verbs have the coreferential O argument replaced by the POSS+*ho*. The reduction in the valency of the verb lies in the fact that there is no semantic differentiation between the A and O arguments.

Apart from the direct reflexive, certain verbs, specifically grooming, or body care actions (Kemmer 1993) are used reflexively and, consequently, demonstrate (semantic) valency decreasing properties. Examples of such verbs in Akan are: *pepa* 'wipe', *sera* 'smear/use lotion or oil on the body', *twutwuw* 'wash (as with a sponge/washcloth)', *siesie* 'dress up'. The actions coded by these verbs can be carried out by an AGENT on a PATIENT entity or the AGENT can carry it out on themselves.

(25) a. Aba twùtwúẁ-ẁ   Ekua nó   hó.
         Aba wash.wash-COMPL Ekua 3SG.POSS self

         'Aba washed Ekua.'

      b. Aba twùtwúẁ-ẁ   nó   hó.
         Aba wash.wash-COMPL 3SG.POSS self

         'Aba washed someone/herself.'

(26) a. Efua sérà-à   Kwesi nó   hó.
         Efua smear-COMPL Kwesi 3SG.POSS self

         'Efua used lotion on Kwesi.'

      b. Efua sérà-à   nó   hó.
         Efua smear-COMPL 3SG.POSS self

         'Efua used lotion on someone/herself.'

In (25a) and (26a), the entities in subject position, as the AGENTS, carry out the activities on certain other individuals, the entities in the PATIENT role. In (25b) and (26b), the replacement of the full NP with the reflexive pronoun in the post-verbal position creates ambiguities. In each sentence, the referent of the reflexive pronoun could be the entity in the subject position, the AGENT; or it could be an individual already mentioned in the context of the discourse; that is Ekua in (25a) and Kwesi in (26a).

Where the referents of the reflexive pronouns in (25b) and (26b) are the AGENTS in the subject positions in the two sentences, we can argue for a reduction in the valency of the verbs *sera* 'smear/use lotion or oil on the body' and *twutwuw* 'wash (as with a sponge/washcloth)', on the basis of coreferentiality and non-individuation of the participants involved in the situation.

In addition to the reduction in the valency of the verb *sera* through reflexivization, its valency can also be reduced through its use in an intransitive construction as (27b) shows.

(27) a. Efua sérà-à   nó   hó.
          Efua smear-COMPL 3SG.POSS self

      'Efua used lotion on herself.'

  b. Efua sérà-è.
      Efua smear-COMPL

      'Efua used lotion (on herself).'

There is another feature of the verb *sera* pertaining to valency adjustment that will be discussed in §4.5.2

4.2 Reciprocals

The reciprocal in Akan is formed in ways similar to the reflexive. It also uses the morpheme *ho*. But unlike the reflexive, the possessive pronoun that combines with *ho* is in the plural (28). Similar to the reflexive, the reciprocal creates reduced valency due to coreferentiality and lack of individuation.

(28) a. Hwɛ́, émí   nà wó-égyà   yè-húń-ǹ   hɛ́ń   hó béŕ â
          look 1SG.EMPH and 2SG.POSS-father 1PL.SUBJ-see-COMPL 1PL.POSS self time REL
          mí-dzí-ì   m̀-fé   èdùònú ànáń nó.
          1SG.SUBJ-eat-COMPL PL-year twenty four DCM

      'Your father and I got to know each other when I was twenty-four years.' (Martin 1936: 9)

  b. M̀-bèráńtsé nà ǹ-kàtàásíá nó nyé   Araba Akɔm hám-èè
      PL-man   and PL-woman DEF accompany Araba Akɔm quarrel-COMPL
      yè-yáẁ-ẁ   hɔ̀ń   hó mà ɔ̀-yɛ́-ɛ̀   àsɛ́ḿ wɔ̀ skúùl.
      REDU-insult-COMPL 3PL.POSS self that 3SG.SUBJ-make-COMPL issue in school

      'The young men and young women quarreled with Araba Akɔm and insulted each other such that it became an issue in the school.' (Martin 1936: 21)

4.3 Anticausative/Inchoative

The anticausative (or decausative, inchoative, spontaneous, pseudopassive, being the various ways in which this type of construction has been labelled) works by removing the AGENT argument in the construction. In his characterization of the causative/inchoative alternation, Haspelmath states that:

> An inchoative/causative verb pair is defined semantically: it is a pair of verbs which express the same basic situation (generally a change in state, more rarely a going-on) and differ only in that the causative verb meaning includes an Agent participant

who causes the situation, whereas the inchoative verb meaning excludes a causing Agent and presents the situation as occurring spontaneously … Inchoative verbs are generally intransitive and causative verbs are transitive … (Haspelmath 1993: 90)

As pointed out by Haspelmath & Müller-Bardey (2004: 1132):

In many languages there is a strong requirement for all sentences to have subjects. When in such languages a valency-changing category removes the agent argument from the subject position, the patient argument must take up the subject position instead.

Akan is a typical example of such languages. When a sentence has a single argument, that argument is always in the subject position. In Akan anticausatives the THEME argument is the subject of the sentence.

Unlike some other languages where there is a derivational process to indicate either the demotion of the AGENT argument or its introduction, in Akan there is no change in the morphology of the verb stem to reflect the process of anticausative. This puts Akan into what Haspelmath (1993: 91) describes as non-directed alternation, that is, where "… neither the inchoative nor the causative verb is derived from the other."

The verbs in the language that can be used in the anticausative construction include: *bɔ* 'break', *bu* 'break', *bue* 'open', *butu* 'overturn', *dum* 'put out', *hyew* 'burn', *koa* 'bend', *kyea* 'bend', *nane* 'melt', *sɛɛ* 'destroy', *te* 'tear, pluck', *tuei* 'puncture', *woso* 'shake, vibrate'.

Example (29a) and (30a) are causative and (29b) and (30b) are anticausative/inchoative.

(29)   a.   nà   wɔ̀-bò-bɔ́-ɔ̀   ǹ-kùràbá nó   â   ɔ́-dzé-dzè
          and 3PL.SUBJ-REDU-break-COMPL PL-jar   DEF REL 3PL.SUBJ-REDU-hold
          hɔ́ń   ńsá-mú nó.
          3PL.POSS hand-in DCM
          '… and they broke the jars that were in their hands.' (Judges 7:19 Fante Bible; Bible Society of Ghana 1974)

   b.   ǹ-kùràbá nó   â   ɔ́-dzé-dzè   hɔ́ń   ńsá-mú nó
       PL-jar   DEF REL 3PL.SUBJ-REDU-hold 3PL.POSS hand   DCM
       bò-bɔ́-è.
       REDU-break-COMPL
       '… and the jars in their hands broke.'

(30)   a.   nà   Moses hyéẁ-ẁ   tsíŕ nó.
          and Moses burn-COMPL head DEF
          '… and Moses burned the head.' (Leviticus 8:20 Fante Bible; Bible Society of Ghana 1974)

   b.   nà   tsíŕ nó hyéẁ-èè.
       and head DEF burn-COMPL
       '. . . and the head burned.'

4.4 Impersonal constructions

Another means of valency decrease is through impersonal constructions. The notion of impersonal construction adopted here follows Siewierska (2008; 2011) and Malchukov & Ogawa (2011). These are constructions that do not have a referential subject. Malchukov & Ogawa (2011), following Keenan (1976), argue that the subject in an impersonal construction deviates from the prototype subject. In Keenan's (1976) approach, the canonical subject is expected to have the following properties (Malchukov & Ogawa 2011: 23):

- a referential argument
- a definite NP
- topical
- animate
- agentive

Based on cross-linguistic studies, various coding strategies of impersonal constructions have been isolated. One of these is the pronominal impersonal (Siewierska 2011). In some languages this involves the use of a regular personal pronoun as the subject of the construction. Akan does this by using the regular 3 person plural subject pronoun, as illustrated in (31).

(31) a. Wɔ̀-á-kyèr̀   èwì-fó nó.
3PL.SUBJ-PERF-catch thief-PL DEF

'They have arrested the thieves.'/'The thieves have been arrested.'

  b. Wɔ̀-á-tò   ésíkyìré nó   bó   mú.
3PL.SUBJ-PERF-raise sugar   3SG.POSS price in

'They have increased the price of sugar.'/'The price of sugar has been increased.'

The subject pronouns in (31) are non-referential and non-individuated (Hopper & Thompson 1980). Example (32) below is taken from the Apostles' Creed of the Christian faith.

(32) Wɔ̀-bɔ́-ɔ̀   Nò   m̀bèàmúdúá mù, Ò-wú-ì,   wò-síé-è
3PL.SUBJ-hit-COMPL 3SG.OBJ cross   in   3SG.SUBJ-die 3PL.SUBJ-bury-COMPL
Nò.
3SG.OBJ

'He was crucified, dead and buried.' (Source: Christian Asɔr Ndwom; Methodist Church Ghana 1937)

Even though in the constructions in (31) and (32) there are two arguments, A and O, functionally, there is reduced valency because the subjects are not prototypical subjects. The sentences have AGENT subjects. Nonetheless, the AGENTS involved are not distinct because the pronominal form used is non-referential. This makes it impossible to identify the referent, an obvious way of downgrading the AGENT argument.

4.5 Object omission/suppression

Decreasing valency can also involve verbs that potentially have A and O arguments. However, the O argument remains suppressed either because it is understood or it is known that any one of a range of entities can fill that argument position.

4.5.1 Understood object

Verbs that allow understood objects, falling in the category of Inherent Complement Verbs (*dwanse* 'urinate', *bow* 'be drunk', *nye* 'defecate', *dɔr* 'be fatty', and sa 'dance') can occur with both overt A and O arguments. However, in many instances, the O argument is not expressed because speakers know what it is, thereby reducing the grammatical valence of the verb. The verbs and their inherent complements are listed below:

(33)   a. dwàǹsè dwáńsé
         'urinate' 'urine'

      b. bòẁ   ǹsá
         'be drunk' 'drink/alcohol'

      c. nyè   bíń
         'defecate' 'faeces'

      d. dɔ̀rè   sèràdéɛ́
         'be fatty' 'fat'

      e. sà   àsá
         'dance' 'a dance'

(34)   a. Bànyíń nó á-bòẁ   ǹsá.
         man   DEF PERF-be.drunk alcohol

         'The man is drunk (with alcohol).'

      b. Bànyíń nó á-bòẁ.
         man   DEF PERF-be.drunk

         'The man is drunk.'

(35)   a. Àpɔ̀ǹkyé nó á-dɔ̀rè   sèràdéɛ́.
         goat   DEF PERF-be.fatty fat

         'The goat is fatty.' (Lit. 'The goat is fatty with fat.') (As)

      b. Àpɔ̀ǹkyé nó á-dɔ́ré.
         goat   DEF PERF-be.fatty

         'The goat is fatty.' (As)

In (34a) and (35a) the verbs are used with their inherent complements. However, in (34b) and (35b), there are no overt complements, reflecting a reduction in the grammatical valence of the verbs *bow* 'be drunk' and *dɔre* 'be fatty'.

4.5.2 Unspecified Object

Akan has a limited number of verbs that demonstrate varying valency: monovalent, bi-valent, and trivalent. The reduction in valence of these verbs revolves around the non-expression of the O argument in all instances. This means that the single argument of the intransitive construction is an AGENT. It is more appropriate then to talk about the suppression of the O argument. Another feature of these verbs is that they tend to take various items as the O argument. Examples of the verbs are: *soa* 'carry', *hyɛ* 'wear, dress up', *sera* 'apply body lotion, smear', *sa* 'administer enema', *son* 'apply herbal nasal drop', *tua* 'douche'.

(36) a. Araba sérà-à   ǹkú.
          Araba smear-COMPL shea.butter
          'Araba applied shea butter on her body.'

   b. Araba sérà-è.
          Araba smear-COMPL
          'Araba applied (some substance) to her body.'

In the first example above, the A and O arguments are expressed. However in the second example the O argument remains unexpressed; the single argument, S, is the AGENT. The ditransitive use of these verbs is illustrated in (37):

(37) a. Araba sérà-à   àbòfrá nó ǹkú.
          Araba smear-COMPL child   DEF shea.butter
          'Araba applied shea butter on the child.'

   b. Araba sérà-à   àbòfrá nó.
          Araba smear-COMPL child   DEF
          'Araba applied (some substance) on the child.'

In (37a), the verb has three arguments: A, E, and O, in that order. In (37b), the O argument is unspecified, leaving the A and E arguments.

5 Increasing valency

There are processes that lead to an increase in the arguments of a verb. There are two main ways in Akan that this happens: the introduction of an AGENT through serialization and causativization.

5.1 Causativization

Causative constructions increase valence by adding a causing AGENT to an event. As seen in §4.3, Akan has verbs that permit the anticausative construction. The causative variant would have an AGENT present, as in (37a) and (39a). In (38b) and (39b), the change

in state of the PATIENT entities ' chains and stick, respectively ' is captured without the specification of the responsible AGENT. In (38a) and (39a), on the other hand, the events are presented with the causing AGENTS overtly stated. Effectively, in (38a) and (39a) where we have the causative constructions, the valency of the verbs has been increased by the addition of the causing AGENTS.

(38) a. Causative
      Nà   ò-bù-bú-ù   hɔ́ń   m̀pòkyèrɛ́ mù èsìǹ-ésíń.
      and 3SG.SUBJ-REDU-break-COMPL 3PL.POSS chains   in   REDU-piece
      '… and he broke their chains in pieces.' (Fante Bible Psalm 107:14; Bible Society of Ghana 1974)

      b. Anticausative
      Hɔ́ń   m̀pòkyèrɛ́ mù bù-bú-ù   èsìǹ-ésíń.
      3PL.POSS chains   in   REDU-break-COMPL REDU-piece
      'Their chains broke into pieces.'

(39) a. Causative
      Kofi á-kyéá   àbàá nó.
      Kofi PERF-bend stick DEF
      'Kofi has bent the stick.'

      b. Anticausative
      Àbàá nó á-kyéá.
      stick DEF PERF-bend
      'The stick is bent.'

5.2 Agent introduction through serialization

There are verbs that code the location or spatial configuration of an entity. Examples include: *da* 'lie, be at', *twer* 'lean', *bea* 'lie', *sɛn* 'hang', *hyɛ* ' be in', *gu* 'be in', *tar* 'paste, stick', *fam* 'stick', and *si* 'stand'. For verbs like these, the introduction of an agentive NP requires the use of a serial construction. Even though the resulting construction cannot be said to increase the valency of the verbs, it shows how an agentive argument can be introduced through the syntactic strategy of serialization.

(40) a. Nà ɔ-dzè   m̀-pòmá nó hyè-hyɛ́-ɛ̀   àdáká nó   hó.
      and 3SG.SUBJ-take PL-pole DEF REDU-put-COMPL box   3SG.POSS self
      'And he took the poles and put them on the ark.'/'And he put the poles on the ark.' (Fante Bible, Exodus 40: 20; Bible Society of Ghana 1974)

      b. M̀-pòmá nó hyè-hyè   àdáká nó   hó.
      PL-pole DEF REDU-CONT box   3SG.POSS self
      'The poles are on the ark.'

(41) a. Wɔ̀-dè   wɔ̀ń   á-kyɛ́ḿ   sèǹ-sɛ́nè   w-à-fásúó   hó
          3PL.SUBJ-take 3PL.POSS PL-shield REDU-hang/HAB 3SG.POSS-PL-wall self
          hyìá.
          meet/HAB
          'They hang their shields around your walls.' (Asante Bible, Ezekiel 27: 11;
          Bible Society of Ghana 1964)

      b. Wɔ̀ń   á-kyɛ́ḿ   sèǹ-sɛ́nè   w-à-fásúó   hó hyìá.
          3PL.POSS PL-shield REDU-hang/HAB 3SG.POSS-PL-wall self meet/HAB
          'Their shields hang around your walls.' (As)

(42) a. Kàǹdzéá nó sì   póń nó dó.
          lantern DEF stand/CONT   table DEF on
          'The lantern is on the table.'

      b. Ato dzè kàǹdzéá nó sí-ì   póń nó dó.
          Ato take lantern DEF stand-COMPL table DEF on
          'Ato placed the lantern on the table.'

6 Valence adjustment through reduplication

One valency adjusting morphological process in Akan with limited application is verbal
reduplication. So far, only two verbs have been identified in the language that change
their valence when reduplicated. The verbs are *da* 'sleep' and *di* 'eat'. Their reduplicated
forms are *deda* 'put to sleep' and *didi* 'eat'.

(43) a. Ama dà-à   àwìá.
          Ama sleep-COMPL afternoon
          'Ama slept in the afternoon.' (As)

      b. Ama dè-dá-à   àbòfrá nó àwìá.
          Ama REDU-sleep-COMPL child DEF afternoon
          'Ama put the child to sleep in the afternoon.' (As)

As shown in (43b), the reduplication of the verb *da* 'sleep' is a means by which an
AGENT argument is introduced.

The behavior of *di* when reduplicated is the reverse of the reduplication of *da*. When
di is reduplicated, it loses the capacity to have a PATIENT argument.

(44) a. Kofi dí-ì   àdùáń nó.
          Kofi eat-COMPL food DEF
          'Kofi ate the food.' (Ak)

      b. Kofi dì-dí-ì.
          Kofi REDU-eat-COMPL
          'Kofi ate.' (Ak)

c. Kofi dì-dí-ì   àdùáń nó.
Kofi REDU-eat-COMPL food   DEF
'Kofi ate the food.'

There is a polysemous use of *didi* where it takes a postpositional phrase in the post-verbal position (45). But the use of the reduplicated form of the verb *di* in this situation is metaphorical and does not contradict the case made about the verb.

(45) a. Òbíárá   dì-dí   n-àdwúmá   hó.
         everybody REDU-eat 3SG.POSS-work self
         'Everybody benefits from their work.' (Ak)
      b. Ǹkrɔ̀fó nó   dì-dí-ì   ɔ̀héń nó   ásé.
         people DEF REDU-eat-COMPL chief DEF under
         'The people sabotaged the chief.' (Ak)

7 Summary and conclusion

The notion of valency has received extensive treatment in the linguistics literature. I set out in this paper to examine the valency of Akan verbs and to investigate the morphosyntactic ways in which the valency of verbs can be modified.

I have shown in the preceding discussion that we can identify verbs in the language that are invariably monovalent ' that is, verbs that take only one core argument, the S argument; those that are bivalent ' requiring two core arguments, A and O; and those that are trivalent, needing three core arguments, that is, A, O, and E.

Apart from the verbs with invariant argument structure, there are many verbs that exhibit variations in the expression of their arguments. I have shown in the paper that overall, the morphosyntactic mechanisms by which the valency of verbs is modified in Akan fit into various cross-linguistic patterns. Akan is not known as a language with complex morphology. Consequently, the valency adjusting processes tend to be more syntactic than morphological. Verbs in the language that can undergo a reduction in the expression of their arguments do so through reflexivization, the use of reciprocals, anticausatives, impersonal constructions, and various forms of object suppression. It has been shown in the paper that where there is valency decrease resulting in only one argument being expressed, the single argument is always the S argument. Increase in verb valency is achieved through causativization, and agentivization through serialization. It has also been demonstrated that the language uses reduplication in a very limited way to adjust verb valency. As stated in the paper, the use of reduplication applies to only two verbs in the language.

Dixon & Aikhenvald (2000b: 25-27), in ending their paper, identify some topics that need to be investigated regarding the notion of valency cross-linguistically:

> Our preliminary impression is that, across the languages of the world, there tend to be more valency-increasing derivations (comitative and applicative) than valency-reducing derivations (passive, antipassive, reflexive, reciprocal, etc.). This needs to

be verified, through study of a large representative sample of languages; if it is true, linguists should seek an explanation. (Dixon & Aikhenvald 2000b: 26)

From what has been presented in this paper, it is obvious that in Akan there are more valency decreasing morphosyntactic strategies than those used to increase valency. Based on my knowledge of the languages that are genetically related to Akan, for example, Ga and Ewe, one would expect a similar tendency. However, this is an issue that needs to be investigated. More broadly, the tools that have been developed by the Leipzig Project and by Nichols and associates need to be applied to Akan and related languages in order to contribute to our further understanding of the notion valency.

Abbreviations

COL	collective	NEG	negation
COMPL	completive	OBJ	objective
CONS	consecutive	OPT	optative
CONT	continuative	PERF	perfect
DCM	dependent clause marker	PL	plural
DEF	definite	POSS	possessive
EMPH	emphatic	PROG	progressive
FUT	future	REDU	reduplication
HAB	habitual	REL	relativizer
IDM	identity marker	SG	singular
IMP	imperative	SUBJ	subject

References

Adi, Kwabena. 1973. *Brako*. Accra-Tema: Ghana Publishing Company.

Agyeman, Nana Ama. 2002. Serial verb constructions in Akan. Trondheim.

Bible Society of Ghana. 1964. *Twerɛkronkron (Asante Bible)*. Accra: Bible Society of Ghana.

Bible Society of Ghana. 1974. *Nwoma Krɔnkrɔn (Fante bible)*. Accra: Bible Society of Ghana.

Comrie, Bernard & Maria Polinsky (eds.). 1993. *Causatives and transitivity*. Amsterdam: John Benjamins.

Dixon, R. M. W. & Alexandra Y. Aikhenvald (eds.). 2000a. *Changing valency: Case studies in transitivity*. Cambridge: Cambridge University Press.

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2000b. Introduction. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Changing valency: Case studies in transitivity*, 1–29. Cambridge: Cambridge University Press.

Dolphyne, Florence A. 1988. *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Accra: Ghana Universities Press.

Essilfie, Thomas. 1977. *Serialization in Fante.* Leeds: Leeds University. (Doctoral dissertation.)

Forson, Barnabas. 1990. On the morphology of Akan serial constructions. In J. P. Hutchison & V. Manfredi (eds.), *Current approaches to African linguistics*, 63–66. Dordrecht: Foris.

van Gelderen, Elly. 2011. Basic valency and language change. (Paper presented at the 9th Association of Linguistic Typology conference, University of Hong Kong, July 21-24, 2011.)

Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity* (Studies in Language Companion Series 23), 87–120. Amsterdam: John Benjamins.

Haspelmath, Martin & Thomas Müller-Bardey. 2004. Valency change. In Geert Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas (eds.), *Morphology: An international handbook on inflection and word-formation*, vol. 2, 1130–1145. Berlin: Walter de Gruyter Mouton.

Hellan, Lars, Dorothée Beermann & Eli S. Andenes. 2003. Towards a typology of serial verb constructions in Akan. In Mary Esther Kropp Dakubu & E. Kweku Osam (eds.), *Studies in the languages of the Volta basin*, vol. 1, 61–86. Legon: University of Ghana.

Hopper, Paul & Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language* 56. 251–299.

Kambon, Obadele. 2012. *Serial verb nominalization in Akan.* Legon: University of Ghana. (Doctoral dissertation).

Keenan, Edward L. 1976. Towards a universal definition of subject. In Charles Li (ed.), *Subject and topic*, 303–334. New York: Academic Press.

Kemmer, Suzanne. 1993. *The middle voice.* Amsterdam/Philadelphia: John Benjamins.

Krampah, D. E. K. 1970. *Mbofra mfa adwen.* Accra: Bureau of Ghana Languages.

Lord, Carol. 1973. Serial verbs in transition. *Studies in African Linguistics* 4. 269–296.

Luraghi, Silvia. 2012. Basic valency orientation and the middle voice in Hittite. *Studies in Language* 36. 1–32.

Malchukov, Andrej & Bernard Comrie (eds.). In press. *Valency classes in the world's languages.* Berlin: Walter de Gruyter Mouton.

Malchukov, Andrej & Akio Ogawa. 2011. Towards a typology of impersonal constructions: A semantic map approach. In Andrej Malchukov & Anna Siewierska (eds.), *Impersonal constructions: A cross-linguistic perspective*, 19–56. Amsterdam: John Benjamins.

Martin, Henry James. 1936. *Ekumfi Benada.* Accra: Bureau of Ghana Languages.

Methodist Church Ghana. 1937. *Christian Asɔr Ndwom (the Fante hymnbook).* Cape Coast: Methodist Book Depot Ltd.

Narogg, Heiko. 2009. Synchrony and diachrony in transitivity pairs. (Paper presented at the 8th Association of Linguistic Typology conference, University of California, Berkeley, July 23-26, 2009.)

Nichols, Johanna, David A. Peterson & Jonathan Barnes. 2004. Transitivizing and detransitivising languages. *Linguistic Typology* 8. 149–211.

Osam, E. Kweku. 1994a. *Aspects of Akan grammar: A functional perspective.* Eugene, OR: University of Oregon. (Doctoral dissertation.)

Osam, E. Kweku. 1994b. From serial verbs to prepositions and the road between. *Sprachtypologie und Universalienforschung* 47. 16–36.

Osam, E. Kweku. 1996. The object relation in Akan. *Afrika und Übersee* 79. 57–83.

Osam, E. Kweku. 1997. Verb serialization and grammatical relations in Akan. In T. Givón (ed.), *Grammatical relations: A functionalist perspective*, 255–281. Amsterdam: John Benjamins.

Osam, E. Kweku. 2000. The status of the indirect object in Akan. *Journal of Asian and African Studies* 59. 171–177.

Osam, E. Kweku. 2002. The ingressive in Akan: A re-consideration. In Felix Ameka & E. Kweku Osam (eds.), *New directions in Ghanaian linguistics*, 113–126. Accra: Black Mask.

Osam, E. Kweku. 2004. *The Trondheim lectures. An introduction to Akan: Its verbal and multiverbal systems.* Legon: University of Ghana.

Osam, E. Kweku. 2008a. Akan as an aspectual language. In Felix Ameka & Mary Esther Kropp Dakubu (eds.), *Aspect and modality in Kwa languages*, 69–89. Amsterdam: John Benjamins.

Osam, E. Kweku. 2008b. Verbal alternations in Akan. *Journal of African Languages and Linguistics* 29. 97–118.

Osam, E. Kweku. 2014. Descriptive features of serial verb constructions in Akan. In Kwesi Yankah, Kofi K. Saah & Nana Aba Appiah Amfo (eds.), *A Legon reader in Ghanaian linguistics*, 11–23. Oxfordshire: Ayebia Clarke Publishing.

Plank, Franz & Aditi Lahiri. 2009. Macroscopic and microscopic typology: Basic valence orientation. (Paper presented at the 8th Association of Linguistic Typology conference, University of California, Berkeley, July 23-26, 2009.)

Schachter, Paul. 1974. A non-transformational account of serial verbs. *Studies in African Linguistics* 5. 253–270.

Siewierska, Anna. 2008. Introduction: Impersonalization: An agent-based vs. A subject-based perspective. *Transactions of the Philological Society* 106. 115–137.

Siewierska, Anna. 2011. Overlap and complementarity in reference impersonals: Man-constructions vs. Third person plural-impersonals in the languages of Europe. In Andrej Malchukov & Anna Siewierska (eds.), *Impersonal constructions: A cross-linguistic perspective*, 57–89. Amsterdam: John Benjamins.

Edoid

Chapter 8

Degema clitics and serial verb constructions at the syntax/phonology interface

Nicholas Rolle

University of California, Berkeley

Ethelbert E. Kari

University of Botswana / University of Port Harcourt

Degema exhibits two distinct clitic patterns in serial verb constructions (SVCs). In one, a set of inflectional proclitics and enclitics attaches to each verb within a SVC, resulting in [CL=V=CL ... CL=V=CL]. We refer to this as the DOUBLE-MARKED SVC PATTERN. This pattern occurs when the verbs are separated by a prosodically heavy object. In a second pattern, an inflectional proclitic attaches to the first verb of the sequence, and an inflectional enclitic attaches to the last verb of the sequence [CL=V ... V=CL], which we refer to as the SINGLE-MARKED SVC PATTERN. This pattern occurs when the verbs are not separated by an overt object, or are separated only by a prosodically light pronoun. At first glance, verbs within the Single-Marked Pattern resemble verb compounds involving verb movement (e.g. Collins 2002). We present two arguments against this verb compound hypothesis: there is unmotivated "blocking" of V2 movement by an intervening object, and the Single-Marked Pattern is found whenever the verbs are not separated by a prosodically heavy object, e.g. under dislocation. Instead, we account for the distribution of clitics through post-syntactic operations, and advocate for what we call the clitic alignment hypothesis. This hypothesis allows us to account for the puzzling fact that prosodically light pronouns may intervene between verbs in the Single-Marked Pattern. We support this hypothesis from the distribution of grammatical tone within verbal complexes.

1 Introduction

Degema exhibits two distinct clitic patterns in serial verb constructions (SVCs). In one, a set of inflectional proclitics and enclitics attaches to each verb within a SVC, resulting in [CL=V=CL ... CL=V=CL]. We refer to this as the DOUBLE-MARKED SVC PATTERN. This pattern occurs when the verbs are separated by a bisyllabic direct object. In the second

pattern, an inflectional proclitic attaches to the first verb of the sequence, and an inflectional enclitic attaches to the last verb of the sequence [CL=V … V=CL], which we refer to as the SINGLE-MARKED SVC PATTERN. This pattern occurs when the verbs are not separated by an overt object, or are separated only by a monosyllabic pronoun (a prosodically light object). With Double-Marked SVC Patterns, it is ungrammatical for the medial clitics to be absent. In contrast, although single-marking is the preferred pattern with Single-Marked SVC Patterns, double-marking is seen as acceptable but not preferred, questionable, mildly ungrammatical, or ungrammatical, depending on the speaker and context.

This paper presents these facts and an analysis which accounts for this distribution of clitics within SVCs. We analyze Degema SVCs involving nested verb phrases (VP shells in a VP complementation structure), where the second verb phrase (v_2P) is the complement of the first lexical verb (V1) (Collins 1997; 2002). At first glance, verbs within the Single-Marked Pattern resemble verb compounds. One articulated theory of verb compounds is provided by Collins (2002), in which both verbs in the SVC undergo syntactic head-movement to a higher functional position ($v_1°$) and form a complex head together. We refer to this as the VERB COMPOUND HYPOTHESIS, and present two arguments against it. The first is that there is unmotivated "blocking" of V2 movement when there is an intervening object. The second is that the Single-Marked Pattern is found whenever the verbs are not separated by a prosodically heavy object, e.g. when the object has been dropped or dislocated. These arguments point to the verbs not being a syntactic constituent.

Instead, we account for the distribution of clitics through post-syntactic operations. We advocate for what we call the CLITIC ALIGNMENT HYPOTHESIS, which states that for every clitic and every verbal host within a clause, there must be alignment between those clitics and the verbal host. In the Double-Marked Pattern contexts, because there are two verbal hosts in the clause, clitics align with both of the verbs to satisfy this alignment principle. In Single-Marked Pattern contexts, we understand that adjacent verbs form a type of verb complex, a morphophonological constituent to which the clitics align. This hypothesis allows us to account for the puzzling fact that prosodically light pronouns may intervene between verbs in the Single-Marked Pattern. We can understand this verb complex formation as sensitive to locality, but this locality is measured prosodically and not hierarchically. If the two verbs are separated by a prosodically heavy object, the verbs are not sufficiently local for the creation of the verbal complex, and therefore the Double-Marked Pattern results. If, however, we assume that monosyllabic pronouns are prosodically deficient, i.e. they do not project their own phonological word (ω). They are therefore transparent to the formation of the verbal complex.

Finally, we provide additional evidence for this constituency from the distribution of grammatical tone within verbal complexes, though we note that a full tonology of Degema has not been completed at this point.

2 Degema clitics in serial verb constructions

Degema is an Edoid language of Nigeria spoken by approximately 22,000 people (Kari 2004: 5). The language is largely head-initial with S(Aux)VO order and adjuncts (adverbials, complement phrases, prepositional phrases) follow the object. Focalized and topicalized constituents occur in the left periphery. Tense, aspect, modality, and negation are expressed through independent auxiliaries, tone patterns, and/or clitics on the verb. In this section, we provide a descriptive overview of clitics and their distribution within serial verb constructions.[1]

2.1 Overview of verbal clitics

Degema has a number of clitics which canonically appear adjacent to verbs, previously described in Kari (2002a; 2002b; 2002c; 2002d; 2003b; 2004; 2005). We discuss two types of clitics: subject agreement proclitics and tense/aspect enclitics. Proclitics agree with the subject in number, person, and humanness, and occur obligatorily in canonical finite contexts. Proclitics form two sets, what Kari (2004: 333-335) calls a Set 1 (/mV/ set, where V stands for vowel) and a Set 2 (/V/ set), provided in Table 1. Generally, Set 1 proclitics all begin with /m/, and are used in positive non-past constructions, whereas Set 2 are vowel initial except first person singular, and appear elsewhere. The proclitic receives its ATR value from its verbal host. We do not discuss here tonal alternations, nor a set of third person subject proclitics used with non-human referents. The only elements which may intervene between the lexical verb and the proclitic are auxiliaries.

Table 1: Degema subject agreement proclitics

	1st		2nd		3rd	
	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
Singular	*me/mẹ*	*mi/mị*	*mu/mụ*	*u/ụ*	*mo/mọ*	*o/ọ*
Plural	*me/mẹ*	*e/ẹ*	*ma/mạ*	*a/ạ*	*me/mẹ*	*e/ẹ*

In addition to verbal proclitics, Degema also has a series of verbal enclitics which attach to the right edge of the verb. These enclitics form a heterogeneous semantic class, unlike subject agreement proclitics. We will only discuss the tense/aspect enclitics =(V)n FACTATIVE (FE) designating past perfective with eventive verbs and present imperfective with stative verbs, and =*te/tẹ* PERFECT (PRF) 'has done'. We do not discuss additional

---

[1] Degema consonant conventions are <ḅ>=/ɓ/, <ḍ>=/ɗ/, <nw>=/ŋʷ/, <ny>=/ɲ/, <y>=/j/, <ñ>=/ŋ/, and <v>=/β/. Degema is an Advanced Tongue Root harmony language, contrasting [+ATR] /i e ɜ o u/ vs. [-ATR] /ɪ ɛ a ɔ ʊ/. Vowels with Retracted Tongue Root [-ATR] are written with a dot, e.g. <ẹ> [ɛ]. This dot is placed only under the first vowel within the word, although all vowels in that word agree for ATR. Degema orthography marks high tone with an acute accent ´ and downstepped high with a macron ¯ ; low tone is not marked.

enclitics such as =*tu/tụ* NEGATIVE IMPERATIVE (NIE) 'don't' (see Kari 2004). Under specific phonological conditions, factative =(*V*)*n* copies the final vowel of the verb and appears with a specific tonal pattern, discussed below (see also Kari 2004: 340-342 for discussion of their segmental and tonal allomorphs).

Example (1a) illustrates these clitics appearing verb-adjacent on the matrix verb *kpẹri* 'tell' and the embedded verb *tạ* 'go'. Example (1b) shows that these enclitics must attach outside of verbal "extension" suffixes, e.g. the reciprocal suffix −(*v*)*Vñine* RPS 'each other'.[2]

(1) Tense/aspect clitic adjacent to verbs

  a. **ọ**=kpérí=**té** óyi   mámū Ohoso **ọ**=tá=**té**   mú éki.
     3SG=tell=**PRF** her/him that Ohoso **3SG**=go=**PRF** to market

     '(S)he has told her/him that Ohoso has gone to market.' (Kari 2004: 63)

  b. e=gbóm-**(*én)**-óñíné=**ēn**.
     3PL=bite-**(*FE)**-RPS=**FE**

     'They bit each other.' (Kari 2004: 149)

Proclitics combine with enclitics to form distinct tense/aspect meanings, co-occurring with specific tone patterns. Set 2 proclitics combine with the factative enclitic =(*V*)*n* and the perfect enclitic =*te,* wherein the proclitic receives low tone and the verb receives high tone. Set 1 proclitics appear on the verb without one of these enclitics to convey present tense/habitual aspect or future tense, wherein the verb gets high tone and the proclitic gets high tone (except 1st person singular, which is always low). This is shown in (2)−(6). (We provide the proclitic set number within the gloss when pertinent, though usually leave it out.)

(2) Factative
   mị=dị́=**īn**.
   1SG.SET2=eat=**FE**

   'I ate.' (Kari 1997: 44)

(3) Perfect
   ọ=dị́=**tē**.
   3SG.SET2=eat=**PRF**

   '(S)he has eaten.' (Kari 2004: 284)

(4) Present Tense/Habitual
   mẹ=dị́   ịdị́yóm mị́na.
   1SG.SET1=eat food   now

   'I am eating now.' (Kari 1997: 45)

---

[2] The publication source of the Degema data is provided after each example. Those examples which do not list any publication source are native speaker interpretations by the second author not previously published, or interpretations in conjunction with other Degema speakers.

(5) Future
mẹ=ḍí   íḍíyóm úḍē.
1SG.SET1=eat food   tomorrow
'I shall eat tomorrow.' (Kari 1997: 45)

(6) Present Tense/Habitual ~ Future
mẹ=ḍí   á.
1SG.SET1=eat NPM
'I eat.' / 'I shall eat.' (Kari 1997: 45)

Example (6) illustrates the post-verbal particle *a* NPM 'non-past marker', which appears when a verb is at the end of a clause. The distribution of this particle is complex and is often not overtly realized (see Kari 2004: 278–280).

The placement of the tense/aspect enclitic depends on the type of object. When the verb precedes a vowel initial bisyllabic pronoun (*ọyi* 3SG, *eni* 1PL) in object position or with any object complement noun phrase or adjunct, the enclitic attaches next to the verb and before this complement/adjunct. This was seen in (1a) above where the enclitic =*te* appears directly adjacent to the verb *kpẹri* 'tell'. In contrast, if the verb precedes a monosyllabic pronoun in object position, the enclitic attaches to the right edge of that pronoun, and not directly next to the verb. This is shown in (7), in which the enclitic appears to the right of the pronoun *many* 'you' (pl.) and not the verb *mọn* 'see'.

(7) Surface position of enclitics with monosyllabic pronoun in object position
ọ=món   mány=**dē**. (cf. *ọ=món=**dē** mány)
3SG=see you=PRF
'(S)he has seen you (pl.).' (Kari 2004: 341)

The generalizations about clitic placement for each pronoun are summarized in Table 2. The shaded cells in this table indicate those pronouns which enclitics must attach to when present.[3] The subscript sigma $_\sigma$ indicates the pronoun is monosyllabic. The [V pron$_\sigma$=CL] pattern occurs only with pronouns and not with monosyllabic nouns or adverbials.

Table 2: Attachment site of tense/aspect enclitic with pronouns in object position

	1	2	3	XP{NP/CP/PP/etc.}
SG	*méē/mẹ́ē*	*wọ́ō*	*ọyí*	
	V pron$_\sigma$=CL	V pron$_\sigma$=CL	V=CL pron	V=CL XP
PL	*ení*	*máāny/mẹ́āny*	*ḅáāw/ḅẹ́āw*	
	V=CL pron	V pron$_\sigma$=CL	V pron$_\sigma$=CL	

---

[3] For differences in the Usokun dialect of Degema, see Offah (2000).

2.2 Serial verb construction clitic patterns

Like many West African languages, Degema exhibits robust verb serialization, expressing exhaustion/completion of a situation, directionals, benefactives, verbal comparison, comitatives, instrumentals, accompanimentals, refusal, simultaneousness, abilitatives, consequentials, and event coordination (see Kari 2003a). Resultatives and purposives are not expressed through SVCs in Degema (Kari 2004: 59–60, 206).

SVCs in Degema show two distinct surface patterns with respect to inflectional clitic placement. In the first pattern, a subject agreement proclitic appears before both verbs, and a tense/aspect enclitic appears after both verbs within the SVC. This pattern occurs when there is an intervening bisyllabic direct object between the two verbs. This Double-Marked SVC Pattern is shown in (8). In the second pattern, a proclitic appears only before the first verb, and an enclitic appears only after the second verb. This pattern occurs when there is no intervening object between the verbs, or when the only intervening element is a monosyllabic pronoun. This Single-Marked SVC Pattern is shown in (9).

(8) Double-Marked SVC Pattern: Non-adjacent verbs
Tatane **ọ**=sá=**n**   ẹ̄nám **o**=gbíyé=ē̄n.
Tatane **3sg**=shoot=**fe** animal **3sg**=kill=**fe**

'Tatane shot an animal dead.' / 'Tatane shot and killed an animal.'

(9) Single-Marked SVC Pattern: Adjacent verbs
Ohoso **o**=sóm   túl=**n**   ọ́yi.
Ohoso **3sg**=be.good reach=**fe** him

'Ohoso is as handsome as him.' / 'Ohoso is as good as him.'

In (8), both the verbs *sạ* 'shoot' and *gbiye* 'kill' are marked with a subject agreement proclitic *o=* 3sg, and an enclitic *=(V)n* fe marking factative tense/aspect. These verbs are separated by an object *ẹnam* 'animal', the object of the first verb. In contrast, in (9) only the first verb *som* 'be good' is marked by the proclitic *o=*, whereas only the second verb *tul* 'reach' is marked with the enclitic *=n*. In this case, the two verbs are not separated by an intervening object.

Recall above that monosyllabic pronouns are the only elements which may precede tense/aspect enclitics after the verb. Example (10) illustrates further that when a monosyllabic pronoun e.g. *me* 'me' intervenes between the two verbs in a SVC, this structure too exhibits the Single-Marked Pattern.

(10) Single-Marked SVC Pattern with prosodically light object *me* 'me'
Breno **o**=ḍúw   **mé** tá=ā̄n.
Breno **3sg**=follow **me** go=**fe**

'Breno went with me.' (Kari 2004: 115)

This Single-Marked Pattern happens even when both the first verb and the second verb occur with monosyllabic pronouns in object position, shown in (11) (the factative enclitic *=(V)n* is realized only tonally here due to regular allomorphic changes).

(11) Breno **ọ̀**=tútú mé ḍ̣í ḅáāw.
  Breno **3sg**=be.first me eat them**fe**

  'Breno ate them first before me.'

 Monosyllabic pronouns are the only direct objects which may intervene between verbs within a Single-Marked SVC. This is demonstrated with the bisyllabic pronoun *ọyi* 'her/him' in (12). It is ungrammatical to delete the medial clitics within the Double-Marked Pattern; cf. the minimal pair this forms with (10).

(12) Double-Marked SVC Pattern with bisyllabic object *ọyi* 'her/him'
  mi=ḍúw=*(**n**) óyi *(**mị**)=tá=ān.
  1sg=follow=**fe** her/him **1sg**=go=**fe**

  'I went with her/him.' (Kari 2004: 201)

2.3 SVCs, clitics, and tense/aspect

We illustrated above in examples (2)-(6) the role clitics play in expressing different tense/aspect meanings. Within SVCs, an interesting development can be seen. Examples (8)–(12) established two clitic patterns expressing factative and perfect tense/aspect, namely the Single-Marked and Double-Marked Patterns. In present tense/habitual aspect, however, a Set 2 proclitic is always on the first verb and a Set 1 proclitic is always on the second verb. This takes place regardless of whether the verbs are immediately adjacent, e.g. example (13a), or separated by a pronoun or noun phrase as in (13b)-(13c).

(13) Present/habitual in SVCs – Uniform Double-Marked Pattern

 a. Tẹvúró tẹvuro **ọ**=rékéréké **mǭ**=ḍí á.
   everyday   **3sg.set2**=be.slow **3sg.set1**=eat **npm**

   'Everyday, she eats them slowly.'

 b. Breno **o**=ḍúw-íy   mé **mǭ**=tá.
   Breno **3sg.set2**=follow-**iter** me **3sg.set1**=go

   'Breno goes with me all the time.'

 c. Eni **e**=ḍúw=n   óyi **mḗ**=tá.
   we **1pl.set2**=follow=**fe** him/her **1pl.set1**=go

   'We are going with him.'

 In (13c), the factative enclitic =(*V*)*n* appears on V1, though not all tokens show the presence of this element even where we would expect them given appropriate morphophonological conditions; further research is required.

 We can compare this pattern with the future. Recall that in monoverbal clauses, both present/habitual and future tense are expressed via a Set 1 proclitic, and display the same tonal pattern; cf. (4)–(6) above. In contrast, in SVCs the future tense is expressed as a Set 1 proclitic if the verbs are adjacent (14), or if the verbs are only separated by a monosyllabic object (14b). This shows the Single-Marked Pattern. If the verbs are separated by any

other object, it is expressed as a Double-Marked Pattern with identical Set 1 proclitics on both verbs (15).

(14) Future in SVCs, Single-Marked Pattern

    a. **mọ́**=rékéréké   dị́ á.
       3sg.set1=be.slow eat npm

       '(S)he will eat (them) slowly.'

    b. Breno **mó**=dụ́w   mé ta.
       Breno 3sg.set1=follow me go

       'Breno will go with me.'

(15) Future in SVCs, Double-Marked Pattern
    eni **mé**=dụ́w   oyi **mé**=tá.
    we 1pl.set1=follow him 1pl.set1=go

    'We will go with him/her.'

Table 3 summarizes these clitic patterns, presenting their distribution in both mono-verbal clauses and serial verb constructions, including adjacent verbs (V V), those separated by a monosyllabic pronoun (V σ V), and those separated by bisyllabic pronouns or noun phrases (V $\sigma\sigma$ V). We indicate a Set 1 proclitic with the 3sg *mó*=, and a Set 2 proclitic with 3sg *o*=. Cells exhibiting the Single-Marked SVC Pattern are shaded grey, whereas those exhibiting the Double-Marked Pattern remain white. (In this table, we gloss over the fact that in the present/habitual the first verb is sometimes marked with factative =(*V*)*n*.)

Table 3: Clitic patterns in Serial Verb Constructions

	Mono-verbal	Serial verb constructions		
		V V	V σ V	V σ σ V
Factative	o=V́=n	o=V́ V́=n	o=V́ ó́ V́=n	o=V́=n σ σ o=V́=n
Perfect	o=V́=tē	o=V́ V́=tē	o=V́ ó́ V́=tē	o=V́=tē σ σ o=V́=tē
Present/Habitual	mó=V́	o=V́ mō=V́	o=V́ ó́ mō=V́	o=V́ σ σ mó=V́
Future	mó=V́	mó=V́ V́	mó=V́ ó́ V́	mó=V́ σ σ mó=V́

We can generalize that if verbs in a SVC are marked with different sets of proclitics, they always show the Double-Marked Pattern, seen with present/habitual. In contrast, if the verbs receive the same set of proclitics, they only exhibit this Double-Marked Pattern if there is a prosodically heavy element intervening between the two verbs. If there is not, they are only singly-marked as a unit, and in this way have the same distribution as verbs in monoverbal clauses.

2.4 Variation in clitic patterns

We remark here on some specific locations of variation in these patterns not previously discussed (Kari 1997; 2003a; 2003b; 2004). First, we highlight a case in which no variation is found. If a SVC shows a Double-Marked Pattern, it is ungrammatical to delete the medial clitics and f orm a Single-Marked Pattern, shown below.

(16) Ungrammatical deletion of medial clitics (repeated from (12) above)
mi=ḍúw=*(**n**) óyi *(**mị**)=tá=**ān**.
'I went with her/him.' (Kari 2004: 201)

In contrast, patterns which show a distinct Single-Marked Pattern in [V V] and [V σ V] contexts exhibit variation to some degree. Examples are provided in (17)-(18) in which the Single-Marked Pattern in a [V σ V] construction is also accepted with a Double-Marked Pattern, though not preferred to the Single-Marked Pattern.

(17) Single-Marked Pattern
Tatane **ọ**=sá   ḅáw gbíyé=**ēn**.
Tatane 3sg=shoot them dead=**fe**

'Tatane shot them dead.' / 'Tatane shot and killed them.'

(18) Double-Marked Pattern
$^{(?)}$ Tatane **ọ**=sá   ḅáāw   **o**=gbíyé=**ēn**.
   Tatane 3sg=shoot them**fe** 3sg=dead=**fe**

'Tatane shot them dead.' / 'Tatane shot and killed them.'

In these contexts, the Double-Marked Pattern is typically not preferable to the Single-Marked one, and often sounds unnatural. The degree to which this sounds unnatural/dispreferred is notated via an acceptability scale (?)~?~?* before the sentence, where (?) is acceptable but dispreferred, ? is unnatural and dispreferred, and ?* is grammatically questionable. In the examples below, the grammaticality value on the left of the forward slash is the second author's judgment, whereas those on the right are judgments in consultation with other Degema speakers. Those with only one value solely represent the second author's judgment.

With the factative tense/aspect enclitic =(V)n, if the two verbs are adjacent and no object intervenes between them, a Double-Marked Pattern is interpreted as questionable, as in (19).

(19) Factative [V V] – Double-Marked Pattern acceptability interpretation
$^{?/?}$* Breno **o**=síré=**n**   **ọ**=tá=**n**.
   Breno 3sg=run=**fe** 3sg=go=**fe**

'Breno ran and went.'

As shown in (17)-(18), if the two verbs are separated by a monosyllabic pronoun i.e. [V σ V], the interpretations are somewhat better and more acceptable. Examples (20)-(21) shows the range of interpretations from (?) to ?*, for both tense/aspect enclitics.

(20) Factative [V σ V] – Double-Marked Pattern acceptability interpretations

    a. ^{(?)/?} Breno **ọ**=vón   mẹ́=ēn **o**=yí=īn.
         Breno 3sg=take me=fe 3sg=come=fe

      'Breno brought me.'

    b. ^{(?)/?} **o**=kótú   mé=ēn **ọ**=kpérí=n   īnúm.
         3sg=call me=fe 3sg=tell=fe something

      'She called me and told me something.'

    c. [?]* Breno **ọ**=tútú   mé=ēn **ọ**=ḍ́ị   bạ̄āw.
        Breno 3sg=be.first me=fe 3sg=eat them\fe

      'Breno ate them first before me.'

(21) Perfect [V σ V] – Double-Marked Pattern acceptability interpretation
    ^{(?)/?} Breno **ọ**=vón   mẹ́=**tē**   **o**=yí=**tē**.
     Breno 3sg=take me=prf 3sg=come=prf

    'Breno has brought me.'

In the future tense shown in (22), in a [V V] context only the Single-Marked Pattern is grammatical (compare 22a-22b), whereas in the [V σ V] context the Double-Marked Pattern is marginally grammatical but sounds unnatural (compare (22c)-(22d)).

(22) Future tense

    a. * **mọ́**=rékéréké   **mō**=ḍ́ị   á.
        3sg.set1=be.slow 3sg.set1=eat npm

      '(S)he will eat them slowly.'

    b. **mọ́**=rékéréké Ø ḍ́ị á.

      '(S)he will eat them slowly.'

    c. [?] Breno **mó**=ḍúw   mé **mọ́**=tá.
        Breno 3sg.set1=follow me 3sg.set1=go

      'Breno will go with me.'

    d. Breno **mó**=ḍúw mé Ø tạ́.

      'Breno will go with me.' (cf. 14b)

Regardless of the specific judgment, these Double-Marked Patterns stand in stark contrast to cases in which only one of the two clitics is doubled. These cases are unambiguously ungrammatical. Ungrammatical examples in (23) show the doubling of only the proclitic *o=* (23a), or only the enclitic *=te* (23b).

(23) Ungrammatical doubling of only one clitic

    a. Ohoso **o**=sóm   (***o**=)túl=n   ọ́yi.
       Ohoso 3sg=be.good (***3sg**=)reach=fe him

      'Ohoso is as handsome as him.'

b. Ohoso o=yí(*=tē)   kótú=té óyi.
   Ohoso 3SG=come(*=PRF) call=PRF him

   'Ohoso has come and called him.' (Kari 2003a: 285)

These patterns are summarized as a whole in Table 4. In summary, even when the Double-Marked Pattern is accepted, it is dispreferred, often sounding unnatural to grammatically questionable. We take these observations to show the Double-Marked-Pattern in this context to be highly marked, compared to the unmarked Single-Marked Pattern. The role of event and information structure may help illuminate this variability, though remains outside of the present study.

Table 4: Summary of variation in SVC clitic patterns

		SVC	Single	Double
Factative Perfect	o=V́=n o=V́=te	V V	✓	?~?*
		V σ V	✓	(?)~?~?*
		V σσ V	*	✓
Present/Habitual	o=V́ mō=V́	V V	*	✓
		V σ V	*	✓
		V σσ V	*	✓
Future	mó=V́	V V	✓	*
		V σ V	✓	?
		V σσ V	*	✓

3 Syntax of Degema SVCs

In this section, we situate the two Degema patterns in the larger typological and theoretical syntax literature, and present our assumptions regarding Degema SVC clause structure. We seek to account for those tense/aspects which show both single and double-marking, namely factative, perfect, and future. Present/habitual showing uniform double-marking remains outside of the present scope.

At first glance, the two Degema SVC patterns resemble the core-layer vs. nuclear-layer serial verb construction distinction (Foley & Olson 1985). The Single-Marked SVC Pattern resembles a nuclear-layer SVC by exhibiting singular morphological inflection and contiguity between verbs (i.e. a SINGLE COMPLEX NUCLEUS), while the Double-Marked Pattern resembles a core SVC with a looser relationship between the verbs, exhibiting double morphological inflection and non-contiguity (Foley & Olson 1985: 37–39; Crowley 2002; see summary in Cleary-Kemp 2015: 126-129). In this way, sequences [V pron$_\sigma$], [V V], [V pron$_\sigma$ V], and [V pron$_\sigma$ V pron$_\sigma$] have the same distribution as a single [V]

with respect to single-marking of inflectional categories the sequences, suggesting these complex sequences form a constituent at some level of representation which the clitics are sensitive to.

One way to capture the single-marking in Degema is to treat these SVCs as verb compounds. An example of a resultative verb compound from the nearby language Igbo is provided in (24). Like in the Degema cases, this example shows only one inflectional tense affix -[rù] which takes scope over both verbs.

(24)　ó　**ꜜtú-fù**-rù　　　ákwúꜜkwó́.
　　　he **throw-be.lost**-TNS paper

　　　'He threw paper away.' (Igbo; Lord 1975)

Within the typological literature, the similarity between verb compounds and verb serialization has been widely recognized (Margetts 1999: 101; Crowley 2002: 18; Aikhenvald 2006: a.o.), with Aikhenvald advancing "a general typological framework which encompasses multi-word and one-word SVCs" in order to "breach the artificial (and unhelpful) terminological gap" between the two types (Aikhenvald 2006: 38). Despite surface similarities, we argue below that Degema SVCs do not show properties consistent with that of verb compounding languages.

In what follows, we present our basic assumptions about the clause structure of Degema SVCs, following certain proposals in the generative syntax literature on SVCs. We then present Collins' (2002) analysis of ǂHoan verb clusters, and present two arguments against extending this analysis to Degema.

3.1 Syntactic structure of Degema SVCs

For Degema SVCs, we adopt a structure akin to Collins (1997; 2002) involving nested verb phrases (VP shells in a VP complementation structure, Cleary-Kemp 2015). The hierarchical order of heads in the verbal spine is provided in Figure 1, with a syntactic tree. We employ common generative syntax assumptions in our structure such as lexical verb phrases (VPs) embedded within functional verb phrases (vPs), and the positions of subjects and objects (for overview and justification of these assumptions, see Chomsky 1995; Adger 2003; Radford 2004; among others).[4]

In this structure, the second verb phrase (v_2P) is the complement of the first lexical verb (V1) (they are structurally adjacent and therefore sisters). To illustrate, consider the data from Collins (1997) in (25), from Ewe. Here, the second verb *ɖu* 'eat' is the complement of the first verb *ɖa* 'cook'. Note that surface word order is accounted for through syntactic movement.

(25)　Ewe
　　　M-a　ɖa　nu　ɖu.
　　　I-FUT cook thing eat

　　　'I will cook something and eat it.' (Collins 1997: 490-491)

---

[4] We abstract away from the element in the specifier of v_2P, which we have designated with neutral *e*.

```
                    ...PolP/...AuxP
                          |
                        AspP
                       /    \
                    Asp°    v1P
                           /    \
                        DP        v1'
                     [subject]   /  \
                               v1°   V1P
                                    /    \
                                 (DP)     V1'
                              [object]   /   \
                                       V1°   v2P
                                            /  \
                                          e     v2'
                                               /  \
                                             v2°   V2P
                                                  /   \
                                                V2°   (DP)
                                                    [object]
```
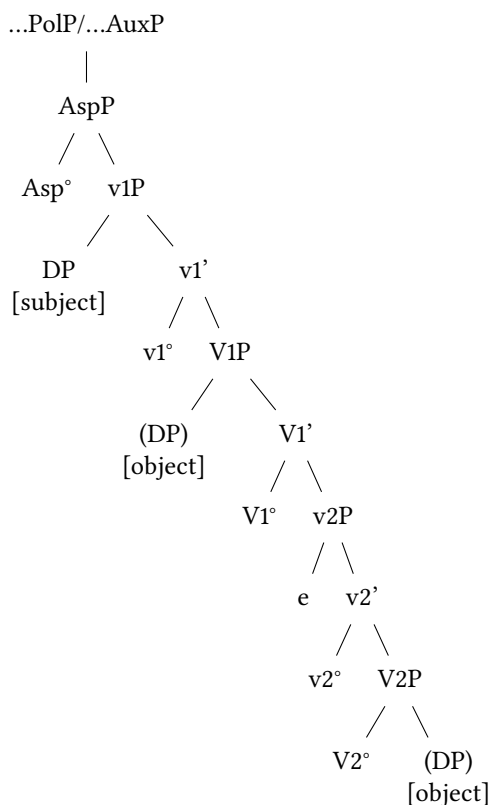
Figure 1: Degema SVC Syntax

For reasons of space, we do not compare Collins' proposal here to other syntactic analyses of SVCs (e.g. Baker 1989; Hiraiwa & Bodomo 2008; Aboh 2009; among others).

With respect to Degema clitics, we adopt that tense/aspect enclitics are in a functional head Asp(ect)° which selects a *v*P as its complement. Functional heads such as Asp° take scope over both verb phrases. Note that although the position of Asp° is above the first verb, as enclitics they will appear after the verb(s). Further, subject agreement proclitics have no designated position in this syntactic structure. We follow Embick & Noyer (2007), Kramer (2010), Norris (2014), among others that agreement features are inserted post-syntactically, and therefore these clitics are purely morphological with no unique syntactic correspondent.

3.2 Arguments against the verb compound hypothesis

With this structure in mind, we can begin to account for the distribution of the clitic SVC patterns. As mentioned above, one possibility is to consider the Single-Marked Pattern as

a verb compound, forming one syntactic constituent. One articulated theory within the generative tradition is Collins' (2002) account of ǂHoan verb compounds. An example is provided in (26) where the two verbs in bold are contiguous, and marked only once by inflectional *a-* PROG.

(26)   ǂHoan

Ma a-**q‖hu** |**'o**   djo   ki   kx'u na.
1SG PROG-**pour put.in** water PARTICLE pot   in

'I am pouring water into the pot.' (Collins 2002: 1)

Under this analysis, the syntactic structure is largely identical to that adopted in Figure 1 above involving nested VPs, in which the second verb phrase (v_2P) is the complement of the first lexical verb (V1). The critical parametric difference is that in ǂHoan the two verbs both undergo syntactic head-movement to a higher functional position ($v_1°$), and together form a complex head. We refer to this as the VERB COMPOUND HYPOTHESIS. Under this approach, we could capture the SVC patterns differences by saying that the Single-Marked Pattern is a verb compound exhibiting verb movement forming a complex head, whereas the Double-Marked Pattern exhibits no such movement.

We present two arguments against the verb movement hypothesis. The first argument is that under the verb compound hypothesis there is unmotivated "blocking" of V2 movement to form a verb compound when there is an intervening object. As described above, the Single-Marked SVC Pattern occurs in Degema only if the structure consists of two adjacent verbs, or if there is a prosodically light intervening object. If there is a prosodically heavy intervening object, the two verbs exhibit the Double-Marked Pattern. Under the verb compound hypothesis, both verbs would undergo verb movement and form a complex head. Because this movement is head-movement, it would be unclear why the presence of an object in the V_1P would block head-movement of $V_2°$ to $v°$, but allow movement of $V_1°$.

Specifically, under the verb compounding analysis we would expect the intransitive V_2 verb *tạ* 'go' to undergo head movement and form a complex verb compound with V_1 *ḍuw* 'follow' in (27a). However, (27b) shows that this is ungrammatical. Examples in (28) show it is equally ungrammatical for two transitive verbs to form a verb compound when an object is present.

(27)   Transitive + Intransitive

a.   Breno o=ḍúw   mé **tạ́**=ān.
Breno 3SG=follow me **go**=FE

'Breno went with me.' (Kari 2004: 115)

b.   *Breno o=ḍúw **tạ́** mé=ēn.

(28) Transitive + Transitive (shared object)

    a. Jzakume ọ́=tam   ị̣díyom ọ=**ḍóny**.
        Jzakume NEG\3SG=chew food   3SG=**swallow**

        'Jzakume did not chew food (and did not) swallow'. (Kari 2004: 110)

    b. *Jzakume ọ́=tam **ḍọny** ị̣díyom.

In other verb compounding languages, the presence of an object also does not block verb compounding, e.g. Juǀʼhoan (Collins 2002), Khwe (Kilian-Hatz 2006), Igbo (Lord 1975), Isu (Kießling 2011), Toqabaqita (Lichtenberk 2006; 2008), and Eastern Kayah Li (Solnit 2006). Noting that V2 in the examples above does not undergo head movement, these data cast doubt on the head movement analysis as a whole.

A second argument against the verb compound hypothesis is that the Single-Marked Pattern is found whenever the verbs are not separated by a prosodically heavy object. In many examples this happens when the first verb is intransitive, though it also is found when V_1 is transitive but the verb does not appear surface-adjacent to its syntactic argument. This can be seen in three circumstances: in object dislocation, as in both content questions (29a)-(29b) and focus constructions (29d); object-drop, i.e. where the object of the verb simply has no overt phonological instantiation (marked in the examples by an underscore), as in (30); and object relative clauses, as in (31). In each of these types, the transitive verb does not appear adjacent to its argument, and consequently appears surface-adjacent to the second verb in the SVC. Under these circumstances, a Single-Marked Pattern occurs.

(29) Dislocation of objects

    a. **Mi**=ḍúw=**n**   ovo **mị̣**=tá=**ān**?
        1SG=follow=**FE** who **1SG**=go=**FE**

        'I went with who?'

    b. Ovó nụ́   **mi**=ḍúw   Ø tá=**ān**?
        who that **1SG**=follow   go=**FE**

        'Who did I go with?'

    c. *Ovó nụ́ **mi**=ḍúūw Ø **mị̣**=tá=**ān**?
        'Who did I go with?'

    d. kụ́ ọ́yi   nụ **mi**=ḍúw   Ø tạ́=**ān**.
        not her/him that **1SG**=follow   go=**FE**

        'It was not her/him that I went with.'

(30) Object-drop
    Ohoso **ọ**=tá   ḍẹ́ __ vọ́ __ yị́ kị́yé=**n** ọ́yi   __.
    Ohoso **3SG**=go buy __ take __ come give=**FE** her/him __

    'Ohoso went and bought (something) and brought (it) to her/him.' (Kari 2004: 121)

(31) Relative clauses
      owéy   nú   **mi**=ḍúw   Ø tá=**tē**
      person that **1SG**=follow   go=**PRF**
      'the person whom I have gone with'

These data highlight that there is no strict association between clitic patterns and the argument structure of verbs in a SVC, but rather surfaces whenever the verbs are sufficiently local on the surface.

4 Phonological alignment of clitics

We have argued above that appealing to syntactic structure and operations alone is insufficient to account for the distribution of inflectional clitics within SVCs in Degema. We therefore understand post-syntactic operations to play a major role in this distribution (assuming morphological and phonological modules containing such post-syntactic operations follow syntax). This will account for why different clitic patterns surface depending largely on how local the verbs are, and why prosodically light pronouns may intervene between the verbs in a Single-Marked Pattern.

Recall from §3 that we assume tense/aspect enclitics are in a high functional Asp° projection above the verb phrases, but that subject agreement proclitics are inserted post-syntactically. Regardless of the origin of these clitics, we must account for how the clitics appear in their surface position. We adopt what we call the CLITIC ALIGNMENT HYPOTHESIS, defined below.

(32) CLITIC ALIGNMENT HYPOTHESIS: For *every* clitic and *every* verbal host within a clause, align the clitic with the verbal host

This principle states that there must be alignment between verbs and clitics within a clause, which takes place post-syntactically. Subject agreement proclitics are pre-specified as attaching to the left-edge, and tense/aspect enclitics to the right-edge. In the Double-Marked Pattern, because there are two verbal hosts in the clause, to satisfy this principle of alignment the clitics must align with both verbs, resulting in doubling of the clitics.

In contrast, in the Single-Marked Pattern proclitics are only found on V1 and enclitics only on V2, which seemingly violates this hypothesis. However, if we understand that verbs which are linearly adjacent can form a single verbal complex and that clitics align to this complex, then these structures do adhere to the clitic alignment hypothesis. Under this hypothesis, verbal hosts can be either simplex verbs or complex verb clusters.

A number of aspects of this hypothesis need to be fleshed out. One concerns which mechanism moves clitics from their original positions, tying to a rich literature on the nature and ontological status of post-syntactic movement, e.g. Nespor & Vogel (1986), Bošković (2001), Anderson's (2005) principles of STRAY ADJUNCTION and FULL INTERPRETATION, mobile morphology in Jenks & Rose (2015), among others. A second aspect is

how we can get clitics to "copy" in the Double-Marked Pattern, a type of verbal concord. We cannot satisfactorily address these issues here, but we sketch below certain aspects of Degema grammar which support our hypothesis.

One concerns the formation of a verbal complex resulting in the Single-Marked SVC Pattern. We presented evidence in §3.2 that verbs showing a Single-Marked Pattern are not the result of syntactic movement and are not a verb compound, i.e. they do not form a syntactic constituent. We do, however, understand this verbal complex as forming a morphological constituent. Rolle (2015) analyses the formation of constituents of verbs and clitics as due to the post-syntactic operation of LOCAL DISLOCATION within Distributed Morphology (Embick & Noyer 2001; 2007), which we adopt here. What is most important for our purposes here is that adjacent verbs form a single constituent (albeit not syntactically).

With these assumptions we can account for the puzzling fact that prosodically light pronouns may intervene between verbs in the Single-Marked Pattern. We can think of the formation of the verb complex as sensitive to locality, but this locality is measured prosodically and not hierarchically. If the two verbs are separated by a prosodically heavy object, the verbs are not sufficiently local for the creation of the verbal complex, and therefore the Double-Marked Pattern results. If however we assume that monosyllabic pronouns are prosodically deficient – i.e. they do not project their own phonological word (ω) – they are therefore transparent to the formation of the verbal complex. This is illustrated in (33) below, using data from the examples in (10) and (12) above.

(33) a. prosodically light pronouns

       CL= (   )=CL
     ($_\omega$)   ($_\omega$) Ø ($_\omega$)
      N   V   **σ** V
      Breno   ḍuw **me** tạ

    b. prosodically heavy pronouns

       CL= (   )=CL   CL= ( )=CL
     ($_\omega$)   ($_\omega$)   ($_\omega$)   ($_\omega$)
      N   V   **σσ**   V
      Breno   ḍuw   **ọyi**   tạ

We can understand this complex verb formation to be subject to inter-speaker variation and other complicating factors, which corresponds to the variation in the Single-Marked Pattern laid out in §2.4.

One might ask if the clitic alignment hypothesis is circular in the sense that we are defining morphological constituency based the distribution of clitics. Against this hypothesis, we find additional evidence for the constituency of the verbs in the verb complex in grammatical tone. In Degema, nouns are lexically contrastive for tone, but tone on verbs is determined solely by its grammatical context and not lexically specified. When verbs appear between a proclitic and an aspectual enclitic, all syllables are marked with high tone (H), shown in the Double-Marked Pattern in (34).

(34)   High tone pattern on verbs between clitics
        Tatane o=**kótú**=té   éni o=**kpérí**=tē   ínúm.
        Tatane 3SG=**call**=PRF us 3SG=**tell**=PRF something
        'Tatane has called us and told (us) something.' (Kari 2003a: 285)

This high tone pattern is also seen in verbal complexes with a Single-Marked Pattern, shown in (35a-b). In these cases, all words between the clitics are marked with high tone.

(35)   High tone pattern on verb complexes between clitics

        a. Ohoso o̩=**tá**   d̩é̩ **vó̩** **yí̩** **kí̩yé**=n ó̩yi.
           Ohoso 3SG=**go buy take come give**=FE her/him

           'Ohoso went and bought (something) and brought (it) to him/her.' (Kari 2004: 121)

        b. Breno o=**d̩úw mé tá̩**=ān.
           Breno 3SG=**follow me go**=FE

           'Breno went with me.' (Kari 2004: 115)

Further evidence comes from grammatical tone expressing negation. Sentential negation is realized tonally by a high-low (HL) pattern: a high tone falls on the proclitic, followed by low tone on the verb. In (36a), the proclitic *mí=* is marked with high tone, and the verb *seneke* 'think' with low tone. (Note that the factative enclitic does not occur under negation.)

(36)   HL negation tone with verb complexes

        a. mí=**seneke**   mú̩ ívom.
           NEG\1SG=**think** in inside

           'I don't think so.' (Kari 2004: 32)

        b. ó=**d̩eri me**   ká̩buló ó=**meme d̩i**   í̩díyóm yo̩.
           NEG\3SG=**know me** because NEG\3SG=**agree eat** food   the

           '(S)he refused to eat the food because (s)he doesn't know me.' (Kari 2004: 45)

In (36b), this HL negation tone also applies when the verb appears with another verb or a prosodically light pronoun, both Single-Marked contexts. Here, all elements in the complex are marked with low tone. We can compare this with Double-Marked contexts when there is a prosodically heavy intervening object. In (37a), the first verb and proclitic *vo̩n* 'take' is marked with the HL negation pattern, whereas the second verb *fi̩ya* 'cut' is marked with H, outside of the scope of negative tone. Example (37b) illustrates that marking both verbs with HL negation is ungrammatical.

(37)   Negation tone in Double-Marked Pattern

        a. Osoabo **ó̩=von**   élege o̩=**fí̩yá**.
           Osoabo NEG\3SG=**take** knife 3SG=**cut**

           'Osoabo did not use a knife to cut something.' (Kari 2004: 111)

b. * Osoabo **ọ́=von** élege **ọ́=fiya**.
Osoabo NEG\3SG=take knife NEG\3SG=cut

We can capture these data by understanding the verbs under the Single-Marked Pattern as forming a morphophonological constituent, for both clitic distribution and grammatical tone scope.[5]

5 Conclusion

We have shown that Degema exhibits two distinct clitic patterns in SVCs: a Double-Marked Pattern found when the verbs are separated by a prosodically heavy object, and a Single-Marked Pattern found when the verbs are not separated by an overt object, or are separated only by a prosodically light pronoun. We showed that although the Single-Marked Pattern superficially resembles verb compounds involving verb movement, there were two strong arguments against this position, and we therefore view these verbs as not forming a syntactic constituent. Instead, we view the verbs as forming a morphophonological complex subject to clitic alignment under a clitic alignment hypothesis. The formation of this verbal complex is subject to locality conditions defined linearly (not hierarchically), which we argued accounts for why prosodically light pronouns may intervene between verbs within the Single-Marked Pattern. We supported this analysis with evidence from grammatical tone scope, though note much more work is needed on Degema tone.

Acknowledgments

---

[5] We note that the full tonology of Degema has not been worked out, and that certain counter-exceptions occur which we cannot account for at present (see Rolle 2015). For example, the tone of the object may affect the tone on verbs in SVCs as in the Single-Marked future tense example below. Here, V1 has high tone, but V2 has a downstepped high:

(i) H - downstepped !H tone pattern on verbs in Single-Marked Pattern
mọ́=**tá** dẹ̀ ạ́bị́.
3SG=**go** buy books
'(S)he will go and buy books.'

The object in this example is lexically LH *ạbị́* /àɓí/ 'books', but is realized as HH *ạ́bị́*. If this object is modified by *yọ* 'the', both verbs have high tone with no downstep: *mọ́=**tá dẹ́** ạ́bị́ yọ* '(S)he will go and buy the books'.

Abbreviations

1,2,3	1st, 2nd, 3rd person	NPM	non-past
CL	clitic	PL	plural
FE	factative tense/aspect	PRF	perfect
FUT	future	PROG	progressive
INAUX	inceptive nonimperative auxiliary	RPS	reciprocal
ITER	iterative	SET1/SET2	Set 1/2 pronouns
NEG	negative	SG	Singular
NIE	negative imperative	TNS	tense

References

Aboh, Enoch Oladé. 2009. Clause structure and verb series. *Linguistic Inquiry* 40(1). 1–33.

Adger, David. 2003. *Core syntax: A Minimalist approach.* Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2006. Serial verb constructions in a typological perspective. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 1–68. Oxford: Oxford University Press.

Anderson, Stephen R. 2005. *Aspects of the theory of clitics.* Oxford: Oxford University Press.

Baker, Mark. 1989. Object sharing and projection in serial verb constructions. *Linguistic Inquiry* 20. 513–553.

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena.* Amsterdam: Elsevier.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: The MIT Press.

Cleary-Kemp, Jessica. 2015. *Serial verb constructions revisited: A case study from Koro.* Berkeley, CA: University of California. (Doctoral dissertation.)

Collins, Chris. 1997. Argument sharing in serial verb constructions. *Linguistic Inquiry* 28(3). 461–497.

Collins, Chris. 2002. Multiple verb movement in ǂHoan. *Linguistic Inquiry* 33(1). 1–29.

Crowley, Terry. 2002. *Serial verbs in Oceanic.* Oxford: Oxford University Press.

Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595.

Embick, David & Rolf Noyer. 2007. Distributed morphology and the syntax-morphology interface. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 289–324. Oxford: Oxford University Press.

Foley, William A. & Mike Olson. 1985. Clausehood and verb serialization. In Johanna Nichols & Anthony C. Woodbury (eds.), *Grammar inside and outside the clause: Some approaches to theory from the field*, 17–60. Cambridge: Cambridge University Press.

Hiraiwa, Ken & Adams Bodomo. 2008. Object-sharing as symmetric sharing: Predicate clefting and serial verbs in Dàgáárè. *Natural Language and Linguistic Theory* 26. 795–832.

Jenks, Peter & Sharon Rose. 2015. Mobile object markers in Moro: The role of tone. *Language* 91(2). 269–307.

Kari, Ethelbert E. 1997. *Degema*. München: Lincom Europa.

Kari, Ethelbert E. 2002a. *The source of Degema clitics*. München: Lincom Europa.

Kari, Ethelbert E. 2002b. *Cliticization, movement, and second position*. München: Lincom Europa.

Kari, Ethelbert E. 2002c. Distinguishing between clitics and affixes in Degema, Nigeria. *African Study Monographs* 23(3). 91–115.

Kari, Ethelbert E. 2002d. Distinguishing between clitics and word in Degema, Nigeria. *African Study Monographs* 23(4). 177–192.

Kari, Ethelbert E. 2003a. Serial verb constructions in Degema, Nigeria. *African Study Monographs* 24(4). 271–289.

Kari, Ethelbert E. 2003b. *Clitics in Degema: A meeting point of phonology, morphology, and syntax*. Tokyo: Research Institute for Languages, Cultures of Asia & Africa (ILCAA).

Kari, Ethelbert E. 2004. *A reference grammar of Degema*. Köln: Rüdiger Köppe Verlag.

Kari, Ethelbert E. 2005. Degema subject markers: Are they prefixes or proclitics? *Journal of West African Languages* 32(1-2). 13–20.

Kießling, Roland. 2011. *Verbal serialisation in Isu West-Ring – a Grassfields language of Cameroon*. Köln: Rüdiger Köppe Verlag.

Kilian-Hatz, Christa. 2006. Serial verb constructions in Khwe Central-Khoisan. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 108–123. Oxford: Oxford University Press.

Kramer, Ruth. 2010. The Amharic definite marker and the syntax-morphology interface. *Syntax* 13(3). 196–240.

Lichtenberk, Frantisek. 2006. Serial verb constructions in Toqabaqita. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 254–272. Oxford: Oxford University Press.

Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita*. Berlin: Mouton de Gruyter.

Lord, Carol. 1975. Igbo verb compounds and the lexicon. *Studies in African Linguistics* 61. 23–48.

Margetts, Anna. 1999. *Valence and transitivity in Saliba, an Oceanic language of Papua New Guinea*. Nijmegen: Max Planck Institute for Psycholinguistics.

Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.

Norris, Mark. 2014. *A theory of nominal concord*. Santa Cruz: University of California at Santa Cruz. (Doctoral Dissertation.)

Offah, Kio Orunyanaa. 2000. *A grammar of Degema*. Port Harcourt, Nigeria: Onyoma Research Publication.

Radford, Andrew. 2004. *Minimalist syntax: Exploring the structure of English*. Cambridge: Cambridge University Press.

Rolle, Nicholas. 2015. Morphological operations forming a conspiracy: Clitic patterns in Degema serial verb constructions. University of California at Berkeley. (Unpublished qualifying paper.)

Solnit, David B. 2006. Verb serialization in Eastern Kayah Li. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 144–159. Oxford: Oxford University Press.

Subpart I-E

Bantu

Chapter 9

Coronal palatalization in Logoori

Eleanor Glewwe

University of California, Los Angeles

Ann M. Aly

University of California, Los Angeles

Logoori (Bantu, Kenya, JE41) exhibits several palatalization processes affecting both coronal and dorsal consonants. These processes give rise to derived instances of [t͡ʃ], [d͡ʒ], and [ʃ]. While the post-alveolar affricates [t͡ʃ] and [d͡ʒ] also correspond to independent phonemes of Logoori, Leung's (1991) rule-based phonological analysis of the language considered [ʃ] to be only an allophone of /h/. We provide new instances of surface [ʃ] that, if also derived from /h/, would lead to a rule-ordering paradox involving the palatalization process Consonant-Glide Reduction and the rule /h/-Plosivization. The ordering paradox can be resolved by claiming either that [ʃ] is in fact a phoneme or that these new instances of [ʃ] are derived from something other than /h/. We opt for the latter analysis and argue that Logoori palatalizes /s/ to [ʃ] before a palatal glide (*sj* → ʃ), a coronal palatalization process not previously described for the language. We also discuss evidence that the phonemicization of [ʃ] is in progress in Logoori.

1 Introduction

This paper examines palatalization processes in Logoori, a Bantu language of western Kenya (JE41, Luyia) (Mould 1981; Bastin 2003). Logoori has a series of postalveolar consonants, [t͡ʃ], [d͡ʒ], and [ʃ], that occur as palatalized allophones of other sounds. The postalveolar affricates [t͡ʃ] and [d͡ʒ] also correspond to independent phonemes, but in the previous literature [ʃ] was considered to be only an allophone of /h/. We review the palatalization processes that have been described for Logoori and identify a rule ordering paradox in an earlier phonological analysis of the language. We then argue, still in a rule-based framework (Chomsky & Halle 1968), that [ʃ] can be derived from /s/ as well as from /h/. The proposed palatalization of /s/ to [ʃ] resolves the rule-ordering paradox. Finally, we discuss evidence that [ʃ] is in the process of becoming an independent phoneme of Logoori.

2 Background

2.1 Palatalization processes in Logoori

Leung (1991), the only phonological description of a Logoori dialect, includes a number of palatalization processes that apply to both coronal consonants and back consonants. While Leung described a different dialect than the one we consider, most of the rules she formalized are active in the dialect under study, with some differences.[1]

Leung found two coronal palatalization processes, which she grouped together as the rule Palatalization of Dental Consonants:

(1) Palatalization of Dental Consonants (Leung 1991: 117)
     j̪ → j / __ {i, u}
     n̪ → ɲ / __ {i, u}

The unusual dental glide /j̪/ that appears in the first part of (1) is never realized in the dialect considered here. Where Leung observed [j̪] on the surface, we observe only the palatal glide [j]. We retain the phoneme /j̪/, however, because nasals assimilate to it in place:

(2) /Nj̪-anz-aa/ → [n̪anzaa]
     1SG-like-PRS
     'I like'

Nevertheless, in the dialect under study, the first part of Leung's Palatalization of Dental Consonants must be subsumed under a more general rule that renders palatal all dental glides that do not undergo other rules.

The second part of Leung's Palatalization of Dental Consonants, whereby the dental nasal becomes palatal before a high vowel, does apply in the dialect we consider, as shown in (3):

(3) /N-man̪-i/ → [maɲi]
     1SG-know-FV
     'I know'

In Leung's phonological description of Logoori, the dental consonants /j̪/ and /n̪/ are the only coronal consonants that palatalize.

For back consonants, Leung posited the following rule, according to which the velar stops and /h/ palatalize before a high front vowel:

---

[1] The data in this paper were gathered in a graduate field methods class at UCLA during the 2014/2015 academic year. Our consultant, Mwabeni Indire, is a male native speaker of Logoori in his early thirties. He was raised in Nairobi, Kenya and also spent time with extended family in Vihiga County. In addition to Logoori, he speaks English and Kiswahili. If according to our consultant an utterance can stand as a complete sentence, we captialize and punctuate the free translation accordingly; otherwise no punctuation is used.

(4) Palatalization of Back Consonants (Leung 1991: 116)

k → t͡ʃ / __ i

g → d͡ʒ / __ i

h → ʃ / __ i

(except before the causative /-iz/)

In the dialect we consider, the velar stops do not undergo Palatalization of Back Consonants. Examples in (5) show that /k/ and /g/ do not palatalize before [i]:

(5) a. /N-handek-i/ → [mbandeki] (*[mbandet͡ʃi])

1SG-write-HOD.PFV

'I wrote.'

b. /a-karag-i/ → [akaɾagi] (*[akaɾad͡ʒi])

CL1-cut-HOD.PFV

'He cut.'

While the dialect Leung described palatalizes the final /k/ of the verb root 'write' and the final /g/ of the verb root 'cut' before the hodiernal perfective suffix /-i/, the dialect we consider does not. However, it does palatalize /h/ before [i], as seen in (6):[2]

(6) /a-hir-aa/ → [aʃiɾaa]

CL1-ride-PRS

'he is riding'

A second palatalization process Leung described is Consonant-Glide Reduction:

(7) Consonant-Glide Reduction (Leung 1991: 116)

kj → t͡ʃ

gj → d͡ʒ

hj → ʃ

Due to the shape of Logoori morphemes, the palatal glides in the targets of Consonant-Glide Reduction are always derived from vowels. That is, Consonant-Glide Reduction is always fed by a gliding process. Logoori has seven vowels, which we transcribe as /i e ɛ a ɔ o u/ (cf. Leung 1991). Vowel length is underlyingly contrastive, making for a total of fourteen vowel phonemes. A high vowel (/i/, /e/, /o/, or /u/) becomes [-syllabic] when it occurs before another vowel, and the formation of the glide induces compensatory lengthening of the second vowel in the sequence. This process is formalized for the front vowels in (8):

---

[2] The real picture is actually more complex. Palatalization of /h/ does not seem to apply across morpheme boundaries in this dialect: /N-veh-i/ '1SG-lie-HOD.PFV' → [mbehi] (*[mbeʃi]). Additionally, even morpheme-internally there is both type and token variation in the application of /h/ palatalization. [e-hiiɾi] 'CL9-ethnic group' does not exhibit palatalization of /h/ while /ma-higa/ 'CL6-cooking stones' is realized as both [mahiga] and [maʃiga].

(8) Gliding
     i/e   V   →   j 2 2
      1   2

Consonant-Glide Reduction is active for both velar stops and /h/ in the dialect considered here. The following derivation demonstrates how Gliding and Consonant-Glide Reduction apply in an example involving /k/:

(9) /ke-ji-siŋ-aa/
     CL7-REFL-wash-PRS
     'It washes itself.'

     ke-i-siŋ-aa   /j/Deletion[3]
     kj-ii-siŋ-aa   Gliding and Compensatory Lengthening
     t͡ʃ-ii-siŋ-aa   Consonant-Glide Reduction
     [t͡ʃ-ii-siŋ-aa]

Note that since Consonant-Glide Reduction is always fed by Gliding, which entails compensatory lengthening of the following vowel, the vowel following a postalveolar affricate derived by Consonant-Glide Reduction should always be long.

The example in (10) shows that /h/ also palatalizes and fuses with a following derived [j] (the Underlying Representation (UR) for the root 'new' will be justified below):[4]

(10) /ke-hia/ → [keʃa][5]
      CL7-new
      'new'

To recapitulate, in this dialect /h/ palatalizes to [ʃ] through both Palatalization of Back Consonants and Consonant-Glide Reduction, but /k/ and /g/ only palatalize to the postalveolar affricates [t͡ʃ] and [d͡ʒ] through Consonant-Glide Reduction. Compared with other Bantu languages, this dialect of Logoori exhibits relatively limited velar palatalization. Hyman & Moxley (1996) proposed a typology of Bantu velar palatalization classifying languages according to the extent of the environments in which they palatalize /k/ and /g/. With velar palatalization occurring only as the fusion of [kj] and [gj], the dialect we consider falls into Hyman & Moxley's (1996) most restrictive category.

From a broader crosslinguistic perspective, the palatalization processes Logoori exhibits are typical. Across languages, the most common types of consonants to be palatalized are coronal and dorsal consonants (Bateman 2011). Logoori shows cases of both

---

[3] /j/Deletion (Leung 1991: 116): /j/ → Ø / V + __

[4] In (10), it is not possible to determine whether the underlying vowel after /h/ in the root 'new' is /i/ or /e/ because both glide to [j] before another vowel, and the underlying vowel never appears on the surface. We have arbitrarily written this vowel as /i/.

[5] The reason the surface form of 'CL7-new' is [keʃa] and not [keʃaa] is because Gliding-induced compensatory lengthening is blocked word-finally (Leung 1991).

coronal and dorsal palatalization, and we will show that there are more cases of coronal palatalization than previously thought. Additionally, palatalization is most commonly triggered by high front vowels or the palatal glide (Bateman 2011), and these are the triggers seen in Logoori's palatalization processes.

2.2 The phonemic status of [t͡ʃ] and [d͡ʒ]

In addition to being allophones of the velar stops, the postalveolar affricates [t͡ʃ] and [d͡ʒ] correspond to independent phonemes of Logoori. Surface postalveolar affricates can be identified as underlying if they are followed by a short vowel. Derived [t͡ʃ] and [d͡ʒ] in this dialect always arise as a result of Consonant-Glide Reduction. As mentioned above, the vowels following these derived affricates are always long because the palatalization rule is fed by Gliding, which triggers compensatory lengthening on the vowel. Therefore, if a postalveolar affricate is not followed by a long vowel, we infer that it is nonderived, i.e. underlying.

Many words containing a postalveolar affricate, particularly [t͡ʃ], exhibit variation in their realization. Consider the following representative set of words, for which we observed variable Surface Representations (SRs):

(11) a. /ke-aŋge/ → [kjaaŋge]/[t͡ʃaaŋge]
          CL7-my
          'my'
      b. /ko-kia/ → [kokja]/[kot͡ʃa]
          CL15-dawn
          'dawn'
      c. /e-kiova/ → [ekjoova]/[et͡ʃoova]
          CL9-outside
          'outside'

In (11a–c), a surface [t͡ʃ] varies with a surface [kj] sequence. In (11a), we know the [t͡ʃ] in the palatalized variant derives from /k/ because the morpheme is the CL7 concord, which appears as [ke] in other words. In (11b) and (11c), there is no evidence from alternations that the roots 'dawn' and 'outside' begin with /k/ underlyingly, but the fact that the [kj] variant is possible suggests that [t͡ʃ] is not underlying in these roots. We therefore propose the underlying forms /kokia/ and /ekiova/ for these words. Cases like (11b) and (11c) are very common in Logoori. In such cases, we take the underlying phoneme to be the velar stop because the postalveolar affricate can vary with the velar stop on the surface.

A few morphemes containing underlying postalveolar affricates are given in (12):

(12) a. /d͡ʒi-/ [a-d͡ʒi-ɾoɾ-aa]
          CL4   3SG-CL4.OBJ-see-PRS
          'He sees it.'

   b. /ʤib/   [ke-ʤib-aa]
      answer cl7-answer-prs

      'It is answering.'

   c. /ʤiriʧi/ [e-ʤiɾiʧi]
      bull   cl9-bull

      'bull'

   d. /ʧɛɛɾe/ [m̩-ʧɛɛɾe]
      rice   cl3-rice

      'rice'

In (12a–c), we know the affricates are underlying because they are followed by short vowels. In (12d), the affricate is followed by a long vowel, but we consider this vowel to be underlyingly long. We analyze the [ʧ] in the root 'rice' as underlying because it does not alternate with [k] and because, unlike (11b) and (11c), it does not exhibit any variation in its realization ([m̩ʧɛɛɾe] cannot be produced as [m̩kjɛɛɾe]).

2.3 The phonemic status of [ʃ]

While the affricates [ʧ] and [ʤ] can be either underlying or derived from /k/ and /g/ by Consonant-Glide Reduction, Leung argues that [ʃ] is *only* ever derived from /h/. That is, according to Leung, [ʃ] is only an allophone of /h/ and does not (also) correspond to a contrastive phoneme like /ʧ/ and /ʤ/. It can be derived from /h/ by Palatalization of Back Consonants, as in [aʃiɾaa] (see ex. 6 above), or it can be derived from an underlying /hV/ sequence by Consonant-Glide Reduction, as in [keʃa] (see ex. 10 above). Leung's claim that [ʃ] is always derived from /h/ rests on two types of evidence. The first type is evidence from alternations. Many instances of surface [ʃ] can be shown to alternate with [h]. The following is a straightforward example:

(13) a. /ko-roh-a/ → [koroha]
         cl15-get.tired-fv

         'to get tired' (Leung 1991: 38)

      b. /ko-roh-i/ → [koroʃi]
         1pl-get.tired-hod.pfv

         'We got tired.' (Leung 1991: 38)

The [ʃ] in (13b) can be identified as being derived from /h/ by Palatalization of Back Consonants since the same segment surfaces as [h] when it does not precede [i], as in (13a).

Other instances of [ʃ] can be shown to be derived from /h/ because they alternate with [b]. Logoori has a rule whereby /h/ becomes [b] after a nasal. This is exemplified by the following partial paradigm:

(14) a. /N-handek-aa/ → [mbandekaa]
      1SG-write-PRS
      'I am writing.'

   b. /o-handek-aa/ → [ohandekaa]
      2SG-write-PRS
      'You are writing.'

   c. /a-handek-aa/ → [ahandekaa]
      CL1-write-PRS
      'He is writing.'

We formalize this rule as /h/-Plosivization (cf. Leung's (1991: 117) Stop Formation):

(15) /h/-Plosivization: h → b / [+nas] __ [6]

Also relevant is Nasal Place Assimilation:

(16) Nasal Place Assimilation (Leung 1991: 116)
      [+nas] → [αplace] / __ [-son, αplace]

Consider again the surface form of 'CL7-new,' given in (10) as [keʃa]. This surface [ʃ] does not precede [i], so it cannot be the result of Palatalization of Back Consonants. We claimed it derived from an underlying /hi/ sequence by Consonant-Glide Reduction but provided no evidence for the UR /hia/ 'new'. In the absence of such evidence, we would be forced to conclude that [ʃ] is underlying and therefore a phoneme of Logoori. However, the evidence exists in the form of the SR of 'CL9-new,' [embja]. The derivations of the CL7 and CL9 forms of 'new' together establish the correct UR of the root 'new' and demonstrate a crucial rule ordering:

(17) /eN-hia/   /ke-hia/
      CL9-new   CL7-new
      'new'   'new'

      eN-hja   ke-hja   Gliding
      eN-bja   -   /h/-Plosivization
      -               ke-ʃa   Consonant-Glide Reduction
      em-bja   -   Nasal Place Assimilation
      [embja]   [keʃa]

The surface alternation between [bj] and [ʃ] in the two forms of 'new' in (17) can only be accounted for by an underlying /hi/. The reason [a] is not long in [embja]

---

[6] The alternation of [h] and [b] has a historical explanation. Logoori /h/ came from *p. The bilabial stop lenited to the glottal fricative everywhere except after nasals, where it was preserved (Hyman 2003). This is why it surfaces in 1SG forms like (14a), which have a nasal prefix. Although the verb root alternations in (14a–c) have a historical account, presumably the root 'write' has been restructured to /handek/ in modern speakers' grammars, so we posit the synchronic rule /h/-Plosivization to explain the alternations.

and [keʃa], despite Gliding having applied, is because Gliding-induced compensatory lengthening is blocked word-finally (Leung 1991). The derivations in (17) also show that /h/-Plosivization must bleed Consonant-Glide Reduction. If Consonant-Glide Reduction were ordered before /h/-Plosivization, the SR of 'CL9-new' would contain a [ʃ], which is not the case. The derivation also shows that Nasal Place Assimilation must follow /h/-Plosivization, since the nasal of the prefix gets its place feature from [b].

The [ʃ]/[b] alternation can also be seen in the paradigm of the verb 'ride':

(18) a. /N-hir-aa/ → [mbiɾaa]
          1SG-ride-PRS
          'I am riding'

      b. /o-hir-aa/ → [oʃiɾaa]
          2SG-ride-PRS
          'you are riding'

      c. /a-hir-aa/ → [aʃiɾaa]
          CL1-ride-PRS
          'he is riding'

The underlying /h/ in the verb root 'ride' surfaces as [b] in the 1SG due to /h/-Plosivization and as [ʃ] in the other persons due to Palatalization of Back Consonants. The fact that the form [mbiɾaa] contains [b] and not [ʃ] illustrates that /h/-Plosivization must also bleed Palatalization of Back Consonants.

The preceding discussion has demonstrated that many surface [ʃ]s in Logoori can be shown to be derived from /h/ because of alternations with [h] and [b].

Leung's second type of evidence for the nonphonemicity of [ʃ] is distributional. In her data, those surface [ʃ]s that do not alternate with known allophones of /h/ always precede [i], the vowel that triggers palatalization of /h/ to [ʃ]. She therefore posits that these [ʃ]s are also derived from /h/. This analysis neatly accounts for the quite limited distribution of these nonalternating [ʃ]s. Thus Leung concludes that all Logoori [ʃ]s are derived from /h/ and that /ʃ/ is not a phoneme.

3 A rule-ordering paradox

In the dialect of Logoori considered here, there are surface [ʃ]s that do not seem to be derived from /h/. This creates a problem for Leung's phonological analysis and at least requires us to say something different about this dialect. Consider the following partial paradigms that feature [ʃ]:

(19) verbs with [ʃ]

      a. [Ø-ʃɔɔm-aa]
          1SG-wail-PRS
          'I am wailing.'

b. [o-ʃɔɔm-aa]
2SG-wail-PRS
'You are wailing.'

c. [a-ʃɔɔm-aa]
CL1-wail-PRS
'He is wailing.'

d. [Ø-ʃoov-aa]
1SG-throw.out-PRS
'I am throwing out'

e. [o-ʃoov-aa]
2SG-throw.out-PRS
'you are throwing out'

f. [a-ʃoov-aa]
CL1-throw.out-PRS
'he is throwing out'

The examples in (19) do not appear in Leung's description, so the forms they would have in the dialect she studied are uncertain. Under Leung's analysis, however, the instances of surface [ʃ] in (19) cannot be derived through Palatalization of Back Consonants because they do not precede [i]. They must therefore be derived from underlying /hV/ sequences through Consonant-Glide Reduction. This would lead us to posit URs for the verb roots 'wail' and 'throw out' that begin with /hi/,[7] like the root of the adjective 'new.' Since the 1SG subject prefix is /N/ and the present tense suffix is /-aa/, the URs of the 1SG forms of 'wail' and 'throw out' in (19a) and (19d) would then be /Nhiɔmaa/ and /Nhiovaa/, respectively. The beginnings of these URs are like that of 'CL9-new' in (17), but while the /Nhi/ sequence in (17) surfaces as [mbj], in (19a) and (19d) the hypothesized /Nhi/ sequences unexpectedly surface as [ʃ].

The absence of a nasal prefix may have an independent explanation. Logoori has a rule that deletes a nasal before [s]:

(20) Nasal Deletion (Leung 1991: 116)
      [+nas] → Ø / __ s

This rule is active in our speaker's dialect:

(21) /N-sɔm-aa/ → [sɔmaa]
      1SG-read-PRS
      'I am reading.'

While Leung does not mention it, it might be argued that nasals delete before all sibilants, that is, [s] and [ʃ]. Thus the reason for the absence of a nasal prefix in the 1SG

---

[7] See Footnote 4.

forms in (19a) and (19d) is the same as the reason for the absence of a nasal prefix in the 1SG form in (21).

Returning now to the forms in (19a) and (19d) in their entirety, if /h/-Plosivization bleeds Consonant-Glide Reduction, as was demonstrated in (17) with the derivation of [embja], /Nhiɔmaa/ and /Nhiovaa/ should be realized as [mbjɔɔmaa] and [mbjoovaa]. Instead, they have the [ʃ]-initial forms given in (19). The alternation between [bj] and [ʃ] that we see for the paradigm of the adjective 'new' does not arise in the paradigms for the verbs 'wail' and 'throw out.' If we assume that all three roots are underlyingly /h/-initial, as demanded by Leung's analysis, we arrive at a rule ordering paradox. This paradox is demonstrated in (22) and (23) with the words 'CL9-new' and 'I am throwing out':

(22) /eN-hia/   /N-hiov-aa/
      CL9-new   1SG-throw.out-PRS
      'new'   'I am throwing out'

      eN-hja   N-hjoov-aa   Gliding
      eN-bja   N-bjoov-aa   /h/-Plosivization
      -                  -   Consonant-Glide Reduction
      -                  -   Nasal Deletion
      em-bja   m-bjoov-aa   Nasal Place Assimilation
      [embja]   *[mbjoovaa]

The ordering of /h/-Plosivization before Consonant-Glide Reduction, established in (17) to derive the correct form of 'CL9-new,' yields the incorrect surface form for 'I am throwing out.' Switching the order of Consonant-Glide Reduction and /h/-Plosivization does not solve the problem:

(23) /eN-hia/   /N-hiov-aa/
      CL9-new   1SG-throw.out-PRS
      'new'   'I am throwing out'

      eN-hja   N-hjoov-aa   Gliding
      eN-ʃa   N-ʃoov-aa   Consonant-Glide Reduction
      -                  -   /h/-Plosivization
      e-ʃa   ʃoov-aa   Nasal Deletion[8]
      -                  -   Nasal Place Assimilation
      [eʃa]   [ʃoovaa]

---

[8] Here the environment of Nasal Deletion is provisionally extended to __ ʃ. See §5 for further discussion.

Now the correct SR of 'I am throwing out' can be derived, but the SR yielded for 'CL9-new' is incorrect. [embja] requires /h/-Plosivization to be ordered before Consonant-Glide Reduction while [ʃoovaa] requires Consonant-Glide Reduction to be ordered before /h/-Plosivization, generating the rule ordering paradox. Note that if Consonant-Glide Reduction were indeed to bleed /h/-Plosivization, there would never be any surface alternations within a paradigm between [bj] and [ʃ]. One might expect speakers to stop maintaining a UR beginning with /hi/ in their grammars if there was no surface evidence pointing to the existence of this UR. However, the paradigm of the adjective 'new' proves that such alternations do exist, necessitating the ordering of /h/-Plosivization before Consonant-Glide Reduction.

If this ordering is correct, a new account of the derivation of verb forms like [ʃoovaa] and [ʃɔɔmaa] must be found. Two explanations suggest themselves. The first is that [ʃ] is a phoneme of Logoori after all, and the [ʃ] in 1SG forms like [ʃoovaa] is underlying. The second is that the [ʃ] in [ʃoovaa] and similar forms is derived from something other than [hj]. We argue for the latter option, proposing that [ʃ] can also be derived from [sj].

4 A new coronal palatalization rule

We call the palatalization and fusion of [sj] to [ʃ] Coronal Consonant-Glide Reduction and formalize it in (24):

(24) Coronal Consonant-Glide Reduction
      sj → ʃ

The evidence for (24) comes from several quarters. First, there is intraspeaker variation that suggests that [ʃ] in [ʃɔɔmaa] and [ʃoovaa] is derived and not underlying. Consider the variants below:

(25) palatalization and fusion of [sj] to [ʃ]

    a. [Ø-ʃoov-aa]
        1SG-throw.out-PRS
        'I am throwing out'

    b. [Ø-ʃjoov-aa]
        1SG-throw.out-PRS
        'I am throwing out'

    c. [Ø-ʃɔɔm-aa]
        1SG-wail-PRS
        'I am wailing.'

    d. [Ø-sjɔɔm-aa]
        1SG-wail-PRS
        'I am wailing.'

The variation between [ʃ] and [ʃj] and between [ʃ] and [sj] in (25) is reminiscent of the variation we see in words with derived postalveolar affricates, such as the 1SG possessive pronoun for CL7 possessums (two of the three variants in (26) were given in (11) above):

(26) a. [t͡ʃ-aaŋge]
          CL7-my
          'my'

      b. [t͡ʃj-aaŋge]
          CL7-my
          'my'

      c. [kj-aaŋge]
          CL7-my
          'my'

The fact that a palatal glide sometimes surfaces in 'I am wailing' and 'I am throwing out' suggests that underlyingly there is a high front vowel between the sibilant and the vowel that is always present in the surface form. The variants in (26) show that, in cases of velar palatalization, the surface form sometimes exhibits a palatalized consonant fused with the glide (as in 26a), a palatalized consonant with the glide still present (as in 26b), and the underlying consonant and the palatal glide, with no application of Consonant-Glide Reduction (as in 26c). These variants all have analogues in (25): (25a) and (25c) show full palatalization and fusion, like (26a); (25b) shows palatalization with incomplete fusion, like (26b); and (25d) shows a form to which neither palatalization nor fusion have applied, like (26c). Crucially, the underlying consonant that is revealed in (25d) is [s], not [h]. If the root of 'wail' was underlying /h/-initial, we would expect a variant form like [hjɔɔmaa], since the [k] in [kjaaŋge] corresponds to the underlying /k/ of the CL7 prefix. In (25d), though, we see [s] instead of [h]. This is an indication that the underlying initial consonant of 'wail' is actually /s/.

Another piece of evidence comes from the forms of the verb 'grind'. Leung (1991) gives the UR of the infinitive form of this verb as /ko-siɛ-a/[9] 'CL15-grind-FV' and its SR as [kosja]. The /i/ in the root glides before /ɛ/, and /ɛ/ in turn deletes before /a/. Leung's SR exhibits no palatalization of /s/ to [ʃ], let alone fusion with [j]. In our data, however, the infinitive form of 'grind' is [koʃja]. Assuming the underlying form of the root 'grind' is still /siɛ/ in this dialect, the form [koʃja] constitutes evidence for the palatalization of /s/ in the language.

There are further, dialect-internal reasons to think the [ʃ] in [koʃja] comes from an underlying /s/. Consider the following surface forms:

(27) a. [ko-ʃ-a][10]
          CL15-burn-FV
          'to burn'

---

[9] Leung (1991) uses special notation for a vowel that is unspecified as to whether it is /i/ or /e/, but here we simply write /i/.

[10] The underlying form of the root 'burn' is /he/.

b. [m-bez-aa]
   1SG-burn-PRS
   'I am burning.'

c. [a-hez-aa]
   CL1-burn-PRS
   'He is burning.'

d. [ko-ʃj-a]
   CL15-grind-FV
   'to grind'

e. [Ø-ʃjɛɛz-aa]
   1SG-grind-PRS
   'I am grinding'

f. [a-ʃjɛɛz-aa]
   CL1-grind-PRS
   'he is grinding'

The verbs 'burn' and 'grind' have very similar infinitive forms, but in inflected forms they diverge. The 1SG present form of 'burn' begins with [mb], pointing to an underlying /h/, while the 1SG present form of 'grind' begins with [ʃj], the same consonants seen in the infinitive. That [koʃa] is underlyingly /ko-he-a/ 'CL15-burn-FV' is confirmed by the 3SG present form [ahezaa] in (27c), which exhibits on the surface the /he/ sequence that undergoes Gliding and Consonant-Glide Reduction to become [ʃ] in the infinitive. The fact that the paradigm of 'grind' in (27d–f) is not parallel to that of 'burn' in (27a–c) strongly suggests that the [ʃ] in [koʃja] does not derive from /h/. It must instead derive from /s/, which is consistent with the UR Leung gives for 'grind.' We note that the forms of 'grind' we observed always seem to show incomplete Consonant-Glide Reduction in that /s/ palatalizes to [ʃ] but the glide does not fuse with the sibilant. This appears to be part of the variation that Logoori shows in the application of its phonological processes.

5 Nasal Deletion revisited

The absence of the nasal prefix in 1SG forms like [ʃoovaa] and [ʃɔɔmaa], together with the necessary ordering of /h/-Plosivization before Consonant-Glide Reduction, constitutes evidence that the initial consonant of the verb roots in these forms is not /h/. We have argued that it is in fact /s/ and drew a parallel between the absence of the nasal 1SG prefix in [ʃoovaa] and [ʃɔɔmaa] and its absence in [sɔmaa] 'I am reading.' Leung's rule of Nasal Deletion, given in (20), is repeated in (28):

(28)   Nasal Deletion: [+nas] → Ø / __ s

This formulation only deletes nasals before [s]. If we retain this version of Nasal Deletion, it must be ordered before Coronal Consonant-Glide Reduction to derive [ʃoovaa]

and [ʃɔɔmaa]. [s] must still be present in the derivation to trigger the deletion of the 1sɢ nasal prefix. This is illustrated in (29) with 'I am throwing out,' whose UR we now give with the root as /siov/:

(29) /N-siov-aa/
      1sɢ-throw.out-prs
      'I am throwing out'

N-sjoov-aa	Gliding
sjoov-aa	Nasal Deletion
ʃoov-aa	Coronal Consonant-Glide Reduction
[ʃoovaa]	

If Coronal Consonant-Glide Reduction were ordered before Nasal Deletion in its current form, it would bleed it, and the 1sɢ nasal prefix would not delete as it must. Ordering Nasal Deletion before Coronal Consonant-Glide Reduction allows us to maintain Nasal Deletion as given in (28) and not expand its environment to include [ʃ] as well as [s].

In fact, though, expanding the environment of Nasal Deletion to include both [s] and [ʃ] actually results in a featurally simpler formulation. Compare (30), the featural equivalent of (28), and (31), whose environment covers both sibilants:

(30) Nasal Deletion: [+nas] → Ø / __ [+cont, +strid, +ant, -voice]

(31) Nasal Deletion: [+nas] → Ø / __ [+cont, +strid, -voice]

The environment that comprises [s] and [ʃ] (shown in (31)) can be described with one less feature than the environment that comprises only [s] (shown in (30)). Specifically, the additional feature [+anterior] is required to isolate [s] in (30). If simpler rules make an analysis more desirable, then the formulation of Nasal Deletion in (31) is to be preferred. If Nasal Deletion is as in (31), then it can be ordered either before or after Coronal Consonant-Glide Reduction and still yield [ʃoovaa].

Notice that if Nasal Deletion deletes nasals before [ʃ] as well as [s], the deletion of the 1sɢ nasal prefix need not be limited to verb roots whose initial consonants are underlyingly /s/. Rather, the rule could delete the 1sɢ nasal prefix before verb roots whose initial consonants are underlyingly /ʃ/. Thus far, we have continued to assume that [ʃ] is always derived, either from /h/, as Leung showed, or from /s/, as we have shown. If Logoori had verb roots that began with /ʃ/ underlyingly, though, the new formulation of Nasal Deletion would ensure that they lacked the 1sɢ nasal prefix in the present tense, which would seem to be the right effect. For instance, if we proposed that the underlying form of the verb root 'throw out' were /ʃoov/ instead of /siov/, Nasal Deletion would correctly delete the 1sɢ prefix, yielding [ʃoovaa] 'I am throwing out.' In the next section, we reconsider the phonemic status of [ʃ] and present evidence that its phonemicization may be in progress.

6 The phonemic status of [ʃ] revisited

In the absence of surface alternations between [s] and [ʃ] in paradigms like that of 'grind' (27d–f), claiming that [ʃ] derives from /sV/ requires accepting that speakers store an abstract UR for which they have little to no evidence. One reason to propose that surface [ʃ] is always derived from /h/ or /s/ and never reflects an underlying phoneme /ʃ/ is simply analytical economy. We do not want to increase the size of the phoneme inventory if we are not forced to.

We have made the case for the palatalization of /s/ to [ʃ], showing that the verb root 'grind' whose UR Leung gave as /siɛ/ is realized as [ʃjɛɛ] in this dialect. The variation between [ʃɔɔmaa] and [sjɔɔmaa] for 'I am wailing' suggests that the UR of the root 'wail' is also /s/-initial. However, it is conceivable that, without consistent alternations like those seen between [bj] and [ʃ] for /hV/ sequences, any underlying /s/ that always surfaces as [ʃ] may restructure to /ʃ/.

In fact, there is evidence of restructuring in the aforementioned cases of [bj]/[ʃ] alternations. Consider the following partial paradigms for the verbs 'haunt' and 'hurry':

(32) a. [ko-ʃookeɾ-a]
           CL15-haunt-FV
           'to haunt'

      b. [m-bjookeɾ-aa]/[Ø-ʃookeɾ-aa]
           1SG-haunt-PRS
           'I am haunting'

      c. [o-ʃookeɾ-aa]
           2SG-haunt-PRS
           'you are haunting'

      d. [a-ʃookeɾ-aa]
           CL1-haunt-PRS
           'he is haunting'

      e. [ko-ʃoog-a]
           CL15-hurry-FV
           'to hurry'

      f. [m-bjoog-aa]/[Ø-ʃoog-aa]
           1SG-hurry-PRS
           'I am hurrying'

      g. [o-ʃoog-aa]
           2SG-hurry-PRS
           'you are hurrying'

      h. [a-ʃoog-aa]
           CL1-hurry-PRS
           'he is hurrying'

The 1sɢ present forms of 'haunt' and 'hurry' (32b and 32f) exhibit variation: they are sometimes produced with an initial [mbj] sequence and sometimes produced with an initial [ʃ]. Like in the paradigm for the adjective 'new,' the alternations between [bj] and [ʃ] in (32) can only be accounted for by an underlying /h/. The 1sɢ forms with [mbj] show that the verb roots 'haunt' and 'hurry' must have the URs /hiok/ and /hiog/, respectively. 'Haunt' and 'hurry' show evidence of paradigm leveling, however. The 1sɢ forms have variants with [ʃ], the consonant that appears root-initially in the other present tense forms of the verbs (see 32c, d, g and h). Moreover, these [ʃ]-initial variants seem to be preferred; we observed more tokens of them than of the [mbj]-initial variants. When [ʃookeɾaa] and [ʃoogaa] are the 1sɢ forms, the paradigms are free of alternations. There is also no more evidence for an underlying /h/, so the verb roots have presumably been restructured.

One possibility is that their new underlying forms are respectively /siok/ and /siog/. The Coronal Consonant-Glide Reduction rule we put forth would apply to these forms, yielding surface paradigms in which the verb roots always began with [ʃ]. Some evidence in favor of an underlying /sV/ sequence, at least for 'haunt,' comes from the fact that we also observed the infinitive [koʃookeɾa] 'to haunt' produced as [koʃjookeɾa] and [kosjookeɾa]. The presence of the glide suggests an additional underlying vowel in the root, and the [s] in [kosjookeɾa] suggests the underlying consonant is /s/, not /ʃ/. We saw the same types of variation in [ʃoovaa]/[ʃjoovaa] 'I am throwing out' and [ʃɔɔmaa]/[sjɔɔmaa] 'I am wailing'.

A second possibility is that the underlying forms of the verb roots 'haunt' and 'hurry' are restructuring to /ʃook/ and /ʃoog/, respectively. If [ʃ] does not alternate with [bj] in the paradigm and if [ʃ] does not vary with [s], there is no reason for speakers to store a UR that begins with anything other than /ʃ/. If this type of restructuring appears to be occurring for /h/-initial roots like /hiok/ 'haunt' and /hiog/ 'hurry,' it seems likely that it could also be occurring for /s/-initial roots like /siɛ/ 'grind,' /siɔm/ 'wail,' and /siov/ 'throw out'. Speakers for whom this restructuring is complete must have /ʃ/ as a phoneme.

According to the analysis of Logoori we presented in the previous sections, [ʃ] is always derived by Palatalization of Back Consonants, Consonant-Glide Reduction, or Coronal Consonant-Glide Reduction. Palatalization of Back Consonants only palatalizes /h/ to [ʃ] before the vowel [i], so in all other cases where [ʃ] appears, it must be derived through a process that is fed by Gliding. As a result, all instances of [ʃ] that appear before a vowel other than [i] should be followed by a long vowel. If a systematic study measuring the duration of vowels following [ʃ] were to uncover that [ʃ] can be followed by short vowels other than [i], we would have to conclude that [ʃ] has phonemicized. Just as the existence of short vowels after the postalveolar affricates [tʃ] and [dʒ] shows that they are independent phonemes of the language, the existence of short vowels besides [i] after [ʃ] would show that it is also an independent phoneme. Such a vowel duration study is a task for future research.

7 Conclusion

Logoori exhibits a range of palatalization processes, including Palatalization of Dental Consonants (j, n̥), Palatalization of Back Consonants (/h/ → [ʃ]), and Consonant-Glide Reduction ({[kj], [gj], [hj]} → {[t͡ʃ], [d͡ʒ], [ʃ]}). While [t͡ʃ] and [d͡ʒ] are independent phonemes as well as allophones of /k/and /g/, Leung's (1991) phonological analysis considered [ʃ] to be only an allophone of /h/. Another allophone of /h/, [b], is derived by postnasal /h/-Plosivization. New data revealed a rule ordering paradox involving /h/-Plosivization and Consonant-Glide Reduction. Forms like /eNhia/ required /h/-Plosivization to precede Consonant-Glide Reduction to yield [embja] while forms like /N-hiov-aa/ required Consonant-Glide Reduction to precede /h/-Plosivization to yield [ʃoovaa]. To resolve this paradox, we argued that certain presumed underlying /h/s were actually /s/s and proposed a new rule, Coronal Consonant-Glide Reduction ([sj] → [ʃ]), thereby demonstrating that coronal palatalization is more widespread in Logoori than previously recognized. Additionally, the phonemic status of [ʃ] in this dialect of Logoori appears to be in flux. In certain cases, underlying /h/ and /s/ may be restructuring to /ʃ/. While [ʃ] was not considered a phoneme of Logoori in the past, it seems on its way to becoming one.

8 Acknowledgements

Abbreviations

1,2,3	first, second, third person	PRS	present
CL	noun class	HOD.PFV	hodiernal perfective
FV	final vowel	REFL	reflexive
OBJ	object	SG	singular
PL	plural		

References

Bastin, Yvonne. 2003. The interlacustrine zone (zone J). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 501–528. New York: Routledge.

Bateman, Nicoleta. 2011. On the typology of palatalization. *Language and Linguistics Compass* 5(8). 588–602.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* New York: Harper & Row.

Hyman, Larry M. 2003. Segmental phonology. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages,* 42–58. New York: Routledge.

Hyman, Larry M. & Jeri Moxley. 1996. The morpheme in phonological change: Velar palatalization in Bantu. *Diachronica* 13(2). 259–282.

Leung, Elizabeth Woon-Yi. 1991. *The tonal phonology of Llogoori: A study of Llogoori verbs* (Working Papers of the Cornell Phonetics Laboratory 6). Ithaca: Cornell University.

Mould, Martin. 1981. Greater Luyia. In Thomas H. Hinnebusch, Derek Nurse & Martin Mould (eds.), *Studies in the classification of Eastern Bantu languages*, 181–256. Hamburg: Helmut Buske Verlag.

Chapter 10

Remote past and phonological processes in Kaonde

Boniface Kawasha

Savannah State University

This paper examines the various morphophonological processes triggered by the remote past suffix *-ile* and the environments in which they occur in Kaonde (Bantu L41), a language spoken in Zambia. The phenomena include vowel height harmony, partial nasal consonant harmony, complete consonant assimilation, alveolar palatalization, and imbrication. Depending on the verb structure and phonological environment, the phenomena affect either the final consonant of the verb root or base, the remote past suffixal consonant, or both the verb stem-final consonant and consonant of the suffix. The remote past suffix *-ile* exhibits at least ten morphophonemic alternations. The final consonant of certain verb roots or bases also alternates between alveolar and palatalized affricate pronunciations when in contact with the high front vowel of the suffix. The imbrication process displays different patterns, and may be accompanied by segment deletion, vowel coalescence, vowel lengthening, and glide formation.

1 Introduction

This paper examines the various types of morphophonological processes triggered by the remote past inflectional suffix *-ile* and the environments in which they occur when the suffix combines with a verb root or base in Kaonde (Bantu L41; Guthrie 1967-71), one of the languages spoken in northwest Zambia. The processes include vowel height harmony, partial nasal consonant harmony, complete consonant assimilation, alveolar palatalization, and imbrication. Depending on the phonological properties of the verb, the phenomena affect either the final consonant of the verb root or verb base, the remote past suffix lateral consonant, or both the verb stem-final consonant and the suffix. The remote past suffix shows at least ten morphophonemic alternations which are *-ile, -ile, -ele, -ine, -ene, -ishe, -eshe, -izhe, -ezhe, -inye* and *-enye*.[1] In addition, the final consonant of a number of verb roots or bases alternates between the voiceless alveolar stop /t/ and

---

[1] The phonemes /ʃ/, /ʒ/, /tʃ/, /dʒ/ and /ɲ/ are orthographically represented by *sh, zh, ch, j,* and *ny*, respectively, in Kaonde. Morphemes are generally cited in their orthographic form, but IPA symbols are used at times to represent specific phonemic or phonetic forms.

the voiceless affricate /tʃ/, and between the voiced alveolar stop /d/, or the lateral /l/ and the voiced affricate /dʒ/ when it comes into contact with the high front vowel of the suffix. With respect to imbrication, the process displays different patterns, and may be accompanied by other processes that include segment deletion, vowel coalescence, vowel lengthening, and glide formation.

In Kaonde, the remote past tense is expressed by means of the inflectional suffix *-ile* in conjunction with the prefix tense maker *-a-*. Though the morpheme is referred to as *perfect* or *perfective* in many Bantu languages, it codes only the remote past tense and cannot be used to express any other meaning in Kaonde (Wright 1977: 137); it is also attested with the same function in other neighboring languages, namely Chokwe, Luvale (White 1947; 1949; 196x; Horton 1949; Yukawa 1987) and Lunda (Fisher 1984; Kawasha 2003).[2] White (1947: 6) points out that "this tense is used to denote a past action prior to yesterday" in Chokwe, Luchazi, Lunda, and Luvale. The same applies to Kaonde. Verb forms in the remote past are shown in (1).[3]

(1) a. n-a-kay-ile
         1SG-TNS-play-RP
         'I played.'

     b. w-a-tang-ile
         CL1-TNS-read-RP
         'He read.'

The sentences in (1) consist of the first person singular subject prefix *n-* and the CL1 (third person) subject prefix *w-* in initial position followed by the tense marker *-a-* attached directly to the verb roots *kay* 'play' and *tang* 'read' with the remote past suffix *-ile* as the last verbal element. Table 1 presents examples of simple verb stems on the left, and their remote past forms in the right-hand column. In the examples in the left-hand column of Table 1, the verb stem is composed of the root and the final vowel *-a*. On the right is the verb form in the remote past consisting of the verb root and the suffix *-ile*.

This paper is organized as follows: §2 discusses the various phonological processes that affect either the vowel /i/ or the consonant /l/ of the remote past suffix *-ile* in Kaonde. §3 examines palatalization which only affects the final consonant of the verb root or base. §4 presents the imbrication process and the concomitant phonological changes it brings about. §5 concludes the paper.

2 Harmony

Kaonde has three types of assimilatory phonological processes whereby segments become identical or more or less identical. They include vowel height harmony, partial

---

[2] (White 196x) is a manuscript without an exact date. The perfective or perfect meaning "does not seem to fit its use in Kaonde" (Wright 1977).

[3] Some of the data used in this paper are based on my own knowledge of the language, while others are drawn from various written sources. I verified them with speakers of Kaonde.

Table 1: Simple and remote past verb stems

Stem		Root-remote past	
j-a	'eat'	*j-ile*	'ate'
fw-a	'die'	*fw-ile*	'died'
fik-a	'arrive'	*fik-ile*	'arrived'
amb-a	'say'	*amb-ile*	'said'
end-a	'move'	*end-ele*	'moved'
imb-a	'sing'	*imb-ile*	'sang'

nasal consonant harmony, and coronal consonant assimilation. They all target either the consonantal segment /l/ or the initial vowel of the suffix by spreading a feature or features rightward from the verb base to the suffix.

2.1 Vowel height harmony

Like most of the languages spoken in Zambia, the process of vowel height harmony involves assimilation of the suffixal vowel /i/ in Kaonde. In other words, a verb root vowel induces assimilation in the vowel of the suffix. If the verb root contains the vowel /i/, /u/, or /a/, the first vowel of the suffix is a high front vowel. However, mid vowels /e/ and /o/ trigger a left-to-right direction of harmony targeting the initial high vowel of the remote past suffix, which assimilates completely to the height of the preceding mid vowel. It becomes mid front /e/ if it is preceded by a mid vowel. This type of height harmony occurs across an intervening consonant.

(2) i → e/ mid vowel (C) ___

Compare the examples in Table 2 and Table 3. Verb roots with non-mid vowels in Table 2 have an inflectional suffix with an initial high front vowel /i/, while those with mid vowels in Table 3 contain a suffix with an initial mid vowel /e/.

2.2 Nasal consonant harmony

Kaonde, like a many other Bantu languages (Greenberg 1951), presents partial nasal consonant harmony that operates on suffixes if a verb root ends in a non-palatal or non-velar nasal /m/ or /n/. This process assimilates the segment /l/ of the suffix to an alveolar nasal by harmonizing it with the preceding nasal segment of the root in a single feature. That is, the base nasal triggers left-to-right harmony targeting the lateral consonant in suffix.[4] Notice that nasal consonants induce nasalization of underlying /l/ of the remote past suffix in Table 4.

---

[4] Unlike some Bantu languages such as Kikongo, Kiyaka, Chokwe, etc. nasalization does not operate in Kaonde if the input is NVC-*ile* or CVNVC-*ile*, where the final consonant of the verb root is not a nasal.

Table 2: No height harmony in remote past suffix

Root-FV		Root-remote past	
amb-a	'speak'	*amb-ile*	'spoke'
lamb-a	'be smooth'	*lamb-ile*	'was smooth'
zh-a	'dance'	*zh-ile*	'danced'
languluk-a	'think'	*languluk-ile*	'thought'
itab-a	'answer'	*itab-ile*	'answered'
bijik-a	'shout'	*bijik-ile*	'shouted'
jimb-a	'deceive'	*jimb-ile*	'deceived'
imb-a	'sing'	*imb-ile*	'sang'
zhiik-a	'bury'	*zhiik-ile*	'buried'
ub-a	'do'	*ub-ile*	'did'
yuk-a	'know'	*yuk-ile*	'knew'

Table 3: Mid vowel height harmony in remote past suffix

keb-a	'look for'	*keb-ele*	'looked for'
fwenk-a	'swim'	*fwenk-ele*	'swam'
kenket-a	'look around'	*kenket-ele*	'looked around'
leet-a	'bring'	*leet-ele*	'brought'
nemb-a	'write'	*nemb-ele*	'wrote'
lob-a	'catch fish'	*lob-ele*	'caught fish'
lomb-a	'ask'	*lomb-ele*	'asked'
owv-a	'wash'	*ovw-ele*	'washed'
sok-a	'burn'	*sok-ele*	'burned'

Harmony does not normally operate on consonant clusters consisting of a nasal and an oral consonant. The final example in part (a) of Table 4 shows nasal harmony even though the consonant /v/ intervenes between the high vowel and the verb stem nasal. This could be attributed to the root ending in a vowel.

2.3 Complete consonant assimilation

In addition to vowel height harmony and nasal consonant harmony, Kaonde also shows complete consonant assimilation across a vowel. This is the type of progressive assimilation that affects the consonant /l/ of the remote past suffix when it attaches to verb stems that contain the alveopalatal fricative consonants /ʃ/, /ʒ/. The verb can be an underived monomorphemic root or a derived causative. The segment assimilates totally to the preceding verb root-final consonant becoming *sh*. Thus, the remote past suffix *-ile* surfaces as [iʃe], and [iʒe]. The changes in Kaonde are schematized in (3):

Table 4: Nasal consonant harmony in remote past suffix

	Root-ꜰᴠ		Root-remote past	
a.	*lam-a*	'keep'	*lam-ine*	'kept'
	kaan-a	'refuse'	*kaan-ine*	'refused'
	jim-a	'cultivate'	*jim-ine*	'cultivated'
	zhindam-a	'be quiet'	*zhindam-ine*	'was/were quiet'
	pum-a	'beat'	*pum-ine*	'beat'
	nun-a	'become fat'	*nun-ine*	'became fat'
	unvw-a	'understand, hear, feel'	*unvw-ine*	'heard'
b.	*nyem-a*	'run (away from)'	*nyem-ene*	'ran, ran away from'
	nem-a	'be heavy'	*nem-ene*	'was/were important'
	sem-a	'give birth'	*sem-ene*	'gave birth'
	ten-a	'mention'	*ten-ene*	'mentioned'
	tom-a	'drink'	*tom-ene*	'drank'
	pon-a	'fall'	*pon-ene*	'fell'

Table 5: Alveopalatal fricative assimilation in remote past

	Root-ꜰᴠ		Root-remote past	
a.	*konsh-a*	'can'	*konsh-esh-e*	'could'
	baanzh-a	'delay'	*banzh-izh-e*	'delayed'
	kozh-a	'make sick, hurt'	*kozh-ezh-e*	'made sick, hurt'
b.	*jiish-a*	'feed'	*jiish-ish-e*	'fed'

(3) /il/ → [iʃ], [iʒ] after verb stems ending in /ʃ/ and /ʒ/.

Some examples of lateral /l/ assimilation to alveopalatal fricative consonants are in Table 5. In Table 5, part a, the root-final consonant spreads rightward all its features to the suffixal consonant /l/, while in part b the consonant of the suffix assimilates completely to the causative suffix of the verb stem.

This same process also affects causativized stems that are derived from a very restricted set of verb roots that still employ the now obsolete causative suffix -*i*. When this morpheme attaches to a verb root ending with a lateral consonant, the segment /l/ of the root palatalizes to the voiced alveopalatal fricative. Compare the underived verb stems in the first column of Table 6 with their causativized verb forms in the second column.

With such causativized verbs, the /l/ of the remote past suffix assimilates completely to the causative consonant when the two suffixes occur together (Table 7).

The imbricated verb form in (4b) displays this type of assimilation when it occurs in the remote past tense.

Table 6: Palatalization of root-final /l/ in causative formation

Root-FV		Causative Stem-FV	
jil-a	'make cry'	*jizh-a*	'make cry'
bweel-a	'bring back'	*bweezh-a*	'bring back'
tweel-a	'enter'	*tweezh-a*	'cause to enter, insert'

Table 7: Alveopalatal fricative assimilation in remote past causative

Causative verb-FV		Causativized verb-remote past	
jizh-a	'make cry'	*jizh-izhe*	'made cry'
bweezh-a	'bring back'	*bwezh-ezhe*	'brought back'
tweezh-a	'bring in, insert'	*twezh-ezhe*	'brought in'

(4) a. Stem-FV
mweesh-a 'show'

b. Remote past assimilation
mwesh-eshe 'showed'

Unlike languages such as Bemba and Tonga which show double frication as a result of suffixation of the causative morpheme *-i-* (Hyman 1995; 2003) in conjunction with either the applicative suffix or the perfective suffix, the phenomenon also affects certain underived verbs.[5]

The /l/ of the remote past also assimilates to a stem-final palatalized alveolar nasal consonant *ny*. First, as in many Bantu languages, palatalization of the verb stem consonant is triggered by the affixation of a second type of causative suffix *-y-*, though this is not productive in Kaonde where it only attaches to some verb stems ending in the alveolar nasal /n/ and to some extent to the bilabial nasal /m/. However, this palatalization affects only the alveolar nasal and not /m/. Assimilation of the remote past /l/ to the stem consonant is illustrated by the data in Table 8, where the consonant of the remote past tense suffix *-ile* turns into *ny*.

3 Palatalization

Palatalization is another phonological change that alveolar consonants are susceptible to as a result of suffixation of the remote past morpheme in Kaonde. The suffix triggers this phenomenon by palatalizing the root/stem-final consonant (Table 9).[6] This morphophonological process changes either their place of articulation from alveolar to

---

[5] An anonymous reviewer has suggested that this may be a case of cyclic derivation induced by an underlying causative /i/. However, the Kaonde assimilatory process affects both underived and derived verb forms, suggesting it is not due to an underlying causative.

[6] This applies also to the causative *-ish*, applicative *-il*, and neuter *-ik*.

Table 8: Palatalized nasal assimilation in remote past

Root-FV		Root-remote past	
kany-a	'forbid'	*kany-inye*	'forbade'
fweny-a	'approach, move near'	*fweny-enye*	'approached'
chiny-a	'scare'	*chiny-inye*	'scared'

Table 9: Palatalization of root-final coronals

	Root-FV		Root-remote past	
a.	*as-a*	'shoot'	*ash-ile*	'shot'
	fis-a	'hide'	*fish-ile*	'hid'
	sans-a	'sprinkle'	*sansh-ile*	'sprinkled'
b.	*baanz-a*	'make a fire'	*baanzh-ile*	'made a fire'

alveopalatal or both place and manner of articulation. Subject to palatalization are coronal consonants, namely sibilant fricatives /s/ and /z/, the voiceless stop /t/, and the lateral /l/.[7]

The verb stem-final coronal sibilant fricative consonants /s/ and /z/ change only their anterior feature from [+anterior, + coronal] to [-anterior, + coronal]. The two segments palatalize to /ʃ/ and /ʒ/, respectively, if they occur in front of the high front vowel of the remote past suffix.[8] Compare the root-final consonant of the infinitive verbs in the left column of Table 9 and those of the verb forms of the remote past in the right hand column.

The coronal voiceless stop /t/ changes in place of articulation to become an alveopalatal affricate /tʃ/. Thus, verb root/base-final consonants alternate between [+cor, +ant] and [+cor, -ant] (Table 10).

The voiced counterpart of /t/ only occurs after the alveolar nasal /n/, realized as [nd]. It palatalizes to /dʒ/. This can be seen through the alternation of the root-final consonant in Table 11.

Also susceptible to palatalization are verb roots/bases that end in the lateral /l/ which becomes a voiced alveopalatal affricate /dʒ/ before *-ile*. In short, this consonant, which varies between [l] and a central flap [ɾ] or a lateral flap [ɺ], alternates with /dʒ/. Consider mutation of the verb final consonant /l/ to the alveopalatal affricate /dʒ/ in Table 12.[9]

---

[7] Palatalization is also very productive in other languages such as Chokwe and Luvale spoken in the same region in Zambia. In these languages, /t/ followed by /i/ becomes [tʃ], /nd/ changes to /dʒ/, /n/ to /ɲ/, /s/ to /ʃ/, and /z/ becomes [ʒ].

[8] In Nyamwezi (Maganga & Schadeberg 1992), the alveolar fricative /s/ and the lateral /l/ palatalize to [ʃ] and [dʒ], respectively, like in Kaonde. However, in Mashi, if the verb stem-final consonant is an alveolar fricative /s/ or /z/, it changes to velar stop [k] and [g], respectively, preceding the applicative extension *-iz* or *-ez* (Bashi Murhi-Orhakube 2008): *ku-yuus-a* 'to finish' → *ku-yuuk-iz-a* 'to finish for'.

[9] In Lunda the cognates to *shaala* 'remain' and t*waala* 'take' in Table 12 undergo imbrication to become *sheeli* and *tweeli* respectively.

Table 10: Palatalization of root-final /t/

Root/stem-FV		Root-remote past	
angat-a	'take from s.o.'	*angach-ile*	'took from s.o.'
it-a	'call'	*pich-ile*	'called'
kwat-a	'arrest, hold'	*kwach-ile*	'arrested, held'
pit-a	'pass'	*pich-ile*	'passed'
sant-a	'thank'	*sanch-ile*	'thanked'

Table 11: Palatalization of root-final /nt/ [nd]

Root-FV		Root-remote past	
fund-a	'teach'	*funj-ile*	'taught'
yand-a	'suffer'	*yanj-ile*	'suffered'

Table 12: Voiced palatalization of root-final /l/

Root/stem-FV		Root-remote past	
bul-a	'lack'	*buj-ile*	'lacked'
lal-a	'break'	*laj-ile*	'broke'
vul-a	'be plenty'	*vuj-ile*	'was plenty'
vuul-a	'undress'	*vuuj-ile*	'undressed'
vwal-a	'wear'	*vwaj-ile*	'wore'
shal-a	'remain'	*shaj-ile*	'remained'
twal-a	'take'	*twaj-ile*	'took'
ingil-a	'work'	*ingij-ile*	'worked'
jil-a	'cry'	*jij-ile*	'cried'

Table 13: Voiced palatalization of applicative /l/

Root-applicative-FV		Root-applicative-remote past	
fw-il-a	'die for'	*fw-ij-ile*	'died for'
twaj-il-a	'take for'	*twaj-ij-ile*	'took for'
kajip-il-a	'be angry with'	*kajip-ij-ile*	'was angry with'
tambw-il-a	'receive for'	*tambw-ij-ile*	'received for'
pum-in-a	'hit at'	*pum-ij-ile*	'hit at'
kash-il-a	'tie for'	*kash-ij-ile*	'tied for'

Table 14: Height harmony in remote past

Root/Stem-FV		Root/Stem-remote past	
harmony after root			
lek-a	'to leave off'	*lek-ele*	'left off'
leng-a	'to create'	*leng-ele*	'created'
nemb-a	'to write'	*nemb-ele*	'wrote'
komb-a	'to sweep'	*komb-ele*	'swept'
no harmony after applicative			
ketek-el-a	'trust'	*ketekej-ile*	'trusted'
lek-el-a	'to let off'	*lekej-ile*	'let off'
leng-el-a	'to create for'	*lengej-ile*	'created for'
mwek-el-a	'appear to/at'	*mwekej-ile*	'appeared to/for'
nemb-el-a	'to write to/for'	*nembej-ile*	'wrote to/for'
pemb-el-a	'to wait for'	*pembej-ile*	'waited for'
sek-el-a	'rejoice'	*sekej-ile*	'rejoiced'
komb-el-a	'to sweep for'	*kombej-ile*	'swept for'
sok-el-a	'burn for'	*sokej-ile*	'burned for'

The applicative suffix *-il* changes to *-ij* when it precedes the remote past suffix.[10] This is illustrated by the examples in the right column of Table 13 where the inflectional remote past morpheme follows the applicative. The last example in Table 13 possesses two palatalized segments, i.e. the root-final consonant /ʃ/ and /dʒ/ of the applicative extension.

The vowel of the remote past tense suffix does not harmonize in height with the mid vowel of the verb root if the former is preceded by an applicative extension. In other words, vowel harmony only affects the suffix closest to the root. Compare the remote past examples in the third row in the upper part of Table 14 with those containing a sequence of two suffixes (applicative plus remote past) in the lower part. The remote past suffix *-ele* in the right-hand column harmonizes with the vowel of the verb root. If the applicative occurs immediately after the root, it assimilates in height to the root vowel. The remote past maintains its underlying form *-ile*, as it is not in the immediate post-root position.

Nasal harmony in the remote past does not occur if a root carries the applicative before the remote past. In this case, neither the consonant of the applicative extension nor that of the remote past nasalize. Instead, the applicative *-il* palatalizes to *-ij* before the remote past suffix.

---

[10] Like the remote past, the causative, applicative, and perfective suffixes also condition palatalization. For example, based on the root *kas* 'tie', Kaonde has the intensive form *kash-ish-a* 'tie securely', *kash-il-a* 'tie for', and *kash-ij-il-a* 'tie completely for'.

Table 15: Palatalization of applicative /l/ (without nasal harmony)

Root-applicative-ꜰᴠ			palatalization of applicative /l/ before remote past	
tam-in-a	'be bad for'		*tamij-ile*	'was bad for'
pum-in-a	'beat/hit at'		*pumij-ile*	'beat/hit at'
tum-in-a	'send to'		*tumij-ile*	'sent to'
nyem-en-a	'run to/for'		*nyemej-ile*	'ran to/for'
sem-en-a	'give birth in'		*sem-ej-ile*	'gave birth in'

Table 16: No palatalization of root-final consonant after mid vowel

Root-ꜰᴠ		Root-remote past	
end-a	'walk'	*end-ele*	'walked'
send-a	'carry'	*send-ele*	'carried'
pel-a	'grind'	*pel-ele*	'ground'
loot-a	'dream'	*loot-ele*	'dreamed'
leet-a	'bring'	*leet-ele*	'brought'
kos-a	'be strong'	*kos-ele*	'was strong'

As seen in the first column of Table 15, the underlying lateral consonant of the applicative suffix of the verb stems harmonizes with the nasal consonant of the preceding the verb root. The examples in the right-hand column show the applicative /l/ (or /n/) does not change in the applicative if the remote past suffix occurs. Nevertheless, the applicative consonant /l/ close to the root palatalizes to /dʒ/. One might argue that the applicative /l/ has become /n/, which then turns into /dʒ/.

It is to be noted that the remote past suffix does not trigger palatalization of the root-final consonant if the suffix is attached directly to a verb root containing a mid vowel /e/ or /o/. This behavior suggests that vowel height harmony applies before palatalization. Consider the examples in Table 16, where palatalization fails to apply to verb roots containing a mid vowel.

However, palatalization of a stem-final /l/ takes place before the remote past suffix if that /l/ in the stem is part of the suffix or extension, even if the stem has a mid vowel. Most of the extensions have lost meanings and verbs carrying them are lexicalized. The examples in Table 17 show how the consonant /l/ palatalizes to an alveopalatal affricate *j* before the remote past suffix.

Table 17: Palatalization of stem-final consonant after mid vowel

Root-FV		Stem-remote past	
londel-a	'follow'	*londej-ile*	'followed'
imeen-a	'stand for/at'	*imeneej-ile*	'stood for/at'
mwekel-a	'appear at'	*mwekej-ile*	'appeared at'

4 Imbrication

Imbrication is a phonological pattern in some Bantu languages whereby the perfective or remote past suffix is overlaid and fuses with a preceding root or base. Depending on the language, the suffix either splits into two separate parts or the verb root or base loses its last consonant. Normally, suffixation of the remote past suffix *-ile* to a verb root results in the elongation of the verb by at least one syllable, as exemplified by the data in §2. For clarification, consider the uninflected verb forms on the left, and those with regular suffixation of the remote past morpheme on the right in Table 18.

Table 18: Simple and remote past stems

Root-FV		Root-remote past	
sak-a	'want'	*sak-ile*	'wanted'
imb-a	'sing'	*imb-ile*	'sang'
keb-a	'look for'	*keb-ele*	'looked for'
pot-a	'buy'	*pot-ele*	'bought'
tum-a	'send'	*tum-ine*	'sent'

However, this is not the case with several verbs that bear certain types of derivational suffixes, frozen suffixes, and extensions. The suffixation of the remote past morpheme to such verb bases results in a phonological phenomenon termed *imbrication* (Bastin 1983). This process usually brings about sound changes which alter the resulting verb size without increasing its number of syllables. The concomitant effects of the imbrication process, which depend on the property of the verbal suffix or extension may include segmental deletion, segmental insertion, gliding, and vowel coalescence. Consider the examples in Table 19.

As the two examples in Table 19 show, the verb forms in the right column contain the same number of syllables as those on the left column despite the addition of the two-syllable remote past morpheme. This suffixation causes phonological changes in the verb, namely the fusion of the expansion (underlying) *-al* and the inflectional remote past suffix *-ile* into a single chunk resulting in consonantal deletion, which in turn triggers vowel coalescence between /a/ and /i/ accompanied by vowel lengthening.

Table 19: Imbrication in remote past formations

Base-FV		Remote past	
iman-a	'stand'	*im-eene*	'stood'
ikal-a	'sit, stay'	*ik-eele*	'sat, stayed'

In Kaonde, it is worthwhile noting that imbrication is not uniform as it comes in different patterns. The first pattern affects a class of verb stems of more than one syllable of the structure /...C-aC-/ or /...C-aCaC-/. These verbs change their form to /...CeeC-e/ when they attach the remote past suffix. Verbs that belong to this group contain derivational extensions ending in the lateral or alveolar nasal, such as the reciprocal -*añan*, the associative -*akan*, causativized verbs, and those with a frozen stative -*al* or -*an*.[11] Most of what appear to be verb roots "suffixed" with -*al* do not occur on their own and have no independent meaning without /al/. The "suffix" forms a lexicalized unit with the root. When these stems attach the remote past suffix, the initial vowel of the suffix is inserted before the final consonant of the verb stem, and the remote past consonant deletes. This segmental loss results in the fusion of the verb stem low vowel /a/ and the vowel of the suffix /i/ into a single front mid vowel /e/ accompanied by compensatory lengthening. That is, the two adjacent vowels merge into a single vowel that combines properties of both source vowels. This imbrication process can be schematically described as in (5).

(5)   ...C-VC-ile → ...C-ViCe → C-eeCe
    ik-al-ile   ik-aile   ik-eele
    im-an-ile   im-aine   im-eene

Listed in Table 20 are examples of verb bases with frozen suffixes, reciprocal extensions, and associative extensions.

A very small set of monosyllabic verb bases of the form CV:C- or CwV:C- with the frozen suffix -*an* is also subject to imbrication. However, though these verb forms imbricate, they behave slightly different than the verb stems in Table 18 and Table 19. This is because consonant truncation does not result in fusion between the two adjacent vowels as would normally be expected in that specific context. Instead, the two juxtaposed non-identical vowels remain separate segments as they are pronounced individually. This is illustrated schematically in (6).

(6)   C(w)V:C + -*ile* → CVVCV

Examples of monosyllabic verb roots that have the pattern C(w)V:C- are in Table 21. Observe that the sequence consisting of the root /a/ and the suffixal vowel /i/ does not induce vowel fusion.

---

[11] Meeussen (1967: 90) points out that the extension -*ad*- "appears partly as an expansion, partly as a suffix with ill-defined meaning". Though the extensions are identifiable, the seeming roots to which they attach do not have independent meaning. This is attested in Bemba (Hyman 1995).

Table 20: Imbrication

	Stem-FV		Stem-remote past	
a.	*kankaman-a*	'tie together'	*kankam-eene*	'tied together'
	unvwañan-a	'understand each other'	*unvw-añeene*	'understood each other'
b.	*esekan-a*	'be equal in size'	*esek-eene*	'was equal in size'
	kasankan-a	'tie together'	*kasank-eene*	'tied together'
	kwakan-a	'unite'	*kwatak-eene*	'united'
	sambakan-a	'meet'	*sambak-eene*	'met'
	tentekan-a	'shake'	*tentek-eene*	'shook'
c.	*abany-a*	'share'	*abe-eny-e*	'shared'
	esekany-a	'compare'	*esek-eenye*	'compared'
d.	*ikal-a*	'sit, stay'	*ik-eele*	'sat, stayed'
	sangalal-a	'rejoice'	*sangal-eele*	'rejoiced'
	iman-a	'stand, stop'	*im-eene*	'stood, stopped'
	pusan-a	'differ'	*pus-eene*	'differed'

Table 21: Imbrication in monosyllabic C(w)V:C- roots

	Stem-FV		Stem-remote past	
CV:C-	*paan-a*	'share, distribute'	*pa-i-ne*	'shared, distributed'
	taan-a	'meet, find'	*ta-i-ne*	'met, found'
	kaan-a	'refuse'	*ka-i-ne*	'refused'
CwV:C-	*fwaan-a*	'be deadly'	*fwa-i-ne*	'was deadly'
	swaan-a	'inherit, succeed'	*swa-i-ne*	'inherited, succeeded'

Another group of verbs that undergoes imbrication consists of verb stems that are extended with the transitive suffix -*ul* and -*ulul*.[12] Before we discuss these verbs, it is important to look at their intransitive counterparts. The remote past of verb stems with the intransitive counterpart -*uk* and its long form -*uluk* is formed through regular suffixation, as exemplified in Table 22.

On the other hand, the remote past suffix fuses with derived transitive verbs carrying the transitive suffix -*ul,* inducing imbrication and other phonological modifications. The initial vowel of the remote past suffix is placed before the final consonant of the transitive

---

[12] The suffix is commonly known as *separative* in Bantu studies, though it has other functions. I prefer to call it *transitive*.

Table 22: Intransitive verbs with *-uk* and *-uluk*

Stem-FV		Stem-remote past	
jimuk-a	'be alert'	*jimuk-ile*	'was, were alert'
lumbuluk-a	'be clear'	*lumbuluk-ile*	'was clear'
tabuk-a	'be torn'	*tabuk-ile*	'was, were torn'
shinkuk-a	'open'	*shinkuk-ile*	'opened'
chimuk-a	'be snapped'	*chimuk-ile*	'was, were snapped'
languluk-a	'think'	*languluk-ile*	'thought'
kasuluk-a	'be united'	*kasuluk-ile*	'was were united'
chimauk-a	'be sewn'	*chimauk-ile*	'was sewn'

verb stem, and the lateral consonant of the inflectional tense morpheme deletes resulting in gliding of the final vowel /u/ or /o/ into /w/ due to adjacency with the front high vowel of the triggering suffix. This process is accompanied by compensatory lengthening of the front high vowel of the suffix. In addition, vowel height harmony occurs with verb root stems if the first vowel is mid-back /o/. That is, the vowel /i/ of the suffix lowers to the front mid vowel /e/ by partially assimilating a single feature from the preceding vowel. This imbrication process is schematically described in (7).

(7) …C-VC-ile → …C-Vile → …Cw-iile/…Cw-eele
     tamb-ul-a   tamb-uile   tamw-iile
     'receive'   'received'

The examples in Table 23 show imbrication of extended verb stems that contain the transitive suffix *-ul*.

As in most of the Bantu languages, imbrication is also observed with the verb root *món* 'see' even though it is monosyllabic without an extension.

(8) a. Root-FV
         mon-a 'see'

     b. Remote past
         mw-eene 'saw'

Imbrication also affects verb stems extended with the causative suffix *-ish* and other verb roots ending in the voiced alveopalatal fricative /ʒ/. The imbricated verb formation does not show any sign of the remote past tense suffix on the surface. The suffixal segments *-il* are completely truncated and the presence of the final vowel /e/ is the only indication that they are imbricated verb forms in the remote past. This phenomenon is also observed in isiZulu and Sesotho in which certain verbs exhibit imbrication in the perfect without making "changes to the verb stem" (Monich 2015). Furthermore, this phenomenon does not yield an expected vowel lengthening in the derivational extension domain, as is the case of most of the verbs with irregular suffixation. The form CVC-*ish-ile* changes to CVC-*ish-e* with zero suffixation, as exemplified in Table 24.

Table 23: Imbrication in transitive verbs with -ul and -ulul

	Stem-FV		Stem-remote past	
a.	*kasulul-a*	'untie'	*kasulw-iile*	'untied'
	lumbulul-a	'explain'	*lumbulw-iile*	'explained'
	shinkul-a	'open'	*shinkw-iile*	'opened'
	tabul-a	'tear'	*tabw-iile*	'tore'
	tatul-a	'begin'	*tatw-iile*	'began'
b.	*kongol-a*	'borrow'	*kongw-eele*	'borrowed'
	solol-a	'reveal'	*solw-eele*	'revealed'
	songol-a	'marry'	*songw-eele*	'married'
	tongol-a	'choose'	*tongw-eele*	'chose'
c.	*chimun-a*	'snap'	*chimw-iine*	'snapped'
	jimun-a	'alert'	*jimw-iine*	'alerted'
	chimaul-a	'sew'	*chimaw-iine*	'sewed'
	onaun-a	'destroy'	*onaw-iine*	'destroyed''

Table 24: Imbrication in verbs with causative -ish

	Root-CAUS-FV		Remote past	
a.	*kos-esh-a*	'make strong'	*kos-eshe*	'made strong'
	pos-esh-a	'cure'	*pos-eshe*	'cured'
	pot-esh-a	'sell'	*pot-eshe*	'sold'
	keb-esh-a	'cause to search'	*keb-eshe*	'made to search'
	pulush-a	'save'	*pulw-ishe*	'saved'
	unvw-ish-a	'listen attentively'	*unvw-ishe*	'listened attentively'
	took-esh-a	'whiten'	*took-eshe*	'whitened'
b.	*ipwiizh-a*	'ask'	*ipw-iizhe*	'asked'
	kamb-izh-a	'command'	*kamb-izhe*	'commanded'
	katezh-a	'be difficult'	*kat-ezhe*	'was difficult'
	pitaizh-a	'pass beyond'	*pita-izhe*	'passed beyond'

Table 25: Imbrication in remote past passive

Root-passive-ғ৵		Root-remote past-passive	
kas-w-a	'be tied'	*kash-il-w-e*	'was tied'
kelw-a	'be late'	*kel-el-w-e*	'was late'
kolw-a	'be sick'	*kol-el-w-e*	'was sick'
leng-w-a	'be created'	*leng-el-w-e*	'was created'
pum-w-a	'be hit'	*pum-in-w-e*	'was hit'
temw-a	'love'	*tem-en-w-e*	'was loved'

As seen in Table 24, the verb base on the left and the imbricated verb on the right have the same forms CVCVCV-, VCCGVCV- and VCGVCV-. What distinguishes verb forms in the two columns is only the final vowel *-a* versus *-e*.

Verb bases with the passive suffix *-w* or *-iw* imbricate as well. Some verbs with the passive morphology have developed lexicalized or non-compositional meaning, as the *-w* is frozen and has no semantic content. In addition, what appears as a root does not have independent meaning. When attached to such verbs, the *-il* of the remote past suffix splits from the final vowel and is inserted before the verb stem-final consonant. Thus, the passive suffix appears between the two parts of the remote past morpheme. However, since there is no segment deletion if the stem is monosyllabic, there is an increase in the number of syllables. This is also attested in many other Bantu languages such as Bemba (Chebanne 1993; Hyman 1995; Kula 2002). Consider the remote past verb forms in Table 25.

Consonant deletion occurs in longer verb bases displaying the passive morphology. When the remote past suffix is inserted to the immediate left of *-w*, the /l/ of the applicative or other relevant verb stem deletes, inducing either a long vowel [ii] or vowel coalescence of [ai] into [ee], depending on the property of the verb. Imbrication in the examples in Table 26, part a, induces a sequence of identical long vowels [ii]; while in part b it results in vowel coalescence. The same thing is noted with verb stems already suffixed with the applicative morpheme in the passive, as shown in part c of Table 26.

Certain verb bases already affected by the process also imbricate for a second time. In other words, the same verb may undergo double imbrication. The applicative extension triggers the first instance of imbrication, while the inflectional remote past suffix induces the second one (Table 27).

5 Conclusion

This paper has discussed the phonological processes triggered by the suffixation of the remote past tense suffix to verb roots and stems. It shows that the phenomena include vowel height harmony, nasal consonant harmony, complete consonant assimilation, and vowel harmony when preceded by mid vowels in the root. Nasal consonant harmony

Table 26: Imbrication with vowel lengthening in remote past passive

	Root-extension/ applicative-passive-FV		Remote past passive	
a.	*fwainw-a*	'have to'	*fwaiinw-e*	'had to'
	ambiw-a	'be told'	*ambiiw-e*	'was/were told'
	itabiw-a	'be accepted'	*itabiiw-e*	'was/were accepted'
	paanyiiw-a	'be given'	*panyiiw-e*	'was/were given'
	ubiw-a	'be done'	*ub-iiw-e*	'was done'
b.	*kankalw-a*	'be unable to'	*kankeelw-e*	'was unable to'
c.	*nemb-el-w-a*	'be written to'	*nemb-eel-w-e*	'was/were written to'
	twaj-il-w-a	'be taken for'	*twaj-iil-w-e*	'was/were taken for'

Table 27: Double imbrication in applied remote past passive

Root-applicative-passive-FV		Applied passive in remote past	
zhikw-il-w-a	'be unearthed for'	*zhikw-iil-w-e*	'was unearthed for'
sangw-il-w-a	'be resurrected'	*sangw-iil-w-e*	'was resurrected'

is a process that changes the lateral consonant of the suffix to a nasal /n/ following a bilabial or alveolar nasal in the root. Consonant assimilation is a long-distance phenomenon that affects the consonant of the suffix when it is attached to verb stems ending in the alveopalatal fricatives or the palatal nasal. Palatalization affects alveolar stops and the lateral which turns into the alveopalatal affricate. Regarding imbrication, the process exhibits different patterns depending on the extension and suffix that precedes the triggering remote past suffix.

Abbreviations

1SG	first person singular	RP	remote past
CAUS	causative	TNS	tense
FV	final vowel		

Acknowledgements

References

Bashi Murhi-Orhakube, Constantin. 2008. *Réflexions préliminaires sur la rédaction d'un dictionnaire bilingue Mashi-français*. Doctoral dissertation. Lille: Université de Lille III.

Bastin, Yvonne. 1983. *La finale verbale -ide et l'imbrication en Bantou* (Annales, série in-8, sciences humaines 114). Tervuren: Musée royal de l'Afrique central.

Chebanne, Andy M. 1993. Imbrication of suffixes in Setswana. (Paper presented at the 24th Annual Conference on African Linguistics, Columbus, OH, 23-25 July 1993.)

Fisher, M. K. 1984. *Lunda-Ndembu handbook*. Lusaka: Neczam.

Greenberg, Joseph H. 1951. Vowel and nasal harmony in Bantu languages. *Revue Congolaise* 8. 813–820.

Guthrie, Malcolm. 1967-71. *Comparative Bantu. An introduction to the comparative linguistics and prehistory of the Bantu languages. (4 vols.)* Farnborough: Gregg International.

Horton, Alonzo E. 1949. *A grammar of Luvale*. Johannesburg: Witwatersrand University Press.

Hyman, Larry M. 1995. Minimality and the prosodic morphology of Cibemba imbrication. *Journal of African Languages and Linguistics* 16. 3–39.

Hyman, Larry M. 2003. Segmental phonology. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages,* 42–58. New York: Routledge.

Kawasha, Boniface. 2003. *Lunda grammar: A morphosyntactic and semantic analysis*. Doctoral dissertation. Eugene: University of Oregon.

Kula, Nancy C. 2002. *The phonology of verbal derivation in Bemba*. Leiden: Leiden University. (Doctoral dissertation.)

Maganga, Clement & Thilo C. Schadeberg. 1992. *Kinyamwezi: Grammar, texts, vocabulary*. Cologne: Rüdiger Köppe Verlag.

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* III. 79–121.

Monich, Irina. 2015. Comparative morphological analysis of the perfect form in Sesotho and isiZulu. *Natural Language & Linguistic Theory* 33. 1271–1291. DOI: 10.1007/s11049–014–9280–6.

White, Charles M. N. 1947. A comparative survey of the verb forms in four Bantu languages of the west central Bantu group. *African Studies* 6(1). Johannesburg: Witwatersrand University Press, 1–.20.

White, Charles M. N. 1949. *A short Lwena grammar*. London, Cape Town & New York: Longmans, Green & Co.

White, Charles M. N. 196x. *A Chokwe grammar*. Lusaka: White Fathers, Northern Rhodesia & Nyasaland Publications Bureau.

Wright, John L. 1977. Kaonde: A brief description of the language. In Mubanga E. Kashoki & Michael Mann (eds.), *Language in Zambia: Grammatical sketches*, 109–150. Lusaka: Institute for African Studies, University of Zambia.

Yukawa, Yasutoshi. 1987. A tonal study of Luvale verbs. In Yasutoshi Yukawa, Simooya Jerome Hachipola & Ryohei Kagaya (eds.), *Studies in Zambian languages*, 1–34. Tokyo: ILCAA.

Chapter 11

The future in Logoori oral texts

Hannah Sarvasy

Australian National University

Bantu languages are renowned for their complex tense and aspect systems. Tense reference in Bantu languages is also known to have variable application, often depending on information structure or the scale of the time frames involved. With apparently four positive-polarity future tense inflections, the Luyia Bantu language Logoori (JE41) nearly tops the future tense distinction charts for Bantu and other languages. While the existence of these future-related forms is a given in the literature, the semantics and applications of Logoori future tense inflections are as yet undescribed. Logoori speakers also employ several other forms and constructions for denoting future time. This paper examines future time reference in a corpus of Logoori texts.

1 Introduction

Bantu languages are renowned for their complex tense and aspect systems (Nurse 2003; 2008). Tense reference in Bantu languages is also known to have variable application (Besha 1989; Nurse 2003: 101; Crane 2011), often depending on information structure or the scale of the time frames involved (Botne & Kershner 2008; Botne 2013). With apparently four positive-polarity future tense inflections (Mould 1981; Leung 1991; Nurse 2003), the Luyia Bantu language Logoori (JE41) nearly tops the future tense distinction charts for Bantu and other languages (Nurse 2008: 89). The existence of these inflections may be established through elicitation, but the semantics and applications of Logoori future tense inflections are as yet undescribed. With so many future tense forms, one might think that Logoori speakers would not need other auxiliary constructions for future time reference. But Logoori speakers do have several other options for denoting future time, beyond their four dedicated future tense forms. This paper examines future time reference in a corpus of Logoori texts.

Botne & Kershner (2008) and Botne (2013) present examples of tenses in Bantu languages that do not fit the timeline-extending-from-deictic-center model in Comrie (1985). As Botne (2013) points out, a tense system may include forms that fuse epistemic reality and time reference. This is likely the case for Logoori, as well. Three of the Logoori future tenses seem to be staggered along a two-dimensional timeline, overlapping for most of

their ranges but differentiated in the closest time frames. Another one or two "tenses" may differ from these three in epistemic certainty.

Natural speech is the best context to gauge discourse functions of tense inflections, in consultation with a Logoori native speaker. The data used for this paper is a small corpus including 20 Logoori texts of varying lengths from three sources: a) original narratives recorded by UCLA Field Methods consultant Mwabeni Indire; b) original narratives and religious songs recorded by various female and male Logoori speakers in Kenya on behalf of Michael Diercks; and c) interview clips in Logoori from Sandra Nichols and Joseph Ssennyonga's 1976 film *Maragoli*. All texts were transcribed in consultation with Mr. Indire using the software ELAN (Sloetjes & Wittenberg 2008). Throughout, texts are cited by speaker's name and line number. Occasional examples volunteered by Mr. Indire lack this citation.

Logoori is far from monolithic, and at least two major dialect groups are represented in the corpus, judging by binary distinctions in forms. Apparent differences range from phonology — one dialect's palatal glide corresponds to an interdental consonant in the other — to morphology. Based on Mr. Indire's description and reports from Michael Diercks and colleagues, the dialect situation is complex; nowadays, at least, dialect differences do not seem to correspond to geography or village locations. Further, while forms generally differ in maximally two ways, individual speakers may employ different combinations of the binary distinctions. Examples here are drawn from all Logoori dialects present in the corpus. Where a dialect difference may contribute to a difference in morphology, this is noted.

Botne (2013: 12) warns that any discussion of Bantu tense must cover temporal "domain, time region, and time scale" for each form. This is beyond the scope of this paper. Instead, this paper aims at exposing the ways in which various future tense inflections, and other constructions with future reference, are used in the texts corpus.

2 What has been said about Logoori future tenses

Earlier formal descriptions of Logoori future tenses are in Mould (1981: 206), Leung (1991: 174–189, 204–207, 273–284, 285–322), and Nurse (2003: 100).[1] These accord in the number of future tenses listed: four, including Near, Middle, Far/Remote, and "Indefinite" (Mould, Leung) or "Uncertain" (Nurse). All three descriptions emphasize form over meaning or function. Mould (1981) and Nurse (2003) indicate meaning only indirectly through their chosen labels. Leung (1991) briefly describes use of only two of the four forms: her Near Future (1991: 174) and Middle Future (1991: 285).

None of the sources explain the label Indefinite or Uncertain. In general (for instance, in Johnson 1977: 20), indefinite future tenses are those that can apply to many discrete segments of the presumed timeline, while uncertain implies epistemic evaluation of the likelihood of the event. These would seem to be two different concepts. As described in

---

[1] Logoori is assumed here, following Leung (1991), to have seven phonological vowel distinctions. These are written here as: /i ɪ e a o ʊ u/. Leung's /i/ corresponds to /i/ here, while Leung's /I/ corresponds to /ɪ/; Leung's /u/ corresponds to /u/ here, and her /U/ corresponds to /ʊ/.

§4, the relevant form designated by these terms is the rarest in the present corpus, so this paper will not resolve its semantics.

The forms and terminology given by previous sources for the various Logoori future tenses are in Table 1, along with the terminology used here and an additional form in the last column that is introduced in this paper.[2]

Table 1: Logoori future tense forms

	SP-*ra*-__-*a*	*na*-SP-__-*e*	SP-*raka*-__-*e*	SP-*rika*-__-*e*	SP-*ri*-__-*a*	*naa*-SP-__-*e*
Mould 1981	Near Future	Middle Future	Far Future	—	Indefinite Future	—
Leung 1991	Near Future; hodiernal	Middle Future; tomorrow through several years hence	Far Future (variant)	Far Future (variant)	Indefinite Future	—
Nurse 2003	Near Future	Middle Future	—	Far Future	Uncertain Future	—
This paper	Hodiernal + crastinal; also unspecified future	Post-crastinal	Post-crastinal, more distant than Middle Future	—	Uncertain Future?	Post-crastinal, more distant than Middle Future

These forms are discussed in §3–§6, and additional constructions with future time reference present in the corpus are explored in §7.

3 The Near Future that moonlights as General Future

According to Mr. Indire, the form /SP-ra-__-a/ may be understood by speakers to describe actions and events limited to a time between deictic center 'now' and the end of the following day, i.e. 'later today' through the end of 'tomorrow' (crastinal). This is counter to the description by Leung (1991: 174), who describes the Near Future as used only for later the same day, not also 'tomorrow'. In its 'today'-'tomorrow' function, the

---

[2] In Table 1, SP stands for the subject-indexing prefix; the underlining marks the location of the verb stem in the template. Mould and Leung differ in tonal analyses. Since a full tonal analysis of Logoori verbal inflections has not yet been completed, marking of tone is omitted here except when critical to tense distinctions.

Near Future coincides with another inflectional form and with at least two auxiliary constructions; see §7.

Unlike these other hodiernal/crastinal forms, however, Logoori texts reveal that the Near Future has an extended function as the most general of the future tenses, being applicable to any future time, near or distant.[3] One parallel case is found in the Bantu language Kikuyu (Johnson 1977: 19), where the Near Future doubles as Indefinite Future, applicable to any unspecified future time. Farther afield, the Papuan language Nungon shows a similar extended use of the Near Future form for general time (Sarvasy 2014). The Nungon Near Future is either strictly hodiernal, used only for situations that will occur between the present moment and the end of 'today', or a generic future, used for situations that could take place at any future time, hodiernal or otherwise.

The Logoori form /SP-ra-__-a/ is shown with its restricted semantics, referring to only 'later today' through 'tomorrow', in §3.1. Its extended function for general time is discussed in §3.2.

3.1 Strict Near Future /SP-ra-__-a/

Leung writes that the Near Future /SP-ra-__-a/ is "used when speaking of events or actions that are to take place later during the day" (1991: 174). As noted above, Mr. Indire considers this form to be primarily applicable to events that will take place later today or tomorrow. Regardless of whether the Near Future is strictly hodiernal, or hodiernal and crastinal, this application of /SP-ra-__-a/ may be understood as the semantically restricted one, akin to the Kikuyu or Nungon Near Future tenses. An example of the restricted function of the Logoori Near Future is in (1).

(1) Mu-gamba, **a-ra-zj-a** mʊ ɪ-skuru.
3-tomorrow 1-NF-go-FV LOC 9-school

'Tomorrow, **s/he will go** to school.' [Note that in other dialects, *mu-gamba* refers to 'morning'.]

Other forms are also used for hodiernal/crastinal reference; these include a form that appears to be a linking prefix plus the Subjunctive form, and auxiliary verb constructions. These are discussed in §7.

3.2 General Future /SP-ra-__-a/

Throughout the corpus, it is the Near Future form /SP-ra-__-a/ that occurs with application to maximally 'general' time. While the Middle and Remote Future forms may be applied to time periods extending far into the future, they have not yet been found

---

[3] It may always be speculated for a tonal language with unfinished tonal analysis that there could be a grammatical-tone formal distinction between the hodiernal/crastinal and general time functions of the /ra-/ form. This was not perceived by Mr. Indire, however; for him, the form is the same, morphologically and tonally.

to be applicable to time periods closer to deictic center. The Near Future in its general application can apply to any future time, regardless of distance from deictic center.

The first two examples of the Near Future form functioning as a general, not hodiernal/crastinal, future come from a funerary song from the Diercks corpus, sung by Ms. Linette Mbone. The song describes what people will encounter on arriving in heaven. The song is largely framed in the Near Future, and includes numerous verses similar to the following:

(2) V-oosi **va-ra-ɲor-a** ɪ-tadʒi.
    2-all   2-ɴғ-find-ғv 9-crown

    'All **will find** a crown (in heaven).' (Funeral song, ln 5)

(3) Mi-handa **dʒi-ra-v-a** mi-rahɪ.
    4-road   4-ɴғ-be-ғv 4-good

    'The roads **will be** good (in heaven).' (Funeral song, ln 7)

Here, clearly, the Near Future form does not restrict time reference to the hodiernal/crastinal period. Rather, Mr. Indire described these instances as set in "general future" time, equally applicable to many years hence, or much sooner. A person could die later today, or more than fifty years after the speech act.

The next examples also show the Near Future form used outside the restricted hodiernal/crastinal context. In another text from the Diercks corpus, Mr. Benjamin Egadwa explains how to care for a cow. Much of the text is framed in the Subjunctive, with most verbs lacking a tense prefix and bearing the Subjunctive suffix -*e*. When Mr. Egadwa does use tense-marked verbs, however, these are most often Near Future forms, as in the following examples:

(4) Zi-seendi zi-ra, **zi-ra-ku-koɲ-a** kʊ vi-indo vivj-o   vi-ra.
    10-money 10-ᴅᴇᴍ 10-ɴғ-2sɢ-help-ғv ʟᴏᴄ 8-thing 8:ʀᴇᴅ-2sɢ.ᴘᴏss 8-ᴅᴇᴍ

    'That money, **it will help you** with those things of yours.' (B. E. cow care, ln 69)

(5) Kaande, **ɪ-ra-ku-h-a**   za   ma-veere.
    again   9-ɴғ-2sɢ-give-ғv ʟɪɴᴋ 6-milk

    'Again, **it will give you** milk.' (B. E. cow care, ln 45)

(6) ɪ-v-e   ɪn-dahɪ, kɪgera **ʊ-ra-gur-iz-a**   kʊ.
    9-be-ғv 9-good   because 2sɢ-ɴғ-sell-ᴄᴀᴜs-ғv ʟᴏᴄ

    'It is good, because **you will sell** of it.' (B. E. cow care, ln 80)

(7) ʊ-ta-ret-a   kʊ ɪ-ŋombe ɪ-v-e   ɪ-mbarava ha-aŋgo daave,
    2sɢ-ɴᴇɢ-bring-ғv ʟᴏᴄ 9-cow   9-be-ғv 9-fierce   16-home not

    'Don't bring home a cow that is fierce,

    ɪ-ra-ku-ret-er-a   ɪ-ʃida.
    9-ɴғ-2sɢ-bring-ᴀᴘᴘʟɪᴄ-ғv 9-trouble

    it will bring trouble for you.' (B. E. cow care, ln 130)

The Near Future used for general time may occur negated, as:

(8) ɪ-ŋombe si **ɪ-ra-kw-em-a**   ma-veere daave.
     9-cow   NEG 9-NF-2SG-deny-FV 6-milk   not

     'The cow **will not deny you** milk.' (B. E. cow care, ln 32)

In examples (2–8), the Near Future form is employed to describe future times that are not fixed, along the lines of the Kikuyu Near/Indefinite Future (Johnson 1977: 20). The Logoori Near Future thus seems to be a good candidate for the secondary label Indefinite Future! In fact, in his Bantu-wide discussion of -ka- verbal prefixes, Botne (1999: 492) writes: "Llogoori, itself, has an indefinite future formative -ra-."

So what of the Logoori /ri-/ future, which the Mould/Leung label implies is the dedicated Indefinite Future form? The next section addresses this question.

4 The rare /ri-/ future

The /ri-/ future form — Leung and Mould's Indefinite; Nurse's Uncertain — is too rare in the corpus for any revelations here. It is noteworthy that in the same discussion in which he calls the Logoori /ra-/ form a dedicated indefinite future, Botne (1999: 491) also mentions a Logoori "near future formative" -ri-. It is unclear from his paper whether Botne means to reverse Leung's and Mould's labels. In any case, there is no evidence in the corpus or from Mr. Indire's intuitions for the /ri-/ future form being more applicable to hodiernal/crastinal times than the /ra-/ future form.

The /ri-/ future co-occurs with the Near Future form in the funerary song sung by Ms. Mbone, source of examples (2) and (3). There are only two instances of it, however, compared with 24 of the Near Future. Both instances of the /ri-/ form are negated, and the two instances occur in adjacent clauses, shown in (9) and (10):

(9) Va-ra-ger-iz-a   kw-iŋgir-a jo,
     2-NF-try-CAUS-FV 15-enter-FV 9.there

     'They will try to enter there,

     navuzwa si **va-ri-ɲar-a**   mba.
     but   NEG 2-IF-be.able-FV NEG

     but **they will not be able to**.' (Funeral song, lns 11-12)

(10) Si **va-ri-ɲor-a** kʊ vi-hanwa.
     NEG 2-IF-find-FV LOC 8-gift

     'They **will not find** any presents.' (Funeral song, ln 13)

Both of these negated instances of the /ri-/ future refer to a time frame pre-established by the Near Future. Mr. Indire further indicated that if *si va-ri-ɲor-a* lacked the negator *si*, it would sound better to use the Near Future prefix /ra-/, not /ri-/.[4] It is crosslinguistically

---

[4] When asked, Mr. Indire also expressed doubt that the /ri-/ and /ra-/ future forms differ in degree of certainty.

common for less-certain or irrealis forms to be used as negated counterparts to certain or realis forms.

Another of the few instances of the /ri-/ future is found in a song embedded in a legend told by Ms. Hellen Makungu. Here, the form occurs with positive polarity, seen in (11). A woman is quoted as saying that if anyone breaks one of her pots, she will take the culprit to an ogre.

(11) Na-a-vol-a,   mu-ndʊ u-rj-ataɲ-a   ka-rara,
      CONSEC-1-say.APPLIC-FV 1-person 1-5-break-FV 12-one

      'And she said, the person who breaks it just a little,

      n-di-mu-ʃir-a   wa   gʊ-nani.
      1SG-IF-1-take-FV 1.GEN 20-ogre

      I will take her to the ogre.' (Funeral song, lns 11-13)

As will be seen in §6, the /ri-/ future form here occurs in a similar context to the /raka-/ future form; it is unclear why /ri-/ is used here and not /raka-/. Although the function of /ri-/ is unclear from these paltry examples, its very scarcity may be of note. Perhaps its use does relate more to epistemic judgment than the more common future forms.

5 The Middle Future

The Near Future form is much more common in the corpus than the other future forms, such as the Middle Future /na-SP-___-e/. While Leung (1991: 285) writes that the Middle Future applies to "sometime between tomorrow and several years later," Mr. Indire describes the Middle Future as simply referring to any time after tomorrow. Accepting either of these, the Middle Future tense would seem to entirely encompass the reach of any Remote Future tense on a timeline. That is, either the Remote Future or Middle Future could apparently be used to refer to distant times, probably depending on discourse context, but only the Middle Future could refer to nearer times. A Remote Future (as in §6 below) then would have a reduced range within the range of the Middle Future on the timeline. The Middle Future co-occurs with the Near Future in reported speech in at least one text. It is conceivable that the Middle Future is used there for discourse-related purposes such as increased politeness.

The basic time-referencing function of the Middle Future is illustrated in (12). When Michael Diercks had not yet arrived in the Logoori area, a group of singers (led by Ms. Carolyn Chesi) recorded a praise song for him, in his absence. In the song, the singers promise to bestow flowers upon Diercks on his arrival. This bestowal is framed in the Middle Future tense:

(12) **Na-kʊ-sjaj-e,   na-kʊ-sjaj-e**   Majki w-etu   ma-uwa.
      MF-1PL-award-FV MF-1PL-award-FV Mike 1-1PL.POSS 6-flower

      '**We will award**, **we will award** our Mike flowers.' (Praise song, ln 1)

Use of the Middle Future here may be justified according to Mr. Indire's understanding of the Middle Future tense: the singers knew that Diercks would not arrive later that day, nor the next. The Middle Future allows both for his arrival and the giving of the flowers any time beyond the day after the singing act. The Middle Future may also have been used because the singers expected to bestow flowers on Diercks within, perhaps, three years.

The function of the Middle Future form is not always this easily explained, however. In another legend told by Ms. Grace Otieno, Near Future and Middle Future forms coexist within a short passage, shown in (13). A woman has just come upon a group of people. She carries wings she was given by birds, and asks the group whether, if she gives them the wings, they will be able to finish 'knitting' them.[5]

(13) Va-a-ʤib-a   **va-ra-mar-a**.
      2-RP-reply-FV 2-NF-finish-FV

      'They replied (that) **they will finish**.
      Ja-a-va-h-a   nɪ-v-i-ituŋg-a   nɪ-va-a-mar-a.
      1-RP-2-give-FV CONSEC-2-RP-?knit-FV CONSEC-2-RP-finish-FV

      She gave (the wings) to them and they knitted (them) and finished.
      Ja-a-va-vor-er-a   mw-ituŋg-aŋg-i   a-ma-vaha ga-aŋge
      1-RP-2-say-APPLIC-FV 2PL-knit-PROGR-PERF AUG-6-wing 6-1SG.POSS

      She told them, (since) you have knitted my wings,
      na-mu-m-b-e   kɪ?
      MF-2PL-1SG-give-FV what

      what **will you give me**?
      Va-a-vʊgʊr-a ɪ-mburi nɪ-va-a-mu-h-a   n-a-a-zj-a.
      2-RP-take-FV 9-goat   CONSEC-2-RP-1-give-FV CONSEC-1-RP-go-FV

      They took a goat and gave (it) to her and she went.' (G. O. woman and hen, lns 16-17)

The time period in which the knitters complete the knitting of the wings is not specified in the story. The knitters affirm that they 'will finish' using the Near Future form. This could be compatible with either hodiernal/crastinal or general time reference: perhaps the wings will be finished on the same or following day, or perhaps later. Once they finish, however, the woman then asks what they will give her using the Middle Future form, which is usually restricted to time beyond 'tomorrow'.

The story seems to imply that the giving of the goat occurs quite soon after the woman's query, although no time frames are explicitly described. If the woman actually expects an immediate gift, it is possible that she uses the Middle Future rather than the Near Future for politeness, suggesting that the giving need not take place by the morrow.[6]

---

[5] The notion of 'knitting' wings comes from the broad English translation of this text supplied by Michael Diercks's Logoori consultant in Kenya. Mr. Indire was unable to confirm the glosses for these terms.

[6] This conjecture has not been confirmed by Mr. Indire.

6 Remote Future forms

Mould (1981), Leung (1991) and Nurse (2003) accord in describing a Far Future form /SP-raka—e/ or /SP-rika-__-e/. While Mould only lists /SP-raka-__-e/ and Nurse only /SP-rika-__-e/, Leung's consultant was familiar with both forms and found them synonymous and equally acceptable (1991: 204, fn 10). It is thus possible that these are dialectal variants. Neither of these forms was initially supplied by Mr. Indire in discussing the distant future, however. Mr. Indire preferred /naa-SP-__-e/ for occurrences far in the future. This form is clearly similar to the Middle Future form /na-SP-__-e/, but with a long vowel /aa/ in the tense prefix.

6.1 The /raka-/ Remote Future

Along with Botne (1999), Nurse (2008: 85) notes that a form /-ka-/ occurs in future tense formatives in a number of Bantu languages, and that it "sometimes forms a composite marker with other morphemes, and refers predominantly to far future…in systems with multiple future reference." Few instances of the form /SP-raka-__-e/ are found in the present corpus, and no tokens at all occur of the alternative form /SP-rika-__-e/. Mr. Indire indicated that /raka-/ has a further alternative form /aaka-/. Indeed, most speakers in the corpus use only the /aaka-/ form with SPs other than 1SG, and /raka-/ only after 1SG, as [nda(:)ka].

Most instances of the /raka-/ future form occur in a legend recorded by Ms. Hellen Makungu, source of example (11). In the legend, a mother initially threatens her eight daughters with death-by-ogre if they should break any of the pots she made for fetching water. This is shown in (14), where the /raka-/ form occurs in both the protasis and apodosis of a conditional sentence.

(14)   Mu-kana mu-rara neva **a-raka-ataɲ-e** kʊ ɪ-sjoŋgo ɪ-jɪ,
        1-girl   1-one   if   1-RF-break-FV of 9-pot   9-DEM

    'One girl, if **she will break** a pot,

    n-daka-mu-ʃir-e wa   gʊ-nani.
    1SG-RF-1-take-FV 1.GEN 20-ogre

    I will take her to the ogre.' (H. M. woman and eight girls, lns 6-7)

If Botne (1999) is correct in positing that /raka-/ originally comprised the /ra-/ future prefix plus a distal /ka-/, the function of this distal /ka-/ in (14) may be epistemic rather than timeline-related. Epistemically certain events predicted to occur at unspecified future times are described using the /ra-/ future form, as in (2–8). In contrast, both the breaking of the pot and the taking to the ogre here are hypothetical future events.

One daughter does break a pot in the mother's absence. The other girls urge her to wait for the mother and admit her guilt, but she refuses, saying:

(15) A-a! Mama   a-m-bol-e   ndɪ mu-ndʊ **u-raka-ataɲ-e** kʊ ɪ-sjoŋgo
      no   mother 1-1sɢ-say.ᴀᴘᴘʟɪᴄ-ᴘᴇʀꜰ that 1-person 1-ʀꜰ-break-ꜰᴠ of 9-pot

      'Uh-uh! Mother told me that a person **who will break** a pot,

      a-raka-mu-ʃir-e wa   gʊ-nani gʊ-ra-zj-a   kʊ-mu-rj-a.
      1-ʀꜰ-1-take-ꜰᴠ   1.ɢᴇɴ 20-ogre 20-ɴꜰ-go-ꜰᴠ   15-1-eat-ꜰᴠ

      she will take her to the ogre, who will go to eat her.' (H. M. woman and eight
      girls, lns 11-12)

Here, both the breaking of the pot and the bringing to the ogre are framed with the /raka-/ form, but these are followed by the Near Future /ra-/ form for the ogre's subsequent predicted action. It is possible that the Near Future is used here because the ogre's action is expected to follow immediately on receipt of the pot-breaker.

6.2 Another Remote Future: /naa-SP-__-e/

Mr. Indire described yet another Logoori Remote Future form that is missing from extant descriptions. Nurse (2008: 86, fn c) notes that in Bantu E15, E16, and E55, a future formative /naa-/ corresponds to the /raa-/ of closely-related Bantu languages. It may be that the Logoori Remote Future /naa-SP-__-e/ described by Mr. Indire is another instance of this. Its relationship to the /raka-/ future of §6.1 may be one of dialectal variance, since the /naa-/ formative seems to be the primary one used by Mr. Indire for most distant future time reference.

Mr. Indire indicated that the temporal domain of this tense would fall within the domain of the Middle Future, but begin farther from deictic center. This form is unattested in the corpus.

(16) Na-a-kor-e ɪ-gaasi.
      ᴍꜰ-1-do-ꜰᴠ 9-work

      'S/he will do work (after tomorrow).'

(17) **Naa-a-kor-e** ɪ-gaasi.
      ʀꜰ-1-do-ꜰᴠ   9-work

      '**S/he will do** work (at a later date).'

Alternatively to Nurse's proposal, in which /naa-/ relates to a more-common Bantu form, /raa-/, this Remote Future form may be analyzed in two different ways: either the long vowel in /naa-/ represents iconic lengthening of the Middle Future tense prefix vowel to show temporal distance, or /naa-/ actually comprises the Middle Future tense prefix /na-/ followed by a prefix /a(a)-/ indicating remoteness in time. The Logoori Remote Past form, /SP-aa-__-a/, also features a prefixed long /aa-/, although after, not before, the subject prefix (cf. Mould 1981: 206; Leung 1991: 323). In this respect, this Remote Future form mirrors the Remote Past form. If it is innovated (since none of the previous sources mention it), it could represent a kind of leveling on the part of language learners.

7 Further forms used with general and near future times

Logoori texts show a number of additional forms and constructions used with both unspecified/general and near future times. These go beyond the future tense forms shown in Table 1. These forms may not stand in for specific post-crastinal times. Indeed, Nurse (2008: 85) writes that Bantu future-referencing "forms deriving from 'come' and the subjunctive tend to refer to near futures." The Logoori Near Future tense form has potential substitutes for both its general future function (the Subjunctive) and its hodiernal/crastinal future function (/ma-SP-⎵-e/ or an auxiliary construction). In contrast, the Middle Future and Remote Future tense forms have no auxiliary or other formal substitutes with similar time reference.

7.1 Subjunctive for general time

The Subjunctive form lacks a tense prefix but has final suffix *-e* instead of unmarked *-a*, Present Progressive *-aa*, or Perfective *-i*. As seen above, several of the future tense forms also employ a final suffix *-e.*

In the corpus, the Subjunctive form predominates in texts that describe general processes or give instructions/advice. The following example comes from the narrative instructions for bovine care from Mr. Egadwa.

(18) **ʊ-gɪ-h-e**   a-ma-aze... nɪ-ɪ-dook-a   saa saba,
2sg-9-give-subj aug-6-water foc-9-arrive-fv hour seven

'**You (will) give it water** … when (the time) arrives at one o'clock,

ʊ-gɪ-ker-e   a-ma-veere...
2sg-9-milk-subj aug-6-milk

you milk it...' (B. E. cow care, lns 27-29)

The Subjunctive form alone is not found in the corpus with specific future time reference; for reference to a specific time, it apparently must co-occur with auxiliary *kʊmaṇa* 'to know,' as described in §7.3.

7.2 Auxiliary construction with *kʊzja* 'go'

This auxiliary construction used to express future time uses the cross-linguistically-common device of a verb of motion to indicate relatively-near future time (Heine & Kuteva 2004: 161-163). Here, the verb *kʊzja* 'go' is inflected for Present Progressive, and followed by the lexical verb in the Infinitive form. Example (1) showed the verb *kʊzja* with full lexical meaning 'go,' while (19) shows it serving as an auxiliary indicating imminent future time.

(19) Karoono, **a-zez-a**   kʊ-nw-a   ɪ-kahava.
now/today 1-go.pres.progr-fv 15-drink-fv 9-coffee

'Now, **s/he is going** to drink coffee.'

The other auxiliary construction used frequently with future time reference in the corpus involves the verb *kʊmaɲa* 'know,' described in §7.3.

7.3 Future semantics with *kʊmaɲa* plus Subjunctive

The verb *kʊmaɲa* 'to know' is employed as auxiliary in another construction with future time reference in Logoori. Here, the lexical verb following inflected 'know' occurs in the Subjunctive form /SP-___-e/. This auxiliary construction may stand in for the Near Future tense form in either its specified hodiernal/crastinal or general time reference applications. Because the Subjunctive has modal, future-related meaning without *kʊmaɲa* (§7.1), future meaning here apparently comes from the Subjunctive, rather than from 'know,' which occurs unmarked for tense. The verb 'to know' also functions as an auxiliary with telic verbs in non-future contexts. It is unclear whether *kʊmaɲa* as auxiliary in future constructions may only occur with telic lexical verbs.

Example (20) shows the *kʊmaɲa* plus Subjunctive construction used with specified crastinal time reference:

(20) Mu-gamba, **maṇ-a**   **n-zj-e**   mʊ ɪ-skuru.
     3-tomorrow 1SG.know-FV 1SG-go-SUBJ LOC 9-school

     'Tomorrow, **I will go** to school.'

Note that the Subjunctive by itself has only general, unspecified future time reference in the corpus; *kʊmaɲa* here apparently aids in the specification of time, although *kʊmaɲa* constructions may also have general time reference.

The remaining examples here show this construction with general, unspecified time reference. In Mr. Egadwa's text on cow care, the /SP-ra-___-a/ future tense form alternates with the *kʊmaɲa* auxiliary future construction. After issuing instructions framed in the Subjunctive about keeping timely milking schedules, the speaker says:

(21) **ʊ-maṇ-a**   **ʊ-ɲor-e**   mʊ a-ma-veere sja   w-eɲ-aŋg-a.
     2SG-know-FV 2SG-find-SUBJ LOC AUG-6-milk how 2SG-desire-PROGR-FV

     '**You will find** in it milk as you desire.' (B. E. cow care ln 31)

In another procedural text, Ms. Carolyn Chesi describes how she makes the famous Kenyan cornmeal dish *vuchima.* Ms. Chesi uses the *kʊmaɲa* future time construction frequently throughout the text. The following are only a few samples:

(22) **Maṇ-a**   **n-zj-e**   ṃ-mʊreme, n-zj-e   kw-ah-a i-ri-kove,
     1SG.know-FV 1SG-go-SUBJ LOC-farm   1SG-go-SUBJ 15-pick   AUG-5-veg.sp

     '**I will go** to the farm, I will pick rikove,' (C. C. lunchtime ln 6)

(23) Gw-aaka-ɲaar-a,   **ɪ-maṇ-a**   **ɪ-dook-e**   ɪ-saa, ʃimbe saa   taano,
     3-FUTPERF-shrivel-FV 9-know-FV 9-arrive-SUBJ 9hour near   hour five

     'Once (the greens) have shriveled, **it will arrive at**, close to eleven o'clock,

saa   siita, aa-m-bek-e   kʊ   ma-higa.
hour six   CONSEC2-1SG-put-SUBJ LOC 6-stove

twelve o'clock, then I will put (it) on the stove.' (lns 13-16)

(24) N-daaka-rik-a,   **maṇ-a   n-duk-e**.
1SG-FUTPERF-put.water.on-FV 1SG.know-FV 1SG-make.cornmeal-SUBJ

'Once I've put water on, **I will make the cornmeal**.' (lns 37-38)

The verb *kʊmaṇa* 'know' only functions as an auxiliary if its subject and the subject of the following verb are coreferential. Otherwise, it must be interpreted with full lexical value, as in (25).

(25) Zja **ʊ-maṇ-a**   nɪ-ɪ-vet-a   kʊ   mu-ndʊ,
so 2SG-know-FV CONSEC-9-pass-FV LOC 1-person

'So **you know** when (the cow) passes by a person,

ɪ-vor-e   kʊ-mu-duj-a daave.
9-refrain-SUBJ 15-1-hit-FV   not

it will not hit him.' (B. E. cow care, ln 133)

The semantic difference between the *kʊmaṇa* constructions with general future time reference and the Subjunctive alone with such reference is as yet unclear. It is tempting to assume this difference is epistemic or relates to the strength of the assertion, but this is pure conjecture.

Botne (2009) describes a counterpart auxiliary construction using 'to know' in Lwitaxo (Idakho). He identifies two functions of the construction: a "generic" function, "normally/ordinarily/typically do V" and a "culminative" function, "often suggesting a consequence: 'end up Ving,' 'ultimately V,' or 'in the end'" (2009: 93). He writes that the culminative function "focuses on the event as the culmination of a series of events, while the generic does not refer to any specific event" (2009: 95).

The formal difference between the two functions identified by Botne is the presence of the 'focus' marker /n(i)/:[7]

(26)   Generic: SP$_1$-maṇ-a SP$_1$-___-a
Culminative: SP$_1$-maṇ-a n(i)-SP$_1$-___-a

---

[7] This form has been described as marking 'focus' (Dalgish 1979; Nurse 2006; Botne 2009). As Botne writes (2009, fn 5): "There are five different functions associated with the form *ni*: copula (COP), participial (PRT), sequential (SEQ) (often equivalent to 'and' or 'then'), temporal 'when', and focus (FOC)." With *nɪ*-marked verbs that are not preceded by an auxiliary, Botne's sequential use prevails in Logoori narratives. Some speakers use *nɪ-/n-* more than others. (Over-reliance on *nɪ-/n-* seems to be a characteristic of less-proficient speech, with more-proficient speakers varying their narrative devices.)

(i)   *Ma, nɪ-kʊ-tʊr-a   jo,   ma   va-ande nɪ-va-zj-a...*
then CONSEC-1PL-leave-FV 9.there then 2-other CONSEC-2-go-FV

'Then, we left there, then the others went...' (M. I. today tomorrow lns 21-22)

Both of the equivalent constructions — and at least one other related one — are present in the Logoori texts, but Botne's description of their functions may not fit the Logoori data.

At the end of one legend, Ms. Grace Otieno relates of the two protagonists:

(27) Mu-jaajɪ na   mu-kaana **va-a-maṇ-a**   **va-a-romb-a**   kɪ-serero.
     1-boy   CONJ 1-girl   2-RP-know-FV 2-RP-make-FV 7-wedding

     'The boy and the girl **made a wedding**.' (G. O. girls who wanted luck, ln 78)

If this were equivalent to Botne's description of the Idakho forms, we would expect generic meaning, with no specific event indicated. That is clearly not the case. The following lines from the same Logoori story also run counter to the Idakho analysis:

(28) Ja-a-ṇor-a   ɪ-mbwa ha-aŋgo.
     1-RP-find-FV 9-dog   16-home

     'He found the dog at home.

     Ja-a-maṇ-a   **n-a-a-vʊgʊr-a**   ɪ-mbwa jɪjo,
     1-RP-know-FV CONSEC-1-RP-take-FV 9-dog   9.DEM

     He took that dog,

     n-a-a-taaŋg-a   kʊ-m-beṇ-a   na-jo.
     CONSEC-1-RP-begin-FV 15-?1-?stay-FV CONJ-9

     and started to stay with it.' (G. O. girls who wanted luck lns 39-41)

Following Botne's description of Idakho, the presence of /n(ɪ)-/ on the verb following 'know' should indicate "culminative" function. But here, only one event, the finding of the dog, leads up to the taking of it — and surely the 'start to stay with' is more of an end result than is the 'taking'.

It seems that — in Logoori texts, at least — the greatest generalization possible about non-future *kʊmaṇa* auxiliary constructions has to do with the types of lexical verbs used. These may be considered, without exception, to be telic verbs: 'make,' 'arrive,' 'go to the farm,' 'take,' 'make cornmeal,' 'find,' 'remove from stove,' 'transform into,' among others.

7.4 Narrative form /ma-/ or /aa-/ before Subjunctive for near future times

Like other Bantu languages, Logoori is rich in 'narrative' forms that primarily function in sequences of sentences or clauses. Two of these forms are probably dialectal variants of a single form, comprising the discourse particle *ma* plus the inflected verb; this is realized as *ma-* by some speakers and *aa-* by others.

When the narrative prefix *ma-/aa-* combines with the Subjunctive inflection, Mr. Indire considers the resulting form /ma-SP-__-e/ to have similar hodiernal/crastinal application to the /ra-SP-__-a/ Near Future. Here, a narrative context is not necessary. In fact, Mr. Indire used this form rather than the dedicated Near Future in describing his plan for the next day:

(29) Ma mu-gamba **ma-m-bok-e**, **ma-n-zj-e** kʊ ɪ-gaasi j-eende.
  then 3-tomorrow NF2-1SG-wake-SUBJ NF2-1SG-go-SUBJ LOC 9-work 9-other

'Then tomorrow **I will wake up**, **then I will go** to another job.' (M. I. today, tomorrow ln 62)

Although Mr. Indire introduced this form as an actual Near Future tense, in the corpus it most commonly occurs on non-initial verbs in narrative sequences, indicating a temporal or causal connection with the actions denoted by earlier verbs, as in (30):

(30) **Aa-nz-ah-e** neende mu-tere,
  CONSEC2-1SG-uproot-SUBJ with 3-veg.sp
  aa-ṇor-e.
  CONSEC2-separate.leaves.from.stems-SUBJ

'**I will uproot (it)** together with *umutere*, **then I will separate the leaves from the stems**.' (C. C. lunchtime lns 7-8)

In one instance in the corpus, this *aa-* occurs before a verb lacking the subjunctive suffix *-e*, which clearly refers to a nonfuture event:

(31) Vj-age, kw-aat-a m-mba, a-vi-ivi.
  8-barn 1PL-put-FV LOC-house AUG-2-thief

'As for the barns,[8] we have put (them) in the houses: thieves.
  va-ku-ṇo kʊ-no, ʊ-kob-e zi-kwiiri, na va-ku-temag-aa.
  2-2SG-find LOC-DEM 2SG-beat-SUBJ 10-cries CONJ 2-2SG-chop-PRES.PROG

They find you there; cry out, and they are chopping you.
  Aa-kw-aat-a ha vʊ-riri.
  CONSEC2-1PL-put-FV LOC 14-eating

(Thus) we have put (them) in the eating area.' (Nichols & Ssennyonga 1976, barns lns 1-3)

Here, *aa-kw-aat-a* may be best analyzed as a narrative or consecutive form. It may be that pre-SP *aa-* and its variant *ma-* simply mark this narrative form, with relative time understood through other verbal inflections. Mr. Indire did not perceive a difference in meaning between a *ma-___-e* future form uttered as part of a sequence of clauses versus one uttered in isolation.

8 Conclusion

This paper has aimed to describe some of the complex usage patterns of recognized future tense markers in Logoori texts, as well as additional constructions used to indicate future

---

[8] 'Barns' here apparently refers to grain storage; rather than store food outside, it is now being stored indoors.

time. What can be gleaned from elicitation must be augmented with discourse analysis to yield a fuller picture of the functions of tense forms.

The Near Future /SP-ra-__-a/ form is the most common future form found in the corpus; it functions both in contexts where a hodiernal or crastinal future time is specified, and in contexts in which epistemically-certain events are predicted to occur at unspecified future times. Scant evidence indicates that the Middle Future form /na-SP-__-e/ may both serve to indicate post-crastinal time periods and possibly for purposes of politeness. The /SP-raka-__-e/ form is called Far Future in the literature, but in at least one text it seems to differ from the unspecified use of the Near Future primarily in epistemic certainty. In the corpus data it seems nearly interchangeable with the /ri-/ form. A second Remote Future form with a longer vowel than the Middle Future form is unattested in the corpus and is not found in previous descriptions of the language, but was produced with consistent duration contrast by Mr. Indire.

It might seem that speakers of a language with so many future-related inflections would not need to resort to periphrasis to discuss future time. But the Logoori texts corpus shows otherwise; in fact, dedicated tense inflections are supplemented by the Subjunctive, narrative forms and at least two auxiliary constructions to indicate future time.

Abbreviations

1SG, etc.	first person singular, etc.	IF	indefinite future
1, 2, 3, etc.	noun classes	LOC	locative
APPLIC	applicative	MF	middle future
AUG	augment	NEG	negative
CAUS	causative	NF, NF2	near future
CONJ	conjunction	POSS	possessive
CONSEC, CONSEC2	consecutive forms	PROG	progressive
CONT	continuous	RECP	recent past
DEM	demonstrative	RP	remote past
FUTPERF	future perfect	RF	remote future
FV	final vowel	SP	subject prefix
GEN	genitive	SUBJ	subjunctive

Acknowledgments

References

Besha, Ruth Mfumbwa. 1989. *A study of tense and aspect in Shambala*. Berlin: Dietrich Reimer.

Botne, Robert. 1999. Future and distal -ka-'s: Proto-Bantu or nascent form(s)? In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives*, 473–515. Stanford: CLSI.

Botne, Robert. 2009. The curious case of auxiliary -many'a in Lwitaxo. *Studies in African Linguistics* 38(2). 93–112.

Botne, Robert. 2013. On the organizational properties of Bantu T/A systems. Position paper. In *Bantu 5: Fifth International Conference on Bantu Languages*. http://bantu5.sciencesconf.org/conference/bantu5/pages/Bantu5_WS_2_Bantu_TA_position_paper.pdf.

Botne, Robert & Tiffany L. Kershner. 2008. Tense and cognitive space: On the organization of tense/aspect systems in Bantu languages and beyond. *Cognitive Linguistics* 19. 145–218.

Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.

Crane, Thera Marie. 2011. *Beyond time: Temporal and extra-temporal functions of tense and aspect marking in Totela, a Bantu language of Zambia*. Berkeley, CA: University of California. (Doctoral dissertation.)

Dalgish, Gerard. 1979. The syntax and semantics of the morpheme *ni* in Kivunju (Chaga). *Studies in African Linguistics* 10(1). 41–63.

Heine, Bernd & Tania Kuteva. 2004. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Johnson, Marion R. 1977. *A semantic analysis of Kikuyu tense and aspect*. Columbus, OH: The Ohio State University. (Doctoral dissertation.)

Leung, Elizabeth Woon-Yi. 1991. *The tonal phonology of Llogoori: A study of Llogoori verbs* (Working Papers of the Cornell Phonetics Laboratory 6). Ithaca: Cornell University.

Mould, Martin. 1981. Greater Luyia. In Thomas H. Hinnebusch, Derek Nurse & Martin Mould (eds.), *Studies in the classification of Eastern Bantu languages*, 181–256. Hamburg: Helmut Buske Verlag.

Nichols, Sandra & Joseph Ssennyonga. 1976. *Maragoli*. Berkeley: University of California Media Center.

Nurse, Derek. 2003. Aspect and tense in Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages,* 90–102. New York: Routledge.

Nurse, Derek. 2006. Focus in Bantu: Verbal morphology and function. *ZAS Studies in African Linguistics* 43. 189–207.

Nurse, Derek. 2008. *Tense and aspect in Bantu*. Oxford: Oxford University Press.

Sarvasy, Hannah. 2014. *A grammar of Nungon, a Papuan language of Morobe Province, Papua New Guinea.* Townsville, Queensland: James Cook University. (Doctoral dissertation.)

Sloetjes, Han & Peter Wittenberg. 2008. Annotation by category – ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008).* http://www.lat-mpi.eu/tools/elan/, accessed 2014-10-16.

Chapter 12

Quantification in Logoori

Meredith Landman
Pomona College

In this paper I examine how quantification is expressed in Logoori, a Luyia (Bantu) language spoken in western Kenya. I focus on the two universal quantifiers in Logoori, viz., *vuri* 'every' and -*oosi* 'all'. I show that these two quantifiers display a number of syntactic and semantic differences and present a compositional analysis to account for those differences. Throughout, I discuss how the Logoori patterns relate to previous cross-linguistic work on quantification, both on Bantu (Zerbian & Krifka 2008) as well as across languages more generally (Matthewson 2013).

1 Introduction

In this paper I examine how quantification is expressed in Logoori [ISO 639-3 rag], a Luyia (Bantu) language spoken in western Kenya.[1] I focus mainly on the two universal quantifiers in Logoori, namely, *vuri* 'every' (1a) and -*oosi* 'all' (1b).

(1) a. **vuri** muundu a-syeev-i.
         every 1person 1-dance-PST
         'Every person danced.'

     b. vaandu v-**oosi** va-syeev-i.
         2person 2-all 2SA-dance-PST
         'All the people danced.'

As I will show, these two quantifiers display a number of syntactic and semantic differences. Though I focus on the two universal quantifiers, I will also compare their behavior to other adnominal quantifiers in Logoori, such as -*lla* 'one, some' (2a) and -*iinge* 'many, much' (2b).

---

[1] All data are from field notes collected via elicitation interviews with Isaac K. Thomas, a native Logoori speaker in his late 30s. In the orthographic conventions I use here, *ng'* represents a velar nasal [ŋ], *ny* a palatal nasal [ɲ], *y* a palatal glide [j], and *ch* a voiceless palatal affricate [tʃ]; tone is not transcribed.

(2) a. vaandu va-**lla** va-syeev-i.
       2person 2-one SA-dance-PST
       'Some people danced.'

   b. vaandu v-**iinge** va-syeev-i.
       2person 2-many 2SA-dance-PST
       'Many people danced.'

My main aims for this paper are two. First, to present compositional analysis of (universally) quantified nominals in Logoori; and second, to add to our knowledge of how quantification is expressed across languages (cf. Matthewson 2013), and in Bantu specifically (Zerbian & Krifka 2008). As the latter authors point out in their recent survey of Bantu quantification, more work is needed in this area:

> A literature review on quantification in (whatever) Bantu languages reveals that few studies exist which touch upon quantification… The huge variety found among the Bantu languages as well as the gaps in documentation necessitate further detailed work on aspects of quantification. (Zerbian & Krifka 2008: 383, 412)

The remainder of this paper is organized as follows. In §2, I introduce the basic differences between the two universal quantifiers. In §3, I present a compositional analysis of the quantifiers, taking as a starting point work by Matthewson (2013) on quantification across languages. In §4, I consider some additional patterns that fall outside the scope of the proposed analysis; and in §5, I conclude the paper and articulate some questions for further research.

2 Basic differences between the two universal quantifiers

2.1 The main semantic difference: Distributivity

The main semantic difference between *vuri* and *-oosi* regards distributivity (see Gil 1995; Vendler 1962; among others). While *-oosi* is non-distributive (i.e., it permits distributive or collective interpretations), *vuri* is necessarily distributive. Consider, e.g., (3), which is ambiguous. On its distributive reading, (3) is true just in case each person individually lifted a car. On its collective reading, (3) is true just in case all the people together lifted a car.

(3) vaandu v-oosi va-geeng-i mudoga.
     2person 2-all   2SA-lift-PST car
     'All the people lifted a car.' ✓distributive, ✓collective

Vuri, in contrast, only permits a distributive reading:

(4) vuri   muundu a-geeng-i   mudoga.
     every 1-person 1SA-lift-PST car
     'Every person lifted a car.' ✓distributive, ×collective

Accordingly, when a collective reading is forced, for example by adding the adverb *halla* 'together' as in (5), *-oosi* is grammatical (5a) while *vuri* is not (5b).

(5)   a.   vaandu v-oosi va-geeng-i mudoga **halla**.
           2person 2-all   2SA-lift-PST car   together
           'All the people lifted a car together.'

      b.   *vuri muundu a-geeng-i   mudoga **halla**.
           every 1-person 1SA-lift-PST car   together

Similarly, inherently collective predicates such as *kuvugaana* 'to gather, meet' may combine with *-oosi* (6a) but not with *vuri* (6b).

(6)   a.   vaandu v-oosi va-vugaan-i.
           2person 2-all   2SA-gather-PST
           'All the people gathered.'

      b.   *vuri muundu a-vugaan-i.
           every 1-person 1SA-gather-PST

This semantic difference (i.e., distributivity) is also apparent in the types of nominals each quantifier may combine with. As is typical for a distributive universal quantifier (cf. Gil 1995), *vuri* most naturally combines with singular count nouns (7a). If *vuri* combines with a plural (7b) or mass noun (7c), individuated readings result, e.g., groups of books and bottles of water. In contrast, *-oosi* may naturally combine with singular count nouns (8a), plural count nouns (8b), or mass nouns (8c).

(7)   a.   SG count
           soom-i   vuri ki-tabu.
           1SG.SA.read-PST every 7-book
           'I read every book.'

      b.   PL count
           soom-i   vuri vi-tabu.
           1SG.SA.read-PST every 8-book
           'I read every group of books.'

      c.   MASS
           ngur-i   vuri ma-zi.
           1SG.SA.buy-PST every 6-water
           'I bought every (bottle of) water.'

(8)   a.   SG count
           soom-i   ki-tabu ch-oosi.
           1SG.SA.read-PST 7-book 7-all
           'I read all of the book.' or 'I read the whole book.'

b. PL count

soom-i vi-tabu vy-oosi.

1SG.SA.read-PST 8-book 8-all

'I read all the books.'

c. MASS

ngur-i ma-zi g-oosi.

1SG.SA.buy-PST 6-water 6-all

'I bought all the water.'

2.2 Syntactic and morphological differences

Vuri and *-oosi* also display a number of syntactic and morphological differences. I observe five here.

First, *-oosi* is post-nominal (9a-b), while *vuri* is pre-nominal (10a-b).

(9) a. vaandu v-oosi va-syeev-i.

2person 2-all 2SA-dance-PST

'All the people danced.'

b. *v-oosi vaandu va-syeev-i.

2-all 2person 2SA-dance-PST

(10) a. vuri muundu a-syeev-i.

every 1person 1SA-dance-PST

'Everyone danced.'

b. *muundu vuri a-syeev-i.

1person every 1SA-dance-PST

In this respect, *-oosi* patterns with all other adnominal modifiers (such as adjectives, numerals, possessives, demonstratives, and relative clauses). These also canonically appear post-nominally (11a-b).

(11) a. vaandu {yavo/va-lla/va-vere/va-nene} va-syeevi.

2person {2those/2-one/2-two/2-important} 2SA-dance-PST

'{Those/some/two/important} people danced.'

b. *{yavo/va-lla/va-vere/va-nene} vaandu va-syeevi.

{2those/2-one/2-two/2-important} 2person 2SA-dance-PST

A second syntactic difference regards co-occurrence with a pronoun. While *-oosi* may co-occur with a pronoun (12a), *vuri* may not (12b). Here too, *-oosi* patterns with other adnominal modifiers, which also may co-occur with a pronoun (12c).

(12) a. kunyi v-oosi ku-syeev-i.
      we   2-all   1PL.SA-dance-PST
      'We all danced.'

  b. *vuri kunyi {a-syeev-i   / ku-syeevi}.
      every we   {1SA-dance-PST / 1PL.SA-dance-PST}
      (Intended interpretation: 'We each danced.')

  c. kunyi {va-lla/va-vere/va-nene/v-iinge}   ku-syeev-i.
      we   {2-one/2-two/2-important/2-many} 1PL.SA-dance-PST
      'We {some/two/important/many} danced.'

A third difference regards co-occurrence with a null head noun. While *-oosi* may appear on its own, i.e., with a null head noun (13a), *vuri* may not (13b).[2] Here, too, *-oosi* patterns with all other adnominal expressions, which may also appear on their own (13c).

(13) a. v-oosi va-syeev-i.
      2-all   2SA-dance-PST
      'All danced.'

  b. *vuri a-syeev-i.
      every 1SA-dance-PST
      (Intended interpretation: 'Everyone danced.')

  c. {va-lla/va-vere/va-nene/v-iinge}   va-syeev-i.
      {2-one/2-two/2-important/2-many} 1PL.SA-dance-PST
      '{Some/two/important/many} danced.'

Finally, *-oosi* must agree in noun class with the head noun, as the paradigm in Table 1 shows. *Vuri*, in contrast, is morphologically invariant.

Table 1: Noun class agreement paradigm for *-oosi*

SINGULAR	PLURAL
1 *woosi*	2 *voosi*
3 *gwoosi*	4 *joosi*
5 *rioosi*	6 *goosi*
7 *choosi*	8 *vyoosi*
9 *yoosi*	10 *zyoosi*
11 *ruwoosi*	12 *koosi*
13 *twoosi*	14 *vwoosi*
	10 *gwoosi*

---

[2] Examples (13a) and (13c) can only be used when it is clear from the context what the head noun refers to, e.g., in answer to a question like 'How many people danced?'

In this respect, too, *-oosi* behaves like all other adnominal modifiers, which also must agree with the head noun.[3]

2.3 Summary of differences between the two universal quantifiers

Table 2 provides a summary of the differences between *-oosi* and *vuri*. In brief, their semantic properties accord with familiar differences between non-distributive and distributive quantifiers. Syntactically speaking, we see a divide that will factor into the analysis developed below.

Table 2: Summary of differences between *-oosi* and *vuri*

	oosi	*vuri*
Distributive only	yes	no
Combines with SG count	yes	no
Combines with PL count	yes	if individuated
Combines with MASS	yes	if individuated
Post-nominal	yes	no
Co-occurs with a pronoun	yes	no
Occurs on its own	yes	no
Agrees in noun class	yes	no

3 Analysis

In this section, I present a compositional analysis of *vuri* and *-oosi*. I take as a starting point the cross-linguistic generalizations for different types of universal quantifiers observed by Matthewson (2013).

3.1 Cross-linguistic generalizations for universal quantifiers

Matthewson (2013) presents a preliminary typology of quantifiers. She looks at 37 languages from 25 different families and finds that while there is variation in the syntactic behavior of different quantifiers, the syntax/semantics correspondence is not random. Specifically, she reports the following two tendencies for universal quantifiers. First, she observes that distributive universal quantifiers such as English *every* tend to "combine directly with NP, while other quantifiers do not." (Matthewson 2013: 36). That is, distributive quantifiers tend to be determiner quantifiers (henceforth D-quantifiers) (as in

---

[3] However, *-oosi* displays the same agreement morphology as demonstratives, rather than adjectives. This sets *-oosi* apart from the other two Logoori quantifiers, *-lla* 'one, some' and *-iinge* 'many', which do agree like adjectives. A similar pattern is observed by Krifka (1995) and Zerbian & Krifka (2008) for Swahili *-ote* 'all'.

Barwise & Cooper 1981). Syntactically, a D-quantifier heads a DP and combines directly with NP (see e.g. Heim & Kratzer 1998: 146):

(14) $[_{DP}[_{D}every][_{NP}person]]$

Semantically, a D-quantifier combines with an NP predicate, type *<e, t>*, to form a generalized quantifier, type *«e, t>, t>* (Montague 1973; Barwise & Cooper 1981). Accordingly, as a universal quantifier, *every* can be assigned the lexical denotation in (15):

(15) $[\![every]\!] = [\lambda f_{<e,t>}.[\lambda g_{<e,t>}.\forall x[f(x) \rightarrow g(x)]]]$

This denotation would yield a distributive interpretation for *every*, stipulating that quantification is over atomic individuals.[4]

Matthewson's second generalization regards universal quantifiers translated as 'all'. These quantifiers, she observes, tend to combine with a full DP. For example, English *all* can be analyzed syntactically as a Q (henceforth, a Q-quantifier), which combines with a full DP to form a QP (Matthewson 2001):

(16) $[_{QP}[_{Q}all][_{DP}the\ people]]$

Semantically, *all* combines with an individual-denoting DP (such as a definite plural), type *e*, to produce a generalized quantifier, type *«e, t>, t>*. I adopt the formalism of Zimmermann (2014) here, which is based on Matthewson (2001):

(17) $[\![all]\!] = [\lambda y_{e}.[\lambda g_{<e,t>}.\forall x[x \leq y \rightarrow g(x)]]]$

Because *all* quantifies over subparts ($x \leq y$) of the individual denoted by DP, distributive and collective interpretations are both possible. In the case that the subparts are atomic, a distributive interpretation results, and in the case that there is only one subpart (i.e., $x = y$), a collective interpretation results.

In the next subsection, I look at whether Matthewson's generalizations hold for the two Logoori universal quantifiers. As Zimmermann (2014) points out, African languages are under-represented in Matthewson's survey, representing just four of the thirty-seven languages: Igbo (Igboid), Koromfe (Gur), Fongbe (Kwa), and Xhosa (Bantu). Zimmermann (2014) additionally supports Matthewson's generalizations with data from the West African languages Hausa (Zimmermann 2013) and Wolof.

3.2 Do Matthewson's generalizations hold for the Logoori universal quantifiers?

3.2.1 *Vuri* as a D-quantifier

The D-quantifier analysis can naturally be extended to *vuri*. By this account, *vuri* would have the syntax in (18) and the semantics in (19).

---

[4] Distributivity may alternatively come from another source; this is not crucial to my analysis.

(18) $[_{DP}[_D vuri][_{NP} muundu]]$

(19) $\llbracket vuri \rrbracket = [\lambda f_{<e,t>}.[\lambda g_{<e,t>}.\forall x[f(x) \rightarrow g(x)]]]$

This analysis fares well with the properties observed for *vuri* above (summarized in Table 2). That *vuri* is distributive is accounted for, again stipulating that quantification is over atomic individuals in (19). That *vuri* most naturally combines with singular count nouns would be expected, assuming that singular count nouns denote properties of atomic individuals. In the case that *vuri* combines with plurals or mass nouns (which, under standard assumptions, do not denote atomic individuals, see e.g. Link 1983), semantic coercion would yield individuated readings. That *vuri* is necessarily pre-nominal is expected, assuming that DP is head-initial, as phrasal categories in Logoori generally are. That *vuri* may not co-occur with a pronoun would be accounted for if pronouns occupy D, i.e., *vuri* and pronouns occupy the same position. Independent evidence that Logoori pronouns do occupy D is provided by examples like (20), in which pronouns may co-occur with an overt head noun (see Postal 1996, among others, for relevant arguments that such co-occurring pronouns are in D).

(20) kunyi vaana ku-syeev-i.
      we   2child 1PL.SA-dance-PST
      'We children danced.'

This leaves just two properties of *vuri* unaccounted for: (a) that *vuri* may not occur on its own; and (b) that it agrees with the head noun. However, neither of these properties provides evidence against the D-quantifier analysis, either; they are consistent with it, though unaccounted for. Thus, on the whole the D-quantifier account of *vuri* is a good fit.

3.2.2 -*Oosi* as a Q-quantifier

The status of *-oosi* is most interesting here given Matthewson's generalizations, as it is less clear that it behaves like a Q-quantifier. By the Q-quantifier account, *-oosi*, like English *all*, would have the syntax in (21) and the semantics in (22).

(21) $[_{QP}[_{DP} vaandu][_Q voosi]]$

(22) $\llbracket -oosi \rrbracket = [\lambda y_e.[\lambda g_{<e,t>}.\forall x[x \leq y \rightarrow g(x)]]]$

For the most part, a Q-quantifier account is consistent with the properties summarized for *-oosi* in Table 2. That *-oosi* allows for distributive or collective interpretations, and combines with singular, plural, or mass nouns would be accounted for, given its lexical denotation in (22). That *-oosi* may co-occur with a pronoun would be accounted for if pronouns occupy D, since the two would appear in distinct positions. That *-oosi* may appear alone, without the head noun, is also expected, if the head noun can be null. Finally, that *-oosi* agrees in noun class with the head noun would follow if Q agrees. What would not be expected on this account is that QP would be head-final, since Logoori is

otherwise head-initial; this is reason to consider an alternative, and arguably simpler, account. Such an account is detailed in the next section.

3.2.3 An alternative: *-oosi* as a DP-internal modifier

A clear alternative to analyzing *-oosi* as a Q-quantifier is to analyze it instead as a DP-internal adnominal modifier, given the range of properties that *-oosi* shares with all other adnominal modifiers. As a DP-internal modifier, *-oosi* would have the syntax in (23).[5]

(23) $[_{DP}[_{NP}\textit{vaandu}][_{AP}\textit{voosi}]]$

There are, I believe, several points in favor of a DP-internal syntactic analysis for Logoori *-oosi*. First, *-oosi* has the same syntactic distribution as all other adnominal expressions, as observed above. Other adnominal modifiers also are post-nominal, may co-occur with a pronoun, may appear on their own, agree in noun class with the head noun, and combine with singular count, plural count, or mass nouns. All of these properties are consistent with the structure in (23).

A second point in favor of treating *-oosi* as internal to the DP is illustrated by (24), which shows that *vuri* and *-oosi* may actually co-occur within the same nominal phrase. Here, *-oosi* appears to be interpreted within the scope of *vuri*, thus also suggesting that it is positioned within the DP (at least on this interpretation).

(24) soom-i   vuri   ki-tabu ch-oosi.
      1SG.SA.read-PST every 7-book 7-all

      'I read every whole book.'

A third indication that *-oosi* is positioned within DP is that it may precede DP-internal modifiers, such as adjectives or numerals. Although *-oosi* can follow possessives, demonstratives, adjectives, and numerals (25a), it may also precede all of them (25b). This suggests that *-oosi* is internal to DP (or at the very least can be).

(25) a. vaana   {vaange/yavo/va-nene/vya   Chazima/va-vere} v-oosi
          2children {2my/2those/2-important/2of Chazima/2-two}   2-all
          va-gon-aa.
          2SA-sleep-PRS

          'All {my/those/important/of Chazima's/two} children are sleeping.'

      b. vaana   v-oosi {vaange/yavo/va-nene/vya   Chazima/va-vere}
          2children 2-all   {2my/2those/2-important/2of Chazima/2-two}
          va-gon-aa.
          2SA-sleep-PRS

          'All {my/those/important/of Chazima's/two} children are sleeping.'

---

[5] This, in fact, is what Zerbian & Krifka (2008) suggest for Swahili *-ote* 'all, any'. Moreover, they report that "Bantu languages have few genuine quantifiers. Rather, these languages display a range of adnominal modification with quantitative interpretation." (p. 401)

Summarizing, given that *-oosi* (a) behaves syntactically just like all other adnominal modifiers, and (b) may co-occur with *vuri*, the simplest analysis of *-oosi* would be to treat it as a DP-internal modifier (23).

Positioning *-oosi* internal to the DP raises the question, however, of how best to analyze it semantically. As sister to NP, *-oosi* is expected to combine with an NP predicate, type *<e, t>*. I see two possibilities for a compositional analysis here. The first is suggested by Zerbian & Krifka (2008) for Swahili *-ote* 'all, any', which they propose "can be analyzed as constructing the sum individual of all the entities that fall under the noun it applies to (cf. Link 1983)" (p. 401). *-Oosi* might also be analyzed as mapping predicates to sum individuals. Alternatively, Brisson's (1998; 2003) account of English *all* could be extended to Logoori *-oosi*. Brisson argues that English *all* is not a quantifier, but rather restricts the domain of a covert distributive operator on the VP. Though Brisson analyzes English *all* syntactically as a DP adjunct, her semantics could be extended to Logoori *-oosi*.

3.2.4 Summary of the analysis

This subsection summarizes the analysis. *Vuri* is a D-quantifier. This fits Matthewson's generalization for distributive quantifiers. It is also consistent with Zerbian & Krifka's observation that in Bantu, "…marked formatives are used for the expression of the universal quantifier 'every'" (p. 401), as *vuri* is the only quantificational expression in Logoori that is pre-nominal and does not display noun class agreement.

-Oosi, in contrast, is a DP-internal modifier. Its semantics can be modeled either as mapping sets to sum individuals (as Zerbian & Krifka 2008 suggest for Swahili *-ote* 'all'), or as a domain restrictor (Brisson 1998; 2003).

The syntactic and semantic properties of *-oosi* are interesting given Matthewson's cross-linguistic generalizations for universal quantifiers translated as 'all'. Matthewson presents her generalizations as tendencies, and not absolutes, but it is interesting that *-oosi* does not seem to fit the observed Q-quantifier pattern for 'all'-type quantifiers.

4 Other patterns regarding *vuri* and *-oosi*

In this section I document several other patterns regarding *vuri* and *-oosi* in an effort to lay the groundwork for future research on quantification in Logoori. In §4.1 I look at how the two universal quantifiers interact scopally with other quantificational nominals (albeit preliminarily, since judgments for scope are difficult to obtain in fieldwork contexts). In §4.2 I look at how the universal quantifiers interact scopally with negation (again, preliminarily). In §4.3 I observe a range of additional interpretations available for *-oosi*, beyond just 'all', which are not accounted for by the proposed analysis.

4.1 Scope

In this subsection I look at how *vuri* and *-oosi* interact scopally with other nominal quantifiers. The aim is to more comprehensively understand the semantic properties of each quantifier.

Both *vuri* and *-oosi* interact scopally with bare nouns. Consider (26a), for example, in which the bare noun *ridisha* 'window' is subject and *vuri murumu* 'every room' is object. (26a) is scopally ambiguous. On the surface scope reading, *ridisha* 'window' scopes over *vuri murumu* 'every room'. In this case the sentence is true just in case there is one particular window that is in every room (the pragmatically odd reading here). On the inverse scope reading, *vuri murumu* 'every room' scopes above *ridisha* 'window'. In this case the sentence is true just in case every room has a (potentially different) window. The same ambiguity is available for *-oosi*, as (26b) illustrates.

(26)   a.   ri-dirisha ri-vey-e   vuri   mu-rumu.
           5-window 5-have-PRS every 3-room

         i. 'There is one particular window that is in every room.'
         ii. 'Every room has a potentially different window.'

      b.   ri-dirisha ri-vey-e   mu   zi-rumu zy-oosi.
           5-window 5-have-PRS with 3-room   3-all

         i. 'There is one particular window that is in all rooms.'
         ii. 'All rooms have a potentially different window.'

Vuri and *-oosi* also interact scopally with nomimals modified by *-lla* 'one, some'. For example, (27a) is scopally ambiguous. On the surface scope reading, *ridisha llara* 'one window' scopes over *vuri murumu* 'every room'. In this case, the sentence is true just in case there is one particular window that is in every room. On the inverse scope reading, *vuri murumu* 'every room' scopes above *ridisha llara* 'one window', so that the sentence is true just in case every room has one (potentially different) window. The same ambiguity is again available for *-oosi*, as (27b) illustrates.

(27)   a.   ri-dirisha lla-ra ri-vey-e   vuri   mu-rumu.
           5-window 5-one 5-have-PRS every 3-room

         i. 'There is one particular window that is in every room.'
         ii. 'Every room has one potentially different window.'

      b.   ri-dirisha lla-ra ri-vey-e   mu   zi-rumu zy-oosi.
           5-window 5-one 5-have-PRS with 3-room   3-all

         i. 'There is one particular window that is in all rooms.'
         ii. 'All rooms have one potentially different window.'

Summarizing this subsection, both *vuri* and *-oosi* interact scopally with bare nominals and nominals modified by *-lla* 'one, some' (i.e., existentially quantified nominals). In particular, both *vuri* and *-oosi* permit inverse scope interpretations with respect to these nominals.

4.2 Negation

In this section I look at how *vuri* and *-oosi* interact with negation. Like many other Bantu languages (Zerbian & Krifka 2008), there is no counterpart to the negative English determiner *no* in Logoori. There are, instead, different ways of expressing propositions such as 'No one danced.' One option often volunteered by my consultant is to use *-oosi* in combination with the clausal negation *mba* 'NEG' and the morpheme *ku*:[6]

(28) muundu woosi a-syeev-i   ku mba.
       1person 1all   3SA-dance-PST KU NEG
       'No one danced.'

Vuri and *-oosi* behave differently in negated sentences. *Vuri* may occur as subject of a negated sentence (29a), in which case it must scope above negation. It is judged ungrammatical as object (29b).

(29) a. vuri   muundu a-nyar-a   ku mba.
          every 1person 1SA-mess.up-PST KU NEG
          'Every person did not mess up.'

      b. *ya-yaanz-a muundu vuri   mba.
          1SA-like-PRS 1person every NEG

In contrast, *-oosi* is interpreted as an existential (negative polarity item) in negated sentences, whether subject (30a) or object (30b).

(30) a. muundu w-oosi a-nyar-a   ku mba.
          1person 2-all 1SA-mess.up-PST KU NEG
          'No one messed up.'

      b. ya-yaanz-a vaandu v-oosi mba.
          1SA-like-PRS 2person 2-all NEG
          'He doesn't like anyone.'

4.3 A range of interpretations for *-oosi*

Though I have focused on the 'all' interpretation of *-oosi*, there are a number of other interpretations available for *-oosi* in Logoori. I review these briefly here.

First, as observed earlier, *-oosi* can mean 'whole' when it modifies a singular count noun:

(31) a-syoom-i   ki-tabu ch-oosi.
      1SA-read-PST 7-book 7-all
      'She read the whole book.'

---

[6] I have glossed *ku* here as KU because I am unsure of its semantics; see, however, Bowler & Gluckman (2015) for an account of the semantics of *ku*.

Second, as also observed above, *-oosi* is interpreted as an existential (negative polarity item) in a negated sentence:

(32) muundu w-oosi a-nyar-a   ku mba.
      1person 2-all   1SA-mess.up-PST KU NEG
      'No one messed up.'

(33) ya-yaanz-a   vaandu   v-oosi mba.
      1SA-like-PRS 2person 2-all   NEG
      'He doesn't like anyone.'

Finally, *-oosi* permits free choice interpretations in intensional or modal contexts (34).

(34) muundu w-oosi a-nyar-a   ku-syeev-a.
      1person 1-any   1SA-can-PRS INF-dance-PRS
      'Anyone can dance.'

It is possible that a single semantic analysis of *-oosi* accounts for all of its possible interpretations; I leave this issue for future research.

5 Conclusion

In this paper, I have documented and analyzed universally quantified nominals in Logoori. Specifically, I have analyzed *vuri* 'every' as a D-quantifier, and *-oosi* 'all' as a DP-internal modifier. More broadly, the study has added to our knowledge of how quantification is expressed in Bantu, as well as how the Logoori patterns relate to previous cross-linguistic work on quantification, both on Bantu (Zerbian & Krifka 2008) and across languages more generally (Matthewson 2013). The study has articulated the following questions for future research: (i) what exactly are, and what accounts for, scope interactions among Logoori quantifiers; and (ii) can a unified account of the range of interpretations available for *-oosi* be achieved?

Acknowledgements

Abbreviations

Numerals indicate Bantu noun classes.

COMP	complementizer	PST	past
NEG	negation	SA	subject agreement

References

Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics & Philosophy* 4. 159–219.

Bowler, Margit & John Gluckman. 2015. Existential quantification in Logoori: Distribution and semantics of *ku*. (Paper presented at the World Congress of African Linguistics 8, University of Kyoto, 21-24 August 2015.)

Brisson, Christine. 1998. *Distributivity, maximality, and floating quantifiers*. New Brunswick, NJ: Rutgers University. (Doctoral dissertation.)

Brisson, Christine. 2003. Plurals, all, and the nonuniformity of collective predication. *Linguistics and Philosophy* 26. 129–184.

Gil, David. 1995. Universal quantifiers and distributivity. In Emmon Bach, Angelika Kratzer & Barbara H. Partee (eds.), *Quantification in natural languages,* 321–362. Dordrecht: Kluwer.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Krifka, Manfred. 1995. Swahili. In Jaochim Jacobs (ed.), *Syntax: An international handbook of contemporary research 2*, 1397–1418. Berlin: Mouton de Gruyter.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretic approach. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use, and interpretation of language*, 302–323. Berlin: Walter de Gruyter.

Matthewson, Lisa. 2001. Quantification and the nature of cross-linguistic variation. *Natural Language Semantics* 9. 145–189.

Matthewson, Lisa. 2013. Strategies of quantification in St'átimcets and the rest of the world. In Kook-Hee Gil, Stephen Harlow & George Tsoulas (eds.), *Strategies of quantification,* 15–38. Oxford: Oxford University Press.

Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In Jaakko Hintikka, Patrick Suppes & J. M.E. Moravcsik (eds.), *Approaches to natural language,* 221–242. Dordrecht: Reidel Publishing Company.

Postal, Paul. 1996. On so-called 'pronouns' in English. In Francis Dinneen (ed.), *Nineteenth monograph on languages and linguistics*, 177–206. Washington, D.C.: Georgetown University Press.

Vendler, Zeno. 1962. Each and every, all and any. *Mind* 71. 145–60.

Zerbian, Sabine & Manfred Krifka. 2008. Quantification across Bantu languages. In Lisa Matthewson (ed.), *Quantification: A cross-linguistic perspective*, 384–414. Bingley: Emerald Group Publishing Limited.

Zimmermann, Malte. 2013. Strategies of quantification in Hausa (Chadic). In Kook-Hee Gil, Stephen Harlow & George Tsoulas (eds.), *Strategies of quantification,* 39–59. Oxford: Oxford University Press.

Zimmermann, Malte. 2014. Universal and existential quantification in Chadic and beyond. Plenary talk presented at the Annual Conference on African Languages 45, University of Kansas, 17-19 April 2014.

Chapter 13

The syntactic structure of negation in Ndebele

Ross Burkholder

University of Chicago

This paper examines the structure of negation in Ndebele, a Bantu language of Zimbabwe. The paper argues for the presence of a null verbal element in some types of Ndebele clauses. In addition, the claim is made that the two pieces of Ndebele bipartite negation are realizations of two different nodes, a higher NegP node and a lower TP node which is sensitive to polarity. Two primary pieces of evidence support this hypothesis. First, adopting such a structure allows us to treat the negative prefix *-nga-* and the negative suffix *-anga-* as a single morpheme, contrary to current proposals (Buell 2004; 2005; Khumalo 1981; 1982; Sibanda 2004). Second, introducing a null verbal element where proposed provides an explanation for previously unaccounted for data in the realm of adjectival predicates.

1 Introduction

This paper closely examines the structure of negation in Ndebele,[1] a Bantu language spoken in Zimbabwe. The paper argues for the presence of a null verbal element which occurs in some types of Ndebele clauses. In addition, the claim is made that the two pieces of Ndebele bipartite negation are realizations of two different nodes, a higher NegP node and a lower TP node which is sensitive to polarity. There are two primary pieces of evidence in favor of this hypothesis. The first is that adopting such a structure allows us to treat the negative prefix *-nga-* and the negative suffix *-anga-* as a single morpheme, contrary to current proposals (Buell 2004; 2005;[2] Khumalo 1981; Khumalo 1982; Sibanda 2004). A second piece of evidence is that introducing a null verbal element where proposed provides an explanation for previously unaccounted for data in the realm of adjectival predicates. §2 will introduce necessary background information on the language. §3 will introduce our proposed clausal structure as it relates to the negation of simple tense verbal predicates. In addition, §3 will show how the distribution of negative morphemes in Ndebele can be accounted for via agreement with a series of binary features. §4 will show how null verbal projections can be utilized to deal with both

---

[1] Guthrie Code: S.44

[2] Buell (2004; 2005) make this argument for Zulu, an extremely closely related language.

positive and negative adjectival predicates. §5 extends the argument to imperatives. §6 concludes the paper.

2 Background

Negation in Ndebele shows a bipartite distribution, as demonstrated in (1).

(1) **ka**-ngi-pheg-**i**.
     NEG-1.SM-cook-PRS.NEG

     'I do not cook.'

For this reason it will make sense to refer to negative affixes as either prefixal or suffixal, depending on their linear surface relation to the verb root. Thus it can be seen from (1) that *ka-* is a negative prefix, while *i-* is a negative suffix.[3] The negative affix that we will be primarily describing in this paper, however, seems to behave differently from the two demonstrated in (1) in that it can appear as either a prefix or a suffix.

(2) a. isi-lwane be-si-**nga**-nhle.
         7-lion   COP-7.SM-PST.NEG-pretty

         'The lion wasn't pretty.'

     b. um-fana ka-khal-a-**nga**.
         1-boy   NEG-cry-EP-PST.NEG

         'The boy wasn't crying.'

The claim that the two negative morphemes in bold font in example (2) are the same morpheme is particular to this paper. In previous analyses these two morphemes are listed separately, with *nga-* being listed as a negative prefix and *-anga* being listed as a suffix (Sibanda 2004; Khumalo 1981; Khumalo 1982).[4] The goal of this paper will be to show that not only is it possible to analyze the two morphemes in (2) as a single morpheme, but to show that in addition to being a simpler analysis, this analysis is also able to deal with previously unaccounted for data.

3 Analysis of simple tense verbal predicates

3.1 Agree

This analysis makes use of a system of Agree loosely based on the one described in Zeijlstra (2010). In his paper, Zeijlsta agrees with Bošković (2007) that Agreement is unidirectional. Unlike Bošković, however, Zeijlstra argues that this direction is upwards

---

[3] The negative affix *ka-* can appear with or without the initial /k/ sound depending on context; thus, it is referred to as *ka-* throughout.

[4] In addition, the separation of suffixal *-anga* into an epenthetic vowel *-a-* and suffix *-nga* is unique to this paper.

rather than downwards. Zeijlstra argues that one of the advantages of this type of system is that it can account for patterns of negative concord by allowing several lower nodes to Agree with a single higher node without constituting a violation of relativized minimality (Rizzi 1989). It is this aspect of Zeijlstra's system which will be particularly useful to the analysis of negation in this paper. In this paper, a middle ground is struck which preserves the heart of the Agreement system proposed by Zeijlstra and Bošković, but which loosens the directionality requirement to one direction for a single instance of Agree, in general allowing both upward and downward agreement.

Utilizing this analysis of Agree, the realizations of the polarity node can be determined by a series of binary features (Polarity, Past, Verbal) while the realization of tense nodes can be captured by a trinary feature (Tense) and two binary features (Neg, Verbal).

(3) **Agree**: α can Agree with β iff:
   a. α carries at least one uninterpretable feature and β carries a matching interpretable feature.
   b. β c-commands α OR α c-commands β.
   c. β is the closest goal to α.

As mentioned above, Ndebele has a bipartite system of negation. The claim made in this paper is that the higher of these two nodes is a true polarity projection, while the lower node is a tense node which is sensitive to polarity. The realizations for the polarity node, which is sensitive to the tense of the clause and whether the clause is verbal or nonverbal, are given below. In the tables below, the different realizations are given in the first row, and the featural distribution(s) which produce each of these realizations are given in the subsequent rows. Thus each column represents a featural distribution (pluses and minuses) whose morphological realization is the first row in that column. In Table 1, for example, the realization of the morpheme is null (shown in the first row) regardless of the featural distribution (the arrangement of pluses and minuses in each column beneath). This is not the case for the realizations of negative polarity shown in Table 2. In Table 2 there are two possible realizations depending of the featural distribution. Tables for the tense nodes will be given in the relevant sections.

Table 1: Realizations of Positive Polarity

	$\emptyset$-			
Past	+	+	-	-
Verbal	+	-	+	-

3.2 Present tense

This section provides an introduction to the clausal structure proposed in this paper by showing how it accounts for simple present tense sentences such as those given in (4).

Table 2: Realizations of Negative Polarity

	∅-		(k)a-	
Past	+	+	-	-
Verbal	-	+	+	-

(4) a. ngi-pheg-a.
         1.SM-cook-PRS

         'I cook.'

     b. **ka**-ngi-pheg-**i**.
         NEG-1.SM-cook-PRS.NEG

         'I do not cook.'

The realization of present tense morphemes is given in Table 3.

Table 3: Realizations of Present Tense

	-i	-a	-∅	
Neg	+	-	+	-
Verbal	+	+	-	-

In other Bantu studies, the morpheme *-a* found in (4a) is often glossed as a final vowel (Khumalo 1981; Khumalo 1982; Sibanda 2004). In the analysis presented in this paper, however, it will be treated as a realization of the tense node, as demonstrated both in Table 3 and the tree structures given in (5–8). Due to the abundance of Agreement and movement in the examples given, four tree structures are given, each of which highlights particular aspects of the tree. In (5–8), dotted lines are used to show Agreement, while solid lines are used to show overt movement. Arrows are used to show the direction of both movement and Agreement; thus a dotted line with arrows pointing in both directions shows Agreement between two nodes in both directions. Following Chomsky (2001), this analysis makes use of a distinction between interpretable and uninterpretable features. Uninterpretable features, demarcated by a 'u' before the feature name, must be checked in order for the utterance to be grammatical. Uninterpretable features are checked when they Agree with a matching interpretable feature (demarcated by an 'i' preceding the feature name). Strikethrough is used in (6) and (7) to show that uninterpretable features are checked. (5) gives the underlying structure of the tree before Agree or movement have taken place. (6) shows how the proposed system of Agree works to successfully check all uninterpretable features, while (7) demonstrates movement of the verb from the V(erb) to the T(ense) node (V-T movement). (8) is a simplified tree structure representing (4a). This structure exemplifies the kind of simplified tree structure

that will be utilized throughout the remainder of this paper where only the dotted lines of agreement are shown, and uninterpretable features do not use strikethrough even when checked so as to remain clear to the reader. Paths of movement can still be easily noted in the simplified tree structure by the placement of traces (t) coindexed with the lexical material.

(5)

```
                PolP
          ┌──────┴──────┐
          a            TP
      iPol:+neg      ┌───┴────┐
        uV           T        VP
      uVerb          │      ┌──┴──┐
        uT           i    ngi-pheg
                  iT:-past  iVerb
                  iV: +V
                   uPol
```

(6)

```
                  PolP
            ┌──────┴──────┐
            a            TP
        iPol:+neg      ┌───┴────┐
          u̶V̶          T        VP
        u̶V̶e̶r̶b̶         │      ┌──┴──┐
          u̶T̶          i    ngi-pheg
                   iT:-past  iVerb
                   iV: +V
                    u̶P̶o̶l̶
```

(7)

```
                  PolP
            ┌──────┴──────┐
            a            TP
        iPol:+neg      ┌───┴────┐
          u̶V̶          T        VP
        u̶V̶e̶r̶b̶       ┌──┴──┐   ┌──┴──┐
          u̶T̶    ngi-pheg_i  i    t_i
                      iT:-past  iVerb
                      iV: +V
                       u̶P̶o̶l̶
```

(8)
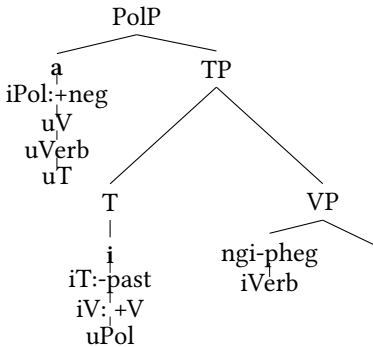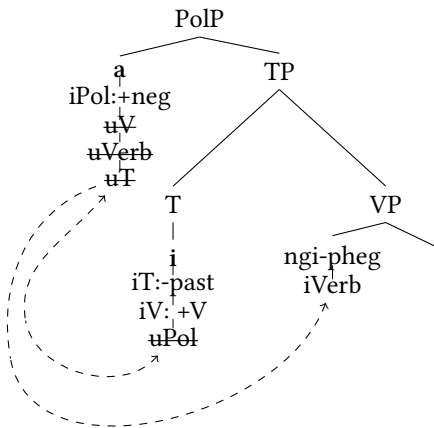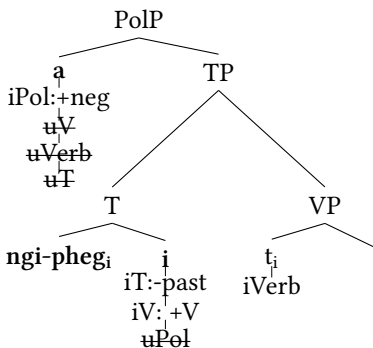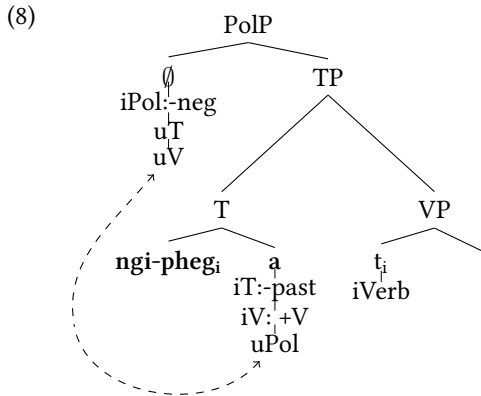
The reasoning for positing that final vowels are generated under tense is as follows. First, both positive and negative realizations of the final vowel contribute semantically to tense. Indeed, in examples such as (4) they are the only overt markers of tense. Furthermore, the complementary distribution of final vowels and negative suffixes makes an argument which treats them as generated under separate nodes undesirable. The solution presented in this paper accounts for these facts by positing that final vowels and negative suffixes are both generated under a tense node which is sensitive to the polarity of the sentence as well as other clausal information (discussed further in §4).

Returning to the discussion of the simple present tense sentences in (4), the system of Agreement discussed in §3.1 is demonstrated. As can be seen from the trees in (6) and (7), there are two primary nodes which participate in Agreement, PolP and TP. Each of these nodes is in an Agreement relationship with the other, and the Agreement relationship is largely symmetrical. Tense carries an interpretable tense feature as well as an interpretable feature which dictates whether the clause is verbal or nonverbal (+V or -V). The uninterpretable polarity feature on the tense node is checked by its Agreement relationship with PolP. PolP carries the interpretable feature polarity as well as two uninterpretable features, tense and uV, which is sensitive to whether or not the clause is verbal. Both of these uninterpretable features are checked by TP. Different realizations of these nodes are treated as separate lexical items, and thus can carry slightly different featural arrangements. Note that the negative polarity node also carries an uninterpretable uVerb feature (that positive polarity does not) which is checked by the verb. In verbal sentences this feature causes no distinguishable difference, and will thus be discussed in §4 on adjectival predicates. In general, relevant featural differences in the different realizations will be discussed in the corresponding sections. The Agreement relationship between PolP and TP will be a key component of the analysis presented here.

3.3 Past tense

The same analysis used for present tense clauses in §3.2 directly applies to past tense clauses (Table 4), as demonstrated in (9–11).

(9) a. mina ngi-khal-e.
         I   1.SM-cry-PST

    'I was crying.'

    b. mina a-ngi-khal-a-**nga**.
        I   NEG-1.SM-cry-EP-PST.NEG

    'I wasn't crying.'

Table 4: Realizations of Past Tense

	-nga-		-e-	∅
Neg	+	+	-	-
Verbal	+	-	+	-

The sentences in (9) are quite similar to the present tense sentences in (4). The trees are provided in (10–11).

(10)
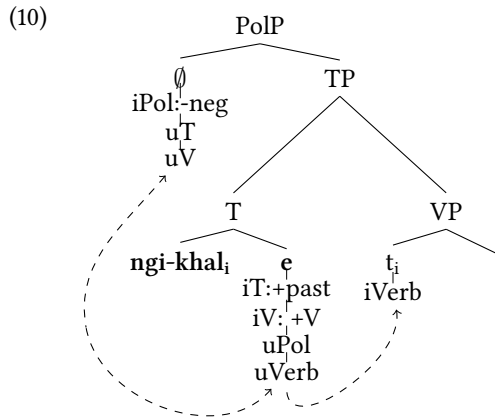
(11)
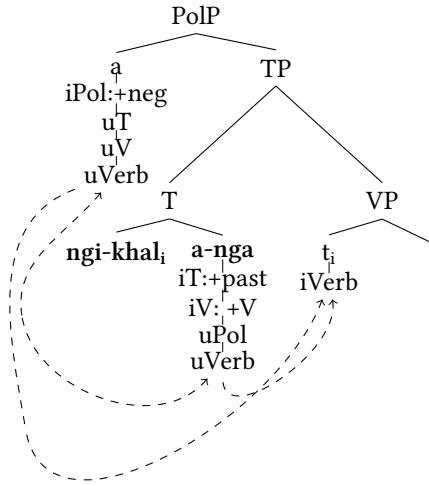
Lexical differences aside, the noteworthy difference from the present tense sentences discussed in §3.2 is the interaction between the verb and the negative past tense suffix *-nga-*.[5] Excluding the final vowel, all verbs in Ndebele end in consonants and maintain a CV syllable structure. Of the four morphemes generated under the tense node, only *-nga-* begins with a consonant. The result of this is that while the other three tense realizations fit into a CV pattern, head movement of a consonant final verb to the consonant initial tense marker *-nga-* would result in a CCV syllable. The result of this problem is that the default vowel *a* is epenthesized between the end of the verb and the beginning of the tense morpheme (Table 5).

Table 5: Realizations of *hamba* 'go'

	Present	Past
Positive	*hamb-a*	*hamb-e*
Negative	*hamb-i*	*hamb-nga* $\longrightarrow$ *hamb-a-nga*

4 Adjectival predicates

The realm of adjectival predicates is particularly important to the argumentation of this paper in that it demonstrates how a null verbal projection can deal with pre-verbal *-nga-*,

---

[5] As mentioned for negative polarity, note that past tense TP also carries an uninterpretable uVerb feature where present tense does not. The reasons for this will be discussed in §4.

as well as showing how this proposal gives an explanation for previously unaccounted for data in Ndebele. In this section, the theory developed in §3 is applied to adjectival predicate data.

4.1 The data and the puzzle

Examples of adjectival predicates in the present and past tense are given in (12) and (13).

(12)   a. isi-lwane si-hle.
            7-lion   7.SM-pretty
            'The lion is pretty.'

        b. isi-lwane **a**-si-∅-si-hle.
            7-lion   NEG-7.SM-∅-7.SM-pretty
            'The lion is not pretty.'

(13)   a. isi-lwane be-si-∅-si-hle.
            7-lion   be.PST-7.SM-∅-7.SM-pretty
            'The lion was pretty.'

        b. isi-lwane be-si-∅-**nga**-si-hle.
            NC-lion   be.PST-7.SM-∅-NEG-7.SM-pretty
            'The lion wasn't pretty.'

The major points of interest in the adjectival predicates shown above are the pre-predicate morpheme *-nga-* showing up in (13b) and the double subject marking which shows up in (12b) and (13a-b). In this paper, it is shown that by positing a null verbalizing projection which takes the first subject marker as a prefix, and the negative morpheme *-nga-* as a suffix, both problems can be solved with a single solution.[6]

4.2 The solution

As mentioned, this paper deals with the puzzles in (12) and (13) by positing a null verbal projection which takes subject marking as a prefix, and tense and polarity information as a suffix, just as regular verbal projections were shown to do in §3. The motivation for the appearance of this null verbal projection is featural in origin. Following Bjorkman's discussion of auxiliaries, a parallel is drawn between the null verb support seen here in adjectival predicates and the familiar do-support phenomenon in English in that both are motivated by a failure in Agreement (Bjorkman 2011). Specifically, the presence of this null verbal element is triggered by the uninterpretable uVerb feature which appears on both past tense and on negative polarity. We saw this feature in §3 on verbal predicates, but in that context there was no failure to agree owing to the already present verbal element. In these adjectival predicates, however, where no verb is present, a null

---

    [6] It should be noted that to many scholars the presence of the pre-predicate *-nga-* is not considered a problem, but is rather a negative prefix. The double subject marking remains a problem regardless.

verbal element is inserted (in the location consistent with our placement of the verb immediately following TP) in order to check a null uVerb feature if there is either a past tense or negative polarity morpheme.[7] Thus we derive the difference seen between (12a) on the one hand, in which no uVerb feature appears, and thus no null verb; and (12b) and (13a–b) on the other hand, in which the presence of a uVerb feature triggers the presence of the null verbal element.

Once inserted, the null verbal element undergoes regular V-T raising, solidifying our claim that *-nga-* is always post-verbal, though on the surface the null verbal element gives the illusion that *-nga-* appears before the adjectival predicate *hle*. This allows us to maintain our claim that there is a single negative morpheme *-nga-* which always appears post-verbally. In addition, the positing of this null verbal projection allows us to account for the presence of the double subject marking seen in examples (12b) and (13a–b) by positing that subject agreement attaches to both the null verbal element as well as the adjective. Examples (14–17) show how this analysis applies to the sentences in (12–13).

(14)

(15)

---
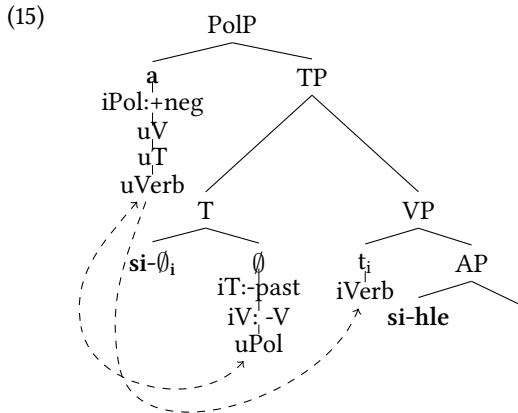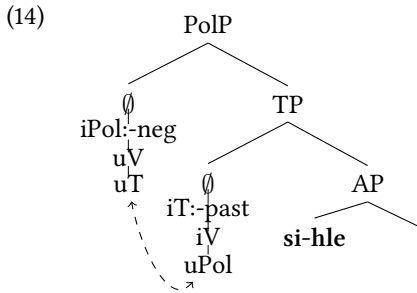
[7] It is important to note that in this analysis, the auxiliary *be* does not have an iVerb feature, and thus cannot check uVerb features in adjectival clauses like (13). See §4.4 for more information on *be*.

(16)

(17)
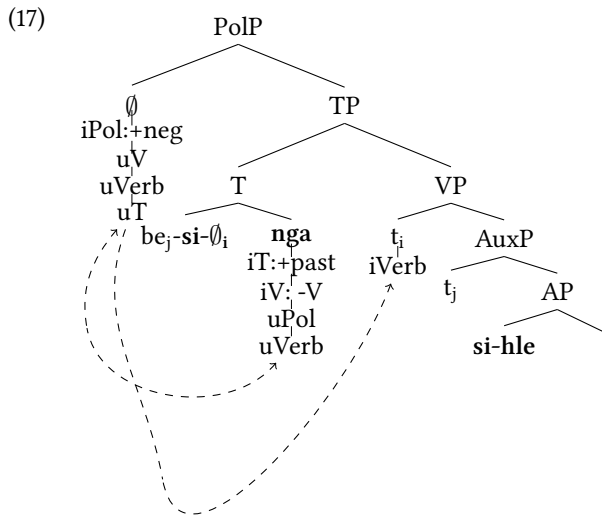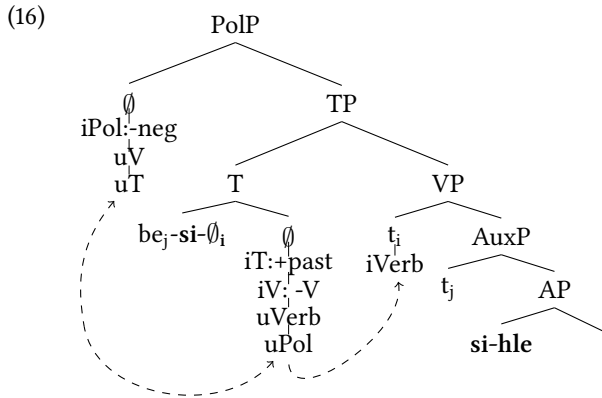
4.2.1 Multiple subject marking

A counter proposal to the one presented in §4.2 would be one that uses *be* as an additional verbal element which triggers the presence of the second subject marking, instead of positing a null verbal element. In this scenario, *be* is generated where our null verbal element is (see 15 and 17), causing a second agreement marker to appear, and in the case of (17), taking *-nga-* as a suffix. The argument would then have *be* moving up to a higher node in order to get the correct surface ordering.

One reason why this solution is not implemented in this paper is evidence that *be* can generate its own subject marking independent from and in addition to that seen in (13a-b). This third subject marking is shown in (18d).

(18)   a.   yenza isi-lwane si-be-si-hle.
           make 7-lion   7.SM-be.PST-7.SM-pretty
           'Make the lion be pretty.'

       b.   yenza isi-lwane si-be-∅-nga-nhle.
           make 7-lion   7.SM-be.PST-∅-NEG-pretty
           'Make the lion not be pretty.'

       c.   yenza isi-lwane si-be-si-∅-nga-nhle.
           make 7-lion   7.SM-be.PST-7.SM-∅-NEG-pretty
           'Make the lion not be pretty.'

       d.   yenza isi-lwane si-be-si-∅-nga-si-hle.
           make 7-lion   7.SM-be.PST-7.SM-∅-NEG-7.SM-pretty
           'Make the lion not be pretty.'

       e.   *yenza isi-lwane si-be-∅-nga-si-hle.
           make   7-lion   7.SM-be.PST-∅-NEG-7.SM-pretty
           'Make the lion not be pretty.'

Additional evidence that *be* generates its own independent subject marking is given by Sibanda, who provides the following forms (Sibanda 2004).

(19)   a.   zibe zihamba ⟶bezihamba 'they were walking/moving/going'
           sibe sikhala ⟶besiskhala 'it was crying'

       b.   ibe ihamba ⟶ibihamba 'it was walking'
           ube ekhala ⟶ubekhala 's/he was crying'

The evidence provided in this section makes it unlikely that the double subject marking can be accounted for by *be* alone.

4.2.2 Selection

Now that more than one type of predicate has been introduced, the selectional properties of TP need to be discussed. One benefit of an approach which accounts for the null verbal projection with respect to repair is that the selectional properties of TP do not then need to account for this phenomenon. Rather, we can posit that present tense adjectival TP selects for AP and past tense adjectival TP selects for a BeP, which in turn selects an AP.[8]

Selection[9]

(20) $TP_{Adj, Past}$: =AP, T -V +Past

(21) $TP_{Adj, Present}$: =BeP, +Adj, T -V -Past

(22)   BeP: =AP, Be

---

[8] BeP is a convenient name for the projection most often realized as *be*; see §4.4 for more discussion of *be*.

[9] For a full list of lexical entries, see the appendix.

If instead of this repair scheme, the motivation for the null verb was accounted for purely via selection, the account would run into troubles. Most importantly, it can be seen that the null verb appears in either the presence of negation or past tense. Thus if this scenario were accounted for purely through selection, both negative polarity and past tense would select for a null verbal element. With this set up, a sentence like (13b) would instead appear as in (23), which is ungrammatical.

(23) *isi-lwane be-si-∅-si-∅-**nga**-si-hle.
      7-lion   be.ᴘꜱᴛ-7.ꜱᴍ-∅-7.ꜱᴍ-∅-ɴᴇɢ-7.ꜱᴍ-pretty
      'The lion wasn't pretty.'

Rather, it seems that a single verbal component is enough to satisfy the requirements of both the NegP and the past tense. This observation fits nicely with the system of Agreement adopted in this paper.

It is worth noting as well that *-nga-* is the only overt realization of tense which appears in non-verbal clauses. The realizations of non-verbal past tense are specifically designed this way in Table 6 in order to prevent the appearance of other overt final vowels appearing before adjectival predicates. The featural specifications for realizations of tense are repeated in Tables 6–8 for convenience.

Table 6: Realizations of Past Tense

	-nga-		-e-	∅
Neg	+	+	-	-
Verbal	+	-	+	-

Table 7: Realizations of Present Tense

	-i-	-a-		∅	
Neg	+	-	+	-	
Verbal	+	+	-	-	

Table 8: Realizations of Future Tense

	-za-			
Neg	+	+	-	-
Verbal	+	-	+	-

4.3 Future tense adjectival predicates

This section looks at the way future tense adjectival predicates are formed in Ndebele. Though these forms show slightly different surface structure than the present and past tense counterparts, they still are able to fit into the system otherwise outlined in this paper.

(24)   a.   isi-lwane si-za-be   si-∅-si-hle.
             7-lion   7.SM-FUT-be.PST 7.SM-∅-7.SM-pretty
             'The lion will be pretty.'

       b.   isi-lwane si-za-be   si-∅-nga-si-hle.
             7-lion   7.SM-FUT-be.PST 7.SM-∅-NEG-7.SM-pretty
             'The lion will not be pretty.'

The difference between future tense adjectival predicates and their counterparts is the future morpheme *-za-*. This morpheme, which originally was an independent verb, has become grammaticalized to serve as the future tense marker, though presumably due to its historical origins it does so outside of the canonical TP position (Sibanda 2004). The trees in (25) and (26) represent the sentences in (24a–b).

(25)
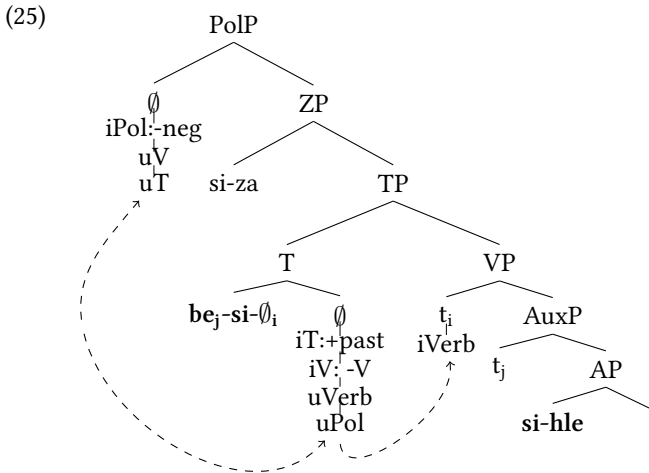
(26)
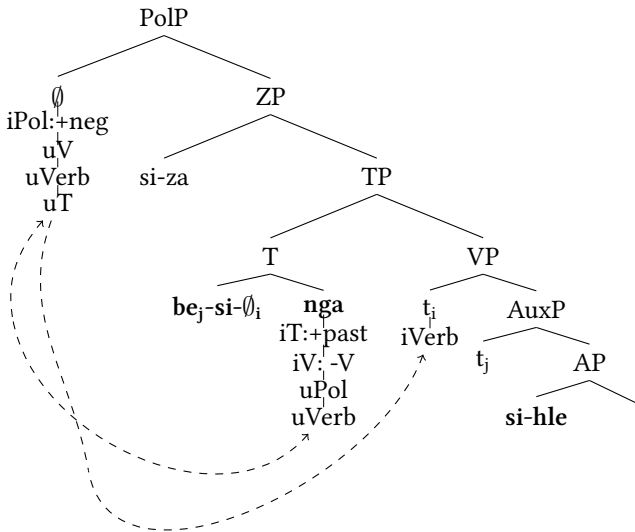
4.4 Accounting for *be*

In §4.2 on past tense adjectival predicates we accounted for the presence of the auxiliary verb *be* by stipulating that it is selected for by a past tense non-verbal tense node. In examples (24a-b), however, we find *be* appearing with a semantically future tense clause. Thus, some explanation needs to be given in order to make sense of the distribution of *be*. On the account presented here the presence of *be* is strictly selected for by a past tense non-verbal tense node. Thus, more needs to be said about ZP. In the analysis above, the node ZP has been introduced as a sort of tense node (which makes it an available to be selected by PolP), which itself selects for a non-future tense node. The ZP tense node differs from regular TPs in a few crucial ways. The first is that while it carries a +/-V feature, in place of the binary +/-past feature, it simply has a +FUT feature. The +/-V feature is crucial to this node in that it dictates what sort of tense node is selected by the ZP. Verbal clause ZPs select a -past TP, while non-verbal clause ZPs select for a past tense node, as seen in the examples above. This selection of the past tense node is what, in turn, allows us to explain the presence of *be* where it would not normally be expected (outside of a semantically past tense clause).

5 Imperatives

This section examines how imperatives are formed and negated in Ndebele. This particular phenomenon will be examined because it is another case in which the negative affix *-nga-* appears pre-verbally in the surface structure. Once again the proposal outlined in the preceding sections will be applied to this phenomenon.

The basic form of the imperative appears in (27–28).

(27) a. hamb-a!
          go-FV
          'Go!'

  b. u-**nga**-hamb-i!
          1.SM-NEG-go-NEG
          'Don't Go!'

(28) a. val-a   um-ngango!
          close-FV 3-door
          'Close the door!'

  b. u-nga-val-i   um-ngango!
          1.SM-NEG-close-NEG 3-door
          'Don't close the door!'

As mentioned above, the negative affix *-nga-* appears in the negated forms in (27b) and (28b). What is puzzling about these facts, given the descriptions and analyses so far, is why *-nga-* should be the affix that negates such sentences, as up to this point *-nga-* has only appeared either in past tense clauses, or future adjectival clauses. While the use of *-nga-* in (27b) and (28b) does not fit the pattern we have seen so far, it could of course be part of a larger pattern not analyzed in this paper.

Aside from these distributional concerns, the form of the negative imperatives above seems completely amenable to the analysis developed here. By positing that in these situations too there is a null verbal element which takes both a subject marker and negative suffix, the surface string can be brought into the analysis. The tree for this is given in (29).

(29)

PolP

∅
iPol:+neg
uV
uVerb
uT

TP

T

u-∅ᵢ

nga
iT:-past
iV: -V
uPol

VP

tᵢ
iV

TP

T

∅-**hambⱼ**
iT:-past
iV: +V
uPol

i

VP

tⱼ
iV
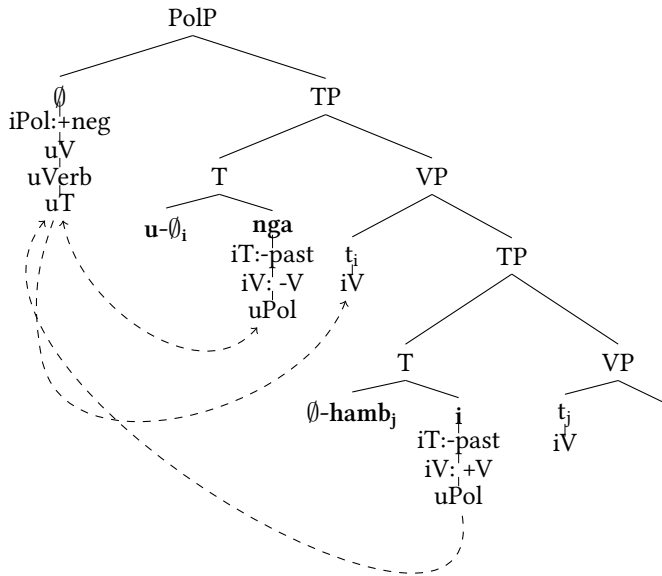
The motivation for such an analysis seems to be primarily lacking in the case of imperatives. Throughout this paper, it has been the tense node (or the tense auxiliary) which has selected for a null verbal element. In the case of imperative sentences like the one above, it is unclear what would motivate the tense node to select for a null verbal element, bringing imperatives in line with the rest of our analysis. Thus it is unclear whether imperative sentences should be analyzed under the same rubric as the other examples in this paper, or whether they are better described as a separate phenomenon.

6 Conclusions

In this paper, negation in Ndebele has been examined across many different subcategories of the language. In §3.1, the distribution of negative morphemes in Ndebele was captured using a complex featural system and bi-directional agreement. In §3.2, a proposal for the clausal structure of negative sentences in Ndebele was proposed with respect to simple tense past and present verbal predicates. In §3.3, this analysis was extended to cases where the negative affix *-nga-* appears pre-verbally instead of post-verbally, as it does in the simple tense sentences.

When adjectival predicates were examined in §4, it was seen that by positing a null verbal projection which hosts both subject agreement and the polarity sensitive tense

node *-nga-*, it was possible to account for seemingly pre-fixal *-nga-* and to explain previously unaccounted for double subject marking in certain adjectival predicates. In §5, the analysis was extended to imperatives, again an area which exhibits pre-verbal *-nga-*. With respect to this phenomenon, however, the motivation for extending the analysis was significantly weaker, enough to call into question the application of the null verbal element as a solution for pre-verbal *-nga-*.

More detailed questions still need to be answered including what the proper analysis of imperatives should be. That is, should they be treated with the approach outlined in this paper, or is this phenomenon merely similar from the perspective of surface structure? Another troubling feature of this analysis is the lack of evidence for the negative affix *-i-* to appear pre-verbally. The claim made in this paper is that both *-nga-* and *-i-* are generated in a tense sensitive tense node. All things being equal, one would assume that finding *-nga-* pre-verbally would be an indicator that *-i-* should appear there as well, though there might be confounding variables at work which prevent *-i-* from surfacing overtly in this position.[10]

This paper is intended as a contribution to the field of Bantu linguistics and to studies on polarity. As this paper presents an alternate analysis of polarity and tense in Ndebele, future research should be able to ascertain whether or not aspects of this analysis, such as the treatment of final vowels as tense nodes, can be extended to other Nguni or Bantu languages. Other languages which exhibit bipartite negation would also serve as targets for comparison with respect to the analysis of bipartite negation developed here. Such comparisons and questions are left for future research.

Appendix

This appendix shows the selectional properties of the various syntactic constructions used in this paper.

Be:	=AP, Be	TP_{-Main}:	=BeP -M, T -M
BeP_{-Main}:	=PolP, Be -M	TP_V:	=VP, T +V
PolP:	=TP, Pol	ZP_{+V}:	=TP PRES +V, T
$TP_{Adj, Past}$:	=AP, T -V +Past	ZP_{-V}:	=TP PAST -V, T
$TP_{Adj, Present}$:	=BeP +Adj, T -V -Past		

Abbreviations

Numbers signify noun classes (e.g. 1, 7)

---

[10] See Buell (2004) for some evidence that *-i-* might indeed appear pre-verbally

COP	copula	NEG	negation
EP	epenthetic (vowel)	PRS	present tense
FUT	future tense	PST	past tense
FV	final vowel	SM	subject marker

References

Bjorkman, Bronwyn. 2011. *Be-ing default: The morphosyntax of auxiliaries.* Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation.)

Bošković, Željko. 2007. On the locality and motivation of Move and agree: An even more minimal theory. *Linguistic Inquiry* 38. 589–644.

Buell, Leston C. 2004. The Zulu verb within the constraints of the LCA. (Paper presented at the 35th Annual Conference on African Linguistics, Boston, MA.) http://www.fizzylogic.com/users/bulbul/applications/papers/acal-2004-zulu-ip-paper.pdf, accessed 2013-05-28.

Buell, Leston C. 2005. *Issues in Zulu verbal morphosyntax.* Los Angeles: UCLA. (Doctoral dissertation.)

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Khumalo, J. S.M. 1982. Zulu tonology, part 2. *African Studies* 41(1). 3–126.

Khumalo, James S. M. 1981. Zulu tonology, part 1. *African Studies* 40(2). 53–130.

Rizzi, Luigi. 1989. *Relativised minimality.* Cambridge, MA: The MIT Press.

Sibanda, Galen. 2004. *Verbal phonology and morphology of Ndebele.* Berkeley: University of California at Berkeley. (Doctoral dissertation.)

Zeijlstra, Hedde. 2010. There is only one way to agree. (Paper presented at at GLOW 33, Wroclaw, Poland.)

Chapter 14

The productivity of the reversive extension in Standard Swahili

Deo Ngonyani

Michigan State University

Nancy Jumwa Ngowa

Pwani University

This paper is a study of the Swahili reversive verb derivation *-ul*, which is also known as conversive or separative. Attachment of this suffix to a root X derives a verb with the meaning 'undo X.' The paper explores the productivity of the reversive using data generated in two elicitation tests: (a) a coinage test and (b) an acceptability judgment test. For the coinage test, we created a questionnaire of 20 nonsense verbs from which subjects were instructed to coin their reversives. Subjects were able to create the reversive forms without difficulty. For the second questionnaire, we created reversive forms using 20 most frequently used verbs from SALAMA of Helsinki Corpus of Swahili. For each of these verbs, a reversive verb was created. Subjects were asked to identify each derived verb as either (a) an existing word, (b) a possible word or (c) impossible. Only three words polled as existing in the language. This revealed that the main constraint for the 20 verbs is that the reversive can attach only to verbs with semantic meaning that may be un-done or reversed. From these tests, we conclude that although the reversive applies to a restricted set of verbs, it is productive and available for creation of new words.

1 Introduction

This is a study of the reversive (ʀᴇᴠ) derivation in Standard Swahili, a verbal derivation that is illustrated in the contrast between (1a) and (1b).

(1) a. M-toto a-li-fung-a   dirisha.
         1-child 1sᴍ-ᴘᴛ-shut-ꜰᴠ 5.window

         'The child shut the window.'

     b. M-toto a-li-fung-u-a   dirisha.
         1-child 1sᴍ-ᴘᴛ-shut-ʀᴇᴠ-ꜰᴠ 5.window

         'The child opened the window.'

The suffix *-u* appears on the root *fung* 'shut' in (1b) with the semantic effect of reversing the action denoted by the root in (1a). Therefore, *fung* 'shut' is reversed into *fungu* 'open'. This derivation is also known as a separative (Schadeberg 2003), conversive (Ashton 1947), or inversive (Doke 1935).

The reversive is part of the system of derivations in Bantu languages, which Bantuists refer to as "verb extensions" (Guthrie 1962). There are three types of derivational suffixes:

- Valence-increasing suffixes (applicative, causative)
- Valence-reducing suffixes (passive, reciprocal, stative)
- Non-valence-changing suffixes (reversive, contactive, static)

While valence-increasing suffixes and valence-reducing suffixes are considered productive, non-valence-changing suffixes are regarded as not productive (Shepardson 1986). Most of the meanings of such suffixes seem to be fossilized.

Apart from the reversive reading, the reversive extension exhibits a range of idiosyncratic meanings, as well as lexicalized forms in Swahili. Its idiosyncrasies in meaning, and its supposed unproductivity, have led dictionary makers to list reversive verbs as separate entries, not related to the roots or not derived like forms involving other derivational affixes (for example TUKI 2001; 2004; Mdee et al. 2009). Although the reversive words are not listed as derivations in the dictionaries, the consistency and transparency of the semantics of many reversive words is so obvious that it is impossible to ignore. With respect to productivity, the status of the reversive is not clear. Shepardson (1986) considers the reversive as not productive, while Schadeberg (1973) claims that it is productive.

The objectives of this paper are two-fold. The first objective is to describe the reversive derivation, its manifestation, and its semantics. The second objective is to explore its productivity. We argue that the reversive in Swahili is a productive affix that can be used by speakers to create new words. We demonstrate that the apparent relative unproductivity is due to structural restrictions.

In §2, we provide an overview of productivity theory and methodology, while in §3 we present the basic phonological, morphological, syntactic, and semantic facts about the reversive. The data for this section are words from dictionaries of Standard Swahili. In §4, the methodology for measuring productivity used in this study is described, and results are discussed in §5. In §6, we make concluding remarks.

2 Morphological productivity

Studies of morphological productivity are based on the intuition that some word-formation processes or rules are used a great deal more than others. For example, in English *-ness* and *-ity* both can be used to derive nouns from adjectives. Consequently, we get *happy → happiness* and *sensitive → sensitivity*. The suffix *-ness* is used a lot more than *-ity*. Words like *sensitiveness* may be recognizable, but *happity sounds not English. In this section, we define productivity, identify its factors and constraints, and discuss how productivity is measured.

Several definitions of productivity have been advanced. Among them are:

> The productivity of a word-formation process can be defined as its general potential to be used to create new words and as the degree to which this potential is exploited by the speakers. (Plag 2006: 127)

> The productivity of a morphological process (whether inflectional or derivational) has to do with how much (or, in the limiting case, whether) it is used in the creation of forms which are not listed in the lexicon. (Bauer 2005: 315)

> A morphological rule or pattern is said to be productive if (and to the extent that) it can be applied to new bases and new words can be formed by it. (Haspelmath & Sims 2010: 114)

These and several other definitions converge on two senses of productivity, namely, the potential to create new words and the extent to which speakers use the word-formation process or rule to create new words (Corbin 1987 cited in Bauer 2001). If a rule or process can be used in the creation of new words at a particular point in time, it is said to be available. A rule that is available is considered productive (Bauer 2001: 205; Plag 2006). Speakers can use the affix or rule to create new words. If an affix or rule can no longer be used for creation of new words, it is not available. Such a rule is not productive. A rule that is available should be used by the speech community rather than just in a single person's idiosyncratic language. Availability is a *yes-no* issue. That is to say, a rule is either available or not available. Two English suffixes *-ion* (as in *action*) and *-ment* (as in *judgment*) illustrate this contrast. Although both are widely attested with many words, *-ment* is not available for creation of new words while *−ion* can still be used (Bauer 2001).

In contrast, the extent to which a rule or process is used is not a *yes-no* issue. In productivity studies, this is known as profitability. "The profitability of a morphological process reflects the extent to which its availability is exploited in language use, and may be subject unpredictably to extra-systemic factors" (Bauer 2001: 211). In this sense, profitability is attested along a continuum. For example, the suffixes cited earlier, *-ness* and *-ity,* are both used today in the creation of new words. Measurements have revealed that *-ness* is more frequently used than *-ity*. Therefore, *-ness* is more profitable than *-ity* (Plag 2006). Our concern in this study is primarily on the availability (i.e. productivity rather than profitability) of the reversive to create new words.

Three features are considered factors that promote productivity. They are (a) TRANS-PARENCY, (b) the FREQUENCY OF THE BASE, and (c) the USEFULNESS of the word (Bauer 2001; Lieber 2009). Bauer (2001) calls them prerequisites of productivity. Transparency refers to clear segmentation and one-to-one meaning-form correspondence. That is, one form appears everywhere in the same shape with the same meaning. For example, *candidness*, *commonness*, and *oddness* all have the same form of the affix *-ness* with the meaning 'state of being X'. This has a better chance of being more productive than *-ity*, as in *timidity*, *locality*, and *oddity*. The suffix *-ity* triggers some phonological changes in

the pronunciation. For example, [ˈtɪmɪd] → [təˈmɪdəti]. Moreover, the resultant meaning is not always 'the state of being X'; for example, *oddity* is not 'the state of being odd', neither is *locality* 'the state of being local'. This makes *-ity* less productive than *-ness*. The frequency of the base is the extent to which bases are available for the rule. If an affix can only be attached to a small set of bases, it is likely to be less productive than an affix that can attach to a wide range of bases. For example, when we compare *-ness* to the deadjectival *-en*, as in *soften* and *blacken*, we see the former affix is used in the derivation of a wide range of adjectives. The latter, however, applies only to a small set of adjectives that are monosyllabic. The bases for *-ness* are more frequent compared to the bases for *-en*. To the extent that a particular process creates words that are needed by the speech community, the word is said to be useful. Neologisms are created to serve a certain naming need (Štekauer 2005). A process that creates such neologisms is productive. We demonstrate in this study that in spite of its transparency and usefulness, the reversive in Swahili does not have too many bases to which it can attach.

In fact, bases and affixes may be subject to several restrictions. These are extensively discussed by, among others, Bauer (2001), Haspelmath & Sims (2010), Lieber (2009), Rainer (2005), and Plag (2006). One kind of restriction that affects the productivity of a rule is phonological. An affix or a rule may be sensitive to certain phonological conditions. For example, the English suffix *-eer* appears on bases that end in /t/, such as *profiteer* and *racketeer*, and not in other environments. Therefore, there is no **gaineer* or **fraudeer*. Another kind of restriction is morphological, such as the presence of some morpheme that may attract some other specific affix. In English, for example, verbs with the affix *-ise* attract the affix *-ation* as a deverbal derivation (Fernández-Domínguez 2013). Syntactic restrictions also restrict productivity. The applicability of affixes and morphological processes may be sensitive to the word class. In specifying the productivity of the suffix *-able*, for example, we must state that it attaches to transitive verbs. Etymology or the origin of the base may play a crucial role in restricting the extent to which a rule is used. In English, the adjective-forming suffix *-ic* is restricted, because it attaches to words that are loans from Latin or French, such as, *specific* and *eclectic*. Another restriction is known as semantic blocking. The existence of a word may block the derivation of another word with the same meaning that uses a particular morphological process. For example, we have *pig→piglet,* but the suffix *–let* is not available for *cow→*cowlet*. In this case, it is believed that since the word *calf* already exists in the lexicon, it blocks the derivation that would yield *cowlet*. Some rules, affixes, or processes are constrained by pragmatic or sociolinguistic conditions. Bauer (2001: 135) cites an example from Dyirbal in which the suffix *-ginary* 'covered with, full of' is only used for something that is unpleasant, for example, *guna-ginay* 'covered in faeces'.

In recent years, several empirical measurements and methods have been developed to gauge the productivity of rules, processes, and affixes. Bauer (2001: 143-161) and Fernández-Domínguez, Díaz-Negrillo & Štekauer (2007) present useful surveys of measurements proposed by various scholars. As noted earlier in the definition of productivity, the measurements can be linked to the two approaches to productivity, namely, productivity as availability of a rule and productivity as profitability or the degree to which a

rule is used. Work that focuses on determining availability is said to take a qualitative approach. The qualitative approach takes into consideration the limitations and constraints of a process. The quantitative approach, on the other hand, assumes that a more productive process will be able to coin more lexemes. Productivity is measured by counting or quantifying attested forms with higher frequencies. In the quantitative approach, productivity is scalar, ranging from very productive to not productive at all. Let it be noted that the characterization of the two approaches does not mean that the qualitative approach does not use any statistical methods (Fernández-Domínguez 2013: 434-437). Our study of the reversive takes into account various qualitative aspects while focusing on whether or not the reversive is available to contemporary speakers of Swahili.

Proposals for measuring productivity can be summarized under three models according to Fernández-Domínguez, Díaz-Negrillo & Štekauer (2007: 35): (a) frequency models, (b) probabilistic models, and (c) onomasiological models. Frequency models are based on the intuition that if an affix is found on many words, it is very productive. Such counting of attested words is done by Plag (1999) with data from the *Oxford English dictionary*. A major issue with this model is that the measures deal with attested words with a particular affix. The nominal suffix *-ment* in English, for instance, is often cited as an example of the flaw in frequency measurements. It appears in many words, such as in *punishment*, *entertainment*, *appointment*, and *commitment*. However, this affix is no longer available for the creation of new words. All the words it appears on were formed between mid-16th and mid-18th century (Bauer 2001: 181). It is, therefore by definition, not productive. Probabilistic models statistically measure the probability of finding a new word by a given morphological process. Productivity is measured by counting neologisms. One such measure is hapax legomenon or unique formations. The insight behind this is that productive rules will produce more unique formations while unproductive rules will not produce such formations (Baayen 1992). The onomasiological model is attributed to Štekauer who sought to focus studies of productivity on meaning first, rather than form first (Štekauer 2005). The model is based on the idea that processes and rules come into play when there is a need for creating a word for what Štekauer calls naming units. A word-formation process that results in creating a new naming unit is therefore considered productive.

Accordingly, there are three main sources of data for studying productivity (Plag 2006; Schroeder & Muehleisen 2010; Bolozky 1999). The first source is dictionaries from which one can search for neologism and words of particular word-formation processes or affixes. For example, Fernández-Domínguez (2013) makes use of the *Oxford English Dictionary*. However, words may appear in the speech of the community that are not yet in the dictionary, because dictionaries tend to use formal language that is often old. The second source are language corpora in which naturally occurring words appear. Statistical methods of calculating frequencies of forms or tokens make use of such databases. Baayen's index of hapax legomena, or words occurring only once in a corpus, is based on the notion of the probability of creation of new words based on neologisms in a corpus (Baayen 1992). Plag (2006) and Fernández-Domínguez (2013) use data from the British National Corpus. The third source of data is elicitation tests. There are two kinds of

elicitation tests. One elicits acceptability judgments on existing or non-existing forms (Aronoff & Schvaneveldt 1978). The other prompts subjects to coin new words within the parameters of some word-formation processes. Aronoff and his associates have sought to investigate productivity as a search for possible words, just as syntax studies investigate possible sentences and phrases (Aronoff 1976). The present study pursues this idea of potentiality and the search for possible words.

Before we describe our test, let us describe the morphological phenomenon we investigate. In the following section, we present a description of the reversive derivation in Swahili.

3 The reversive derivation

In this section, we describe the reversive in terms of its phonological, morphological, syntactic, and semantic features. We also examine constraints with respect to the semantics of the bases to which the reversive may attach. All examples presented here are from three dictionaries of Standard Swahili: *Kamusi ya Kiswahili-Kiingereza,* 'Swahili-English dictionary' (TUKI 2001); *Kamusi ya Kiswahili Sanifu,* 'A dictionary of Standard Swahili' (TUKI 2004), and *Kamusi kamili ya Kiswahili,* 'A complete Swahili dictionary' (Mdee et al. 2009).

As in other Bantu languages, the verb in Swahili exhibits a complex morphological structure consisting of inflectional prefixes, derivational suffixes and inflectional suffixes. Inflectional prefixes include subject markers, object markers, tense, negation, and relative markers. Inflectional suffixes include mood and negation. Derivational suffixes appear between the root and the inflectional suffixes. The following two examples illustrate some of the inflections and derivations.

(2) a. Wa-li-tu-lip-i-a.
     2SM-PT-us-pay-APP-FV
     'They paid for us.'

   b. Wa-lip-ish-eni.
     2OM-pay-CAUS-IMP
     '(You all) make them pay.'

The root for these two examples is *lip* 'pay.' In (2a), the derivational suffix is the applicative *-i* and in (2b) the derivational suffix is the causative *-ish.* Other derivations include the reciprocal, the passive, and the stative. The final vowel *-a* appears after the derivational suffixes, even in citation forms. The final vowel often stands in contrast to the subjunctive mood marker *-e.* It is also replaced by *-eni* in plural imperative. For this reason, *-a* is often considered the indicative marker. There is also a negative suffix *-i* for present tense negation and the question marker *-je.*

The reversive is part of the complex system of verbal derivations, which display varying degrees of productivity. It is realized in four allomorphs, namely, *-u, -o, -ul* and *-ol.* The first two are illustrated in the following set of examples.

(3) *-u* and *-o* allomorphs of the reversive
  a. zib-a 'plug'   zib-u-a 'unplug'
  b. teg-a 'set trap' teg-u-a 'disassemble trap'
  c. pang-a 'arrange' pang-u-a 'disarrange'
  d. fung-a 'shut'   fung-u-a 'open'
  e. chom-a 'stab'   chom-o-a 'pull out'

The examples presented here do not include the inflectional prefixes. The verb roots in (3) all end in consonants. The final vowel *-a* is used for the citation forms of the verbs.

The allomorph *-o* in (3e) has a restricted distribution. It appears when the root has the vowel /o/. The allomorphs *-ul* and *-ol* appear on verbs that also carry other extensions such as causative, applicative, and passive in addition to the reversive. The following examples show these allomorphs of the reversive followed by the applicative derivation.

(4) *-u*, *-o*, *-ul* and *-ol* allomorphs of the reversive
  a. zib-u-a 'unplug'   zib-ul-i-a 'unplug with/for'
  b. teg-u-a 'disassemble trap' teg-ul-i-a 'disassemble trap with/
                         for/on'
  c. pang-u-a 'disarrange'   pang-ul-i-a 'disarrange with/for/on'
  d. chom-o-a 'draw out'   chom-ol-e-a 'draw out with/for/on'
  e. fung-u-a 'open'   fung-ul-i-a 'open with/for/on'

The allomorph *-ol* appears on verbs whose roots have the vowel /o/. The applicative suffix in this case surfaces as *-e.* The other verbs in (4) have *-ul* for the reversive and *-i* for the applicative extension. The liquid /l/ is part of the reversive extension rather than the applicative. This analysis is based on the fact that other extensions that are known to have no /l/ also appear to trigger the *-ul/-ol* reversive allomorphs, as shown in the examples below.

In example (5a-c), the additional extension is causative *-ish*, while in (5d-e) the passive *-iw/-ew* is added.

(5) reversive allomorphs with causative and passive
  a. zib-u-a 'unplug'   zib-ul-ish-a 'cause to unplug'
  b. pang-u-a 'disarrange' pang-ul-ish-a 'cause to disarrange'
  c. fung-u-a 'open'   fung-ul-ish-a 'cause to open'
  d. chom-o-a 'pull out' chom-ol-ew-a 'be pulled out'
  e zib-u-a 'unplug'   zib-ul-iw-a 'be unplugged'

Furthermore, historical comparative evidence indicates that the reversive in Proto-Bantu may have been **-ʊd* (Schadeberg 2003; Meeussen 1967). Nurse & Hinnebusch (1993: 370) also identify **-ul* as the proto-form in Swahili's immediate family of North East Coast Bantu.

The form *-uk* is sometimes cited as an intransitive reversive, in contrast to the transitive reversive described in the foregoing paragraphs (Ashton 1947: 239). Consider the examples below.

(6)   reversive plus stative: *-uk*
  a. zib-u-a 'unplug'   zib-uk-a 'get unplugged'
  b. teg-u-a 'set trap'   teg-uk-a 'get unset'
  c. pang-u-a 'disarrange' pang-uk-a 'get disarranged'
  d. chom-o-a 'draw/pull out' chom-ok-a 'get pulled out'
  e. fung-u-a 'open'   fung-uk-a 'get opened'

However, *-uk* may be regarded as a fusion of the reversive *-ul* and the stative *-ik*. The two are merged with attendant loss of /l/ and /i/. The examples in (6) have *-uk* and *-ok* as the suffix complex forms. They are formed by the fusion of the suffix allomorphs *-u*, *-o* and the stative suffix *-ik (pang-ul-ik → panguk)*.

Semantically, the reversive mostly expresses the undoing of an action denoted by the root. We may characterize this as 'undo X' or 'reverse X' where X is the meaning of the base.

(7)   'undo, reverse X' sense
  a. panga 'arrange'   pang-u-a 'disarrange'
  b. simba 'encode in symbols' simb-u-a 'decode'
  c. tanza 'complicate'   tanz-u-a 'solve, clarify'
  d. tata 'tangle'   tat-u-a 'disentangle'
  e. vaa 'dress'   v-u-a 'undress'

One basic constraint on the reversive derivation, therefore, is that the root must denote an action that is reversible. For example, with *simba* 'encode', the reversive *simbua* leads to 'undoing the code' or 'decoding'. Verbs denoting events that cannot be undone cannot receive the reversive derivation. For example, *sema* 'say' → **semua* 'unsay' is not possible. Indeed, the verbs with meaning that can be undone are few.

The 'undo' or 'reverse an action' reading is not the only meaning for the reversive extension. Other meanings include intensification (Ashton 1947: 239; Polomé 1967: 90; Schadeberg 2003: 78).

(8)   'intensification' sense
  a. kama 'squeeze'   kamua 'squeeze out'
  b. songa 'press'   songoa 'wring'
  c. mega 'cut off piece' megua 'cut off piece'

While *kama* 'squeeze' is an underived stem form, the affixation of *-u* yields *kamua* 'squeeze out,' which is more intensive in meaning. Likewise, *songa* 'press' is intensified into *songoa* 'wring'. From these, we read some notion of removal.

There exist words in the vocabulary that appear to be derived by the reversive suffix, but whose meaning is neither clearly 'reversive' nor 'intensive'. The reputation of the reversive as unproductive is mainly due to such forms in which the meaning of the reversive derivation is not clear.

(9) idiosyncratic meaning
    a. tamba   'narrate'   tambua   'recognize'
    b. remba   'decorate'   rembua   'make eyes at'
    c. enga   'watch'   engua   'skim (e.g. cream off of milk)'
    d. kosa   'err'   kosoa   'criticize'
    e. koma   'stop'   komoa   'do something deliberately to hurt someone'

The meanings of the apparently derived verbs in (9) are not from the result of 'reversing' the action of the base, nor are they cases of 'intensification' of the bases as we saw earlier. Rather, these meanings appear to be idiosyncratic and fossilized.

Shepardson (1986: 103) notes that most occurrences of the reversive form are in fact fossilized. Although we can discern the reversive affix on many bisyllabic roots/stems, the meaning of the base without the supposed reversive suffix is not clear, because the base cannot stand as a verb without this extension. Examples of this appear in (10).

(10) fossilized forms
    a. kagua   'inspect'   *kaga
    b. fufua   'resurrect'   *fufa
    c. kwangua 'scrape off' *kwanga
    d. zindua   'inaugurate' *zinda
    e. chukua   'take'   *chuka
    f. nyofoa   'nibble'   *nyofa

Although *kagua* 'inspect' is a recognizable verb stem, removal of the /u/ yields **kaga,* which is not recognizable. The same thing is true with the other five verbs exemplified in (10). However, the stems are all polysyllabic, a sign that they are at least historically complex.

To sum up, the reversive is part of the system of Swahili verb derivations. It is realized in four forms *-u, -o, -ul* and *-ol.* This affix does not change the argument structure of the root verb. Two meanings are commonly associated with the reversive: (a) 'undo X,' and (b) 'intensive'. This suffix also appears in numerous fossilized verbs.

4 Elicitation experiments

In this section, we report on two elicitation experiments attempting to determine the availability of the reversive derivation to speakers of Standard Swahili. One experiment was a coinage test, while the second experiment was an acceptability judgment test.

4.1 Methods

For subjects, we recruited 56 undergraduate students at Pwani University in Kenya. All of them were students in the Swahili program. They were from diverse geographical backgrounds and therefore may speak different regional dialects. Most reported that they were bilingual in Swahili and other ethnic languages.

For the coinage test, we created two word lists of 10 nonsense words each. They were either disyllabic or trisyllabic, the most common forms of the roots in Swahili. We also ensured that the final consonant or syllable in the nonsense words would not be confused with existing verb extensions. Each wordlist had words such that all five vowels in Swahili were represented. None of the 20 words were found in dictionaries. All words were presented in the infinitive form to ensure they are viewed as verbs. The infinitive is characterized by the infinitive prefix *ku-*, the root, and the final vowel *-a.*

List A Table 1 contains non-derived verbs. Subjects (N=20) were told to assume that these words exist in Swahili and that if they did not know them, it might be due to the fact that one cannot know every word in the language. With that assumption, they were to write the verb that denoted "undoing" of these actions.

Table 1: Nonsense words — List A

Form presented	Predicted subject response for "undoing" sense
kugola	kugoloa
kupeza	kupezua
kujida	kujidua
kumuka	kumukua
kunaba	kunabua
kukweta	kukwetua
kuika	kuikua
kuoha	kuohoa
kubada	kubadua
kuvupa	kuvupua

List B Table 2 was given to another group of university students (N=16). These were also nonsense words, but contained forms corresponding to the reversive extension. Subjects were instructed to write verbs expressing "reverse actions" from what the presented "verbs" would mean.

For the second elicitation test, a questionnaire of 20 verb roots was administered to another group of university students (N=20) at the same institution, inviting their judgments. The verbs were selected from among the most frequent Swahili verbs identified in the SALAMA of Helsinki Corpus of Swahili (HCS). The corpus is made up of over 20 million words from news texts, books containing prose texts of fiction, education, and science from the second half of the 20th century and the 21st century (Hurskainen 2008; 2009). The information that is parsed by its morphological analyzer includes, among other things, parts of speech, inflections, derivations, and etymology. Using Lemmie2.0, a web-based tool for working with a language corpus in the Language Bank of Finland (CSC 2003), we searched HCS for frequencies of verbs. The search generated a list of verbs ranked from the most frequent. From the list, we created reversive derivations for each of the top 20 verbs by attaching the suffix *-ul*. No glosses were presented for the

Table 2: Nonsense words — List B

Form presented	Predicted subject response for "reverse action"
kukemua	kukema
kutubua	kutuba
kubadua	kubada
kuzepua	kuzepa
kulotoa	kulota
kuisua	kuisa
kuholoa	kuhola
kuyunua	kuyuna
kupapua	kupapa
kulikua	kulika

roots. These are roots that all subjects know. The subjects were asked to make a judgment on each derived verb as to whether it was: (a) an existing word, (b) a possible word, or (c) an impossible word. The verbs and the derived forms presented to the students are in Table 3.

4.2 Results

4.2.1 Coinage test

For List A, out of 200 tokens, 60% of the predicted responses were rendered with the reversive form, observing the appropriate vowel harmony for nonce roots that had /o/. For List B, out of the total number of 160 tokens, 65% of the responses removed the supposed reversive suffix, thus producing the predicted non-reversive forms.

The results of both directions of word formation (adding and subtracting a suffixal piece) show that the reversive is active in the speakers' competence or repertoire. They can create new words with it. The idea of working with nonsense words is consistent with the notion that speakers' competence enables them to understand and produce novel utterances they may have never encountered before (Aronoff & Schvaneveldt 1978). However, the domain on which the reversive can be attached is very restricted. The domain is revealed in Elicitation Test 2, to which we now turn.

4.2.2 Judgment test

Only three derived words are recognized by all as words found in the language. They are *kufunga* 'to shut' → *kufungua* 'to open', *kuchoma* 'to stab' → *kuchomoa* 'to pull out', and *kupanga* 'to arrange' → *kupangua* 'to disarrange.' The constraint for productivity in this case is very clear, namely, since the reversive denotes 'undo X', the only verbs that can be derived with this affix are verbs naming actions that can be undone.

Table 3: Verbs for elicitation test 2

	Base	Derived Verb		Base	Derived Verb
1.	*kuwa* 'to be'	kuwua	11.	*kula* 'to eat'	kulua
2.	*kusema* 'to speak'	kusemua	12.	*kudai* 'to claim'	kudiua
3.	*kufanya* 'to do'	kufanyua	13.	*kuenda* 'to go'	kuendua
4.	*kutoa* 'to remove'	kutoloa	14.	*kupa* 'to give'	kupua
5.	*kuweza* 'to be able'	kuwezua	15.	*kupita* 'to pass'	kupitua
6.	*kutaka* 'to want'	kutakua	16.	*kuja* 'to come'	kujua
7.	*kupata* 'to get'	kupatua	17.	*kujua* 'to know'	kujulua
8.	*kufunga* 'to shut'	*kufungua* 'to open'	18.	*kuchoma* 'to stab'	*kuchomoa* 'to pull out'
9.	*kuanza* 'to start'	kuanzua	19.	*kuomba* 'to beg'	kuomboa
10.	*kuona* 'to see'	kuonua	20.	*kupanga* 'to arrange'	*kupangua* 'to disarrange'

5 Discussion

The domain or set of bases to which the reversive can be attached is restricted (Haspelmath & Sims 2010). It is also very significant that 10 subjects identified *kufanyua* 'to undo' from *kufanya* 'to do' as a possible word. None of the dictionaries consulted have this as an entry. However, a general sense of 'undo' is quite possible, it seems. It can be argued that the reason subjects positively indicated that this is a word in the language reflects their intuition that the reversive is still productive.

Earlier on, we noted that the characterization of the reversive as non-productive was, to a great extent, due to the abundance of lexicalized forms associated with it, i.e. forms which are non-predictable (Shepardson 1986). However, lexicalization is not exclusively found with the reversive and the non-valency-changing suffixes. Consider the following

Table 4: Acceptability judgments

	Base	Derived verb	Existing word	Possible word	Impossible
1.	*kuwa* 'to be'	kuwua	0	0	16
2.	*kusema* 'to speak'	kusemua	1	1	14
3.	*kufanya* 'to do'	kufanyua	6	4	6
4.	*kutoa* 'to remove'	kutoloa	1	1	14
5.	*kuweza* 'to be able'	kuwezua	3	5	8
6.	*kutaka* 'to want'	kutakua	1	1	14
7.	*kupata* 'to get'	kupatua	0	0	16
8.	*kufunga* 'to shut'	*kufungua* 'to open'	16	0	0
9.	*kuanza* 'to start'	kuanzua	3	2	11
10.	*kuona* 'to see'	kuonua	0	1	15
11.	*kula* 'to eat'	kulua	0	1	15
12.	*kudai* 'to claim'	kudaiua	0	1	15
13.	*kuenda* 'to go'	kuendua	1	0	15
14.	*kupa* 'to give'	kupua	0	2	14
15.	*kupita* 'to pass'	kupitua	2	0	14
16.	*kuja* 'to come'	kujua	1	1	14
17.	*kujua* 'to know'	kujulua	0	0	16
18.	*kuchoma* 'to stab'	*kuchomoa* 'to pull out'	16	0	0
19.	*kuomba* 'to beg'	kuomboa	1	0	15
20.	*kupanga* 'to arrange'	*kupangua* 'to disarrange'	16	0	0

three examples with apparent valency-changing affixes, namely the applicative, recipro-cal, and stative, respectively.

(11) a. shika 'hold, catch, apprehend'
       b. shikilia 'hold on tight'

(12) a. ona 'see'
       b. onekana 'be seen'

(13) a. tanda 'spread over'
       b. tandika 'lay a cover, make (a bed)'

These examples reveal that lexicalization is not an exclusive feature of the reversive. Even the most productive verbal suffixes have their share of lexicalized non-composi-tional forms. In (11b), the *-ilia* portion has nothing to do with the meaning or function of the applicative or any prepositional meaning, although it does look phonologically like the applicative. However, it is a form of an intensifier. In (12b), the polysyllabic verb appears as if it bears the stative *-ek* and the reciprocal *-an*. One reading of the stative is 'possibility'. While the stative reading is discernible in the meaning of the word in the sense of 'be visible,' the reciprocal reading is completely absent. What looks like the stative suffix is also found in (13b). However, the relationship between the base (13a) and the seemingly derived form (13b) is not compositional. The stative in its two readings is a valence-reducing affix. Nonetheless, in (13b), it appears to increase the valence by introducing a new object; therefore, it is not the stative we know.

To summarize, in this section we have presented results from two elicitation tests. In the first test, a group of speakers was instructed to coin verbs from nonsense bases. The test revealed that the reversive derivation is productively available to contemporary Swahili speakers. In the second test, speakers were asked to make acceptability judg-ments for reversive verb forms derived from the most frequently used verbs. Only three verbs were acceptable to all the respondents. This points to the semantically restricted domain of application of the rule: it can apply only to verbs expressing actions that can conceivably be reversed.

6 Conclusion

In this article, we set out to describe the reversive derivation in Swahili and to determine its productivity. We identified the suffix and its allomorphs *-u, -o, -ul* and *-ol*. The pro-totypical meaning of the reversive is 'undo X.' However, several words were identified bearing this suffix either without the compositional meaning involving reversion or sep-arative, or without an identifiable base. Those are lexicalized forms. The paper reported on tests of productivity of the extension. It demonstrated that the affix, which causes the new verb to read 'undo X,' where X is the meaning of any root, can only be attached to roots designating reversible actions. Its productivity is semantically constrained, but it

is still available for the creation of new verbs. This was demonstrated by the elicitation tests in which subjects coined nonsense words with this affix.

Abbreviations

Numbers in the glosses refer to the conventional Bantu noun classes.

AP	applicative		PT	past tense
CAUS	causative		REC	reciprocal
FV	final vowel		REV	reversive
OM	object marker		SM	subject marker
PASS	passive		ST	stative

Acknowledgments

References

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark & Roger Schvaneveldt. 1978. Testing morphological productivity. *Annals of the New York Academy of Sciences* 318. 106–114.

Ashton, Ethel O. 1947. *Swahili grammar (including intonation)*. London: Longmans, Green & Co.

Baayen, Harald. 1992. Quantitative aspects of morphological productivity. In Geert Booij & Jaan van Marle (eds.), *Yearbook of morphology 1991*, 109–149. Dordrecht: Kluwer.

Bauer, Laurie. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.

Bauer, Laurie. 2005. Productivity: Theories. In Pavol Stekauer & Rochelle Lieber (eds.), *Handbook of Word-Formation*, 315–334. Dordrecht: Springer.

Bolozky, Shmuel. 1999. *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: Brill.

Corbin, Danielle. 1987. *Morphologie dérivationelle et structuration du lexique*. Tübingen: Niemeyer.

CSC. 2003. WWW. Lemmie2.0: User's manual. https://sui.csc.fi/, accessed 2014-05-09.

Doke, Clement. 1935. *Bantu linguistic terminology*. London: Longman, Green & Co.

Fernández-Domínguez, Jesús. 2013. Morphological productivity measurement: Exploring qualitative versus quantitative approaches. *English Studies* 94(4). 422–447.

Fernández-Domínguez, Jesús, Anna Díaz-Negrillo & Pavol Štekauer. 2007. How is low morphological productivity measured? *Atlantis* 21(1). 29–54.

Guthrie, Malcolm. 1962. The status of radical extensions in Bantu languages. *Journal of African Languages* 1(3). 202–220.

Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding morphology*. 2nd edn. London: Hodder Education.

Hurskainen, Arvi. 2008. *SALAMA dictionary compiler - a method for corpus-based dictionary compilation*. Helsinki: Technical Reports in Language Technology, Report No. 2. http://www.njas.helsinki.fi/salama, accessed 2015-06-14.

Hurskainen, Arvi. 2009. Tips for making queries in Helsinki corpus of Swahili. www.njas.helsinki.fi/salama/impl-appl.html, accessed 2014-05-09.

Lieber, Rochelle. 2009. *Introducing morphology*. Cambridge: Cambridge University Press.

Mdee, James S., Adam Shafi, John G. Kiango & Kimani Njogu. 2009. *Kamusi kamili ya Kiswahili*. Nairobi: Longhorn Publishers.

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* III. 79–121.

Nurse, Derek & Thomas J. Hinnebusch. 1993. *Swahili and Sabaki: A linguistic history*. Berkeley: University of California Press.

Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivations*. Berlin: Mouton de Gruyter.

Plag, Ingo. 2006. Productivity. In Bas Aarts & April MacMohan (eds.), *The handbook of English linguistics*, 537–556. Malden: Blackwell Publishing Ltd.

Polomé, Edgar. 1967. *Swahili language handbook*. Washington, DC: Center for Applied Linguistics.

Rainer, Franz. 2005. Constraints on productivity. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of Word-Formation,* 335–352. Dordrect: Springer.

Schadeberg, Thilo C. 1973. *A sketch of Swahili morphology*. Dordrecht: Foris Publications.

Schadeberg, Thilo C. 2003. Derivations. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 71–89. London: Routledge.

Schroeder, Anne & Susanne Muehleisen. 2010. New ways of investigating morphological productivity. *Arbeiten aus Anglistik und Amerikanistik* 35(1). 43–59.

Shepardson, Kenneth. 1986. *Productivity and the Swahili lexicon*. Bloomington, IN: Indiana University. (Doctoral dissertation.)

Štekauer, Pavol. 2005. Onomasiological approach to wordformation. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of Word-Formation,* 207–232. Dordrecht: Springer.

TUKI. 2001. *Kamusi ya Kiswahili-Kiingereza*. Dar es Salaam: Taasisi ya Uchunguzi wa Kiswahili.

TUKI. 2004. *Kamusi ya Kiswahili Sanifu*. 2nd edn. Oxford: Oxford University Press.

Chapter 15

The anticausative alternation in Luragooli

John Gluckman
UCLA

Margit Bowler
UCLA

We discuss the distribution of the verbal suffix *-Vk* in Luragooli (Luyia, Bantu) based on original fieldwork with a native speaker. We show that *-Vk* patterns like an anticausative marker with respect to a number of different diagnostics, including licensing of theta-roles and interaction with lexical aspect. We compare Luragooli to other languages with anticausative morphology and identify different classes of verbs based on their behavior with the *-Vk* suffix.

1 Introduction

This paper addresses the distribution and meaning of the Luragooli (Luyia, Bantu) verbal suffix *-Vk* (variously realized as *-ek*, *-ik*, *-ok*, and *-uk*).[1] This suffix occurs in a number of Bantu languages, including Chichewa (Mchombo 1993; Simango 2009) and Swahili (Seidl & Dimitriadis 2003). *-Vk* constructions in these languages have variously been called statives, middles, neuter-passives, quasi-passives, anticausatives, and intransitivized constructions, among others (Mchombo 1993; Dubinsky & Simango 1996; Seidl & Dimitriadis 2003; Fernando 2013). For now, we will refer to and gloss the suffix with the neutral term *-Vk*.

The primary goal of this paper is to attain a descriptively adequate account of *-Vk* in Luragooli by addressing the following research questions through original fieldwork with a native Luragooli speaker:

---

[1] Luragooli (also called Maragoli, Logoori, Lulogoori, and Logooli) is a Bantu language in the Luyia subfamily, spoken by approximately 618,000 people in Kenya and Tanzania (Lewis, Simons & Fennig 2015).

- What is the distribution of *-Vk*?

- What meaning(s) is/are associated with the use of *-Vk*?

Based on the answers to these two questions, we suggest that Luragooli *-Vk* should be analyzed as an anticausative suffix. That is, *-Vk* can be treated as a marker of intransitivity analogous to the English anticausative in (1b) below. While in English there is no morphological difference between the causative and anticausative forms of the verb *break* (cf. 1a and 1b), in Luragooli the difference between the two is marked by presence versus absence of the *-Vk* suffix, as in (2a-b):

(1) a. John broke the vase.   (LEXICAL) CAUSATIVE

    b. The vase broke.   ANTICAUSATIVE

(2) Luragooli

    a. Sira a-han-i   muriaŋgo.   (LEXICAL) CAUSATIVE
        Sira 1-close-FV 3door

        'Sira closed the door.'

    b. muriaŋgo gu-han-ik-i.   ANTICAUSATIVE
        3door   3-close-VK-FV

        'The door closed.'

We show that the distribution and use of *-Vk* pattern similarly to cross-linguistic diagnostics for anticausative markers. However, we also present a number of further uses that fall outside of the characteristic anticausative domain. It is therefore a matter of ongoing theoretical research as to whether these functions can be subsumed under the anticausative use.

This paper is organized as follows. We present a brief overview of the core anticausative alternation in §2. The remainder of the paper focuses on how the *-Vk* form differs from "plain" intransitives and from valency decreasing processes like passivization. In §3 we look at what sorts of oblique theta-roles are permitted with each of these three types of intransitive verb stem. In §4 we look at how the *-Vk* suffix interacts with Aktionsart, or lexical aspect, and show that *-Vk* anticausatives correlate with a telic reading of the event. In §5 we detail two sub-classes of *-Vk* intransitives which challenge our typology. §6 concludes the paper.

2 Background on anticausatives

The examples in (1) above show the anticausative alternation in English. The verb *break* can appear in a transitive construction where the PATIENT is a direct object (1a) or in an intransitive construction where the PATIENT is the subject (1b). Cross-linguistically, it

is generally true that a verb like *break* can have both a causative and an anticausative form.[2] Likewise, it is generally true that a verb such as *bloom* (3a) tends not to have a (simple) causative counterpart (3b).

(3) a. The flower bloomed.   ANTICAUSATIVE
     b. *The sun bloomed the flower.   CAUSATIVE

What allows a given verb to have an (anti-)causative counterpart is a matter of ongoing research (Smith 1970; Haspelmath 1993; Levin & Rappaport Hovav 1995; Reinhart 1996; Folli 2002; Folli & Harley 2005; Alexiadou, Anagnostopoulou & Schäfer 2006; Schäfer 2008; among others). One of the core debates concerns the number of event subcomponents associated with each form in (1). One influential proposal is that anticausatives lack a CAUSE semantic sub-event (Haspelmath 1993), and hence, a CAUSER/AGENT argument which brings about the result state. Under this approach, the anticausative verb *break* in (1a) means, essentially, 'the vase BECAME broken', while the causative verb *break* in (1b) contains a CAUSE event: 'John CAUSED the vase to BECOME-broken.'[3]

Other proposals argue that causatives and anticausatives are identical in terms of the number of sub-events, and differ in, essentially, thematic structure determined by factors other than types of sub-events. According to these proposals, both verbs in (1) have the meaning 'CAUSE-break' (or 'CAUSE-BECOME-broken').[4] However, the verbs differ in whether or not they encode an external force which acts to bring about this event, i.e., an AGENT (or INSTRUMENT) of the CAUSE sub-event (Levin & Rappaport Hovav 1995; Schäfer 2008). That is, while (1b) encodes that *John* is the AGENT that brings about the CAUSE event of the vase's breaking, (1a) does not encode reference (explicitly or implicitly) to such an argument.

Our study of Luragooli is consistent with this second hypothesis. We contend that Luragooli intransitive verbs with *-Vk* contain a CAUSE event, but lack an external argument which brings this event about. That said, this paper aims for descriptive coverage; as such, we do not take a strong theoretical stance. Our study starts from the well-known typological observation that languages can differentiate between three classes of verbs

---

[2] Throughout the literature on transitivity alternations, there is a great deal of variation in terminology. Here we follow the terminology of Schäfer (2008; 2009) and Alexiadou, Anagnostopoulou & Schäfer (2015). We use the term *causative (verb)* to refer to any transitive verb which is semantically reducible to CAUSE-verb. We will not go into detail on the various possible causative forms in Luragooli. See the appendix in Bowler & Gluckman (2015a) for an overview of these forms. We use the term *anticausative (verb)* to refer to a non-passive intransitive use of a causative verb. If an anticausative form uses special morphology, we refer to this as a *marked anticausative*. Since we aim to be as neutral as possible in our terminology classification, where possible we will try to use the terms *transitive* instead of *causative*, and *marked* and *unmarked intransitive* instead of *marked* and *unmarked anticausative*. We suggest the reader consult Schäfer (2008; 2009) and Alexiadou, Anagnostopoulou & Schäfer (2015) for a more substantive debate about terminological issues.

[3] The second core debate concerns the derivational relationship between the forms in (1). We will not be concerned with this issue here.

[4] Among other things, this allows for the possibility of a spontaneous internal CAUSE.

that display anticausative alternations, i.e., intransitive "causer-less" forms (Haspelmath 1993; Schäfer 2008):[5]

- **Class I**: Intransitive forms that need a special anticausative marker
- **Class II**: Intransitive forms that cannot have an anticausative marker
- **Class III**: Intransitive forms that can optionally have an anticausative marker

Such a partitioning is also present in Luragooli. Three classes of verbs can be distinguished based on how the intransitive version of an otherwise syntactically transitive verb is morphologically expressed:

- **Class I**: Intransitive forms that must occur with *-Vk*
- **Class II**: Intransitive forms that cannot occur with *-Vk*
- **Class III**: Intransitive forms that optionally occur with *-Vk*

We give examples of each of these verb classes in Table 1.[6]

Table 1: Intransitive verb classes in Luragooli based on the distribution of *-Vk*

Class I (intransitive with *-Vk*)	**Class II** (intransitive without *-Vk*)	**Class III** (intransitive with or without *-Vk*)
kwoneka 'to be destroyed'	*kwigora* 'to open'	*kuhana/kuhaneka* 'to close'
kubameka 'to be flattened'	*kumeeda* 'to increase'	*kwiina/kwiineka* 'to sink'
kuzuganyika 'to be mixed'	*kugomagoma* 'to roll'	*kwoma/kwomeka* 'to dry'
kuharagateka 'to be scraped'	*kumera* 'to grow'	*kuzurula/kuzuruleka* 'to wilt'

Thus, at first glance, *-Vk* seems to pattern as we might expect for an anticausative marker. In §3-4 we review further parallels in Luragooli to anticausative alternations that have been observed cross-linguistically. We begin §3 by introducing the basic distribution of *-Vk* in Luragooli.

---

[5] This three-way classification is reported to reflect a scale of "spontaneity," or how likely it is that the event is perceived as needing an external force to bring it about (Haspelmath 1993). Verbs without the marker are expected to be less likely to require an external effort (that is, they are "internally caused"), while verbs with the marker are perceived as requiring some external force to make the event occur.

[6] The forms in Table 1 bear the infinitival prefix *ku-/kw-* and the *−a* final vowel. For a more complete list of all the verbs of the types discussed in this study, see the appendix in Bowler & Gluckman (2015b).

3 The distribution of -*Vk* in Luragooli

The suffix -*Vk* attaches to certain transitive verbs (roots or stems) to form non-passive intransitives, i.e., anticausatives. For instance, the Luragooli transitive verb *kuhana* 'to close' can appear as a (non-passive) intransitive in two different ways:

(4) causative
    Sira a-han-i   muriaŋgo.
    1Sira 1-close-ꜰᴠ 3door

    'Sira closed the door.' [7]

(5) anticausative

    a. plain intransitive
    muriaŋgo gu-han-i.
    3door   3-close-ꜰᴠ

    'The door closed.'

    b. -*Vk* intransitive
    muriaŋgo gu-han-ik-i.
    3door   3-close-Vᴋ-ꜰᴠ

    'The door closed.'/'The door was closed.'

Example (5a) is consistently translated as 'The door closed.' We refer to this form as the *plain intransitive*. However, (5b) is translated more frequently as 'The door was closed.' We refer to this form as the *-Vk intransitive* or *-Vk form*. Curiously, the English passive translation in (5b) is available despite the fact that the verbal passive suffix -*w* is not present. We take this as initial evidence that -*Vk* makes a semantic contribution in addition to its syntactic contribution of detransitivization. The question that this data raises is exactly how to define this semantic contribution. In the remainder of §3, we investigate this question by looking at how oblique theta-roles interact with intransitives in Luragooli. We conclude that the Luragooli -*Vk* form patterns similarly to what is reported for anticausatives cross-linguistically.

3.1 Diagnostics for oblique theta-roles

Our first set of diagnostics concerns the interaction of the three intransitive classes (Table 1) with oblique theta-roles. Anticausatives interact with theta-roles cross-linguistically in consistent ways. Anticausatives generally do not permit ᴀɢᴇɴᴛs or ɪɴsᴛʀᴜᴍᴇɴᴛs in oblique phrases (Levin & Rappaport Hovav 1995). However, they do tend to license

---

[7] Luragooli has an extremely complex tense/aspect system, which we largely ignore in this paper. We also do not mark any tones. Luragooli is analyzed as having 2 tones (high and non-high). See Samuels & Paster (2015) for a description of the Luragooli (verbal) tonal system.

CAUSERS in oblique phrases (Schäfer 2008).[8] In this they differ from passives, which generally permit AGENTS in oblique phrases, but do not permit INSTRUMENTS or CAUSERS.

For example, German passives generally permit AGENTS (6a) but not CAUSERS (i.e. FORCES) in oblique phrases (6b), while German anticausatives permit CAUSERS (FORCES) but not AGENTS (6b, 6c).[9] Note that German anticausatives have an unmarked and marked form; the latter is accomplished with the reflexive *sich*, as in (6c).

(6)   German (Schäfer 2008)

    a.   passive
       Die Tür   wurde von Peter/??vom Windstoß geöffnet.
       the door was   by Peter/by.the wind.gust opened

       'The door was opened by Peter/??by the gust of wind.'

    b.   unmarked anticausative
       Das Segel zerriss (*von Peter/durch   den Sturm)
       the sail   tore   by   Peter/through the storm.

       (*'The sail tore by Peter.')/'The sail tore from the storm.'

    c.   marked anticausative
       Die Tür   öffnete sich (*von Peter/durch   einen Windstoß).
       the door opened REFL by   Peter/through a   wind.gust

       ('The door opened by Peter.')/'The door opened from a gust of wind.'

Note that the availability of a CAUSER (or FORCE) semantic role is not predicted under the proposal that anticausatives lack a CAUSE event (Haspelmath 1993, among others). This type of data is therefore used by Schäfer (2008), among others, to argue that anticausatives do contain a CAUSE event, but do not encode (in their terms, "license") an AGENT which brings this event about.

The following three sub-sections show how oblique theta-roles combine with the three types of intransitives in Luragooli: passives, plain intransitives, and *-Vk* intransitives. The plain and *-Vk* intransitives pattern similarly with respect to the theta-roles they license in oblique phrases, as we expect from anticausatives.

3.2 Oblique AGENTS

In this section, we determine whether an oblique AGENT (that is, an agentive 'by'-phrase) is permitted with each type of Luragooli intransitive construction. Oblique AGENTS are permitted only with the Luragooli passive (7a). Oblique agents are not permitted with either the *-Vk* intransitive (7b) or the plain intransitive (7c):

---

[8] For descriptions of the relevant theta-roles, see Levin & Rappaport Hovav (1995). Also, the reader should be aware that the properties reviewed here are robust crosslinguistic trends, but are not universally true, even within a single language. In this section, we present these diagnostics merely to establish that both intransitive *-Vk* forms are distinct from a true passive in the relevant respect.

[9] German examples are adapted from Schäfer (2008). For reasons of space, we present a slight oversimplification of the data in that there is variability with respect to which verbs permit which oblique theta-roles. We find the same complexities in Luragooli as well.

(7) a. passive

    muriaŋgo gu-han-w-i   (na Sira).

    3door   3-close-PASS-FV by Sira

    'The door was closed (by Sira).'

    b. plain intransitive

    muriaŋgo gu-han-i   (*na Sira).

    3door   3-close-FV by Sira

    'The door closed (*by Sira).'

    c. *-Vk* intransitive

    muriaŋgo gu-han-ik-i   (*na Sira).

    3door   3-close-VK-FV by Sira

    'The door closed (*by Sira).'

3.3 Oblique CAUSERS

Oblique CAUSERS are permitted with both the plain intransitive (8b) and the *-Vk* intransitive (8c), but not with the passive (8a):

(8) a. passive

    muriaŋgo gu-araminy-w-i (*kutorona na imboza).

    3door   3-open-PASS-FV from   by 9wind

    'The door was opened (*because of/from the wind).'

    b. plain intransitive

    muriaŋgo gu-aram-i (kutorona na imboza).

    3door   3-open-FV from   by 9wind

    'The door opened (because of/from the wind).'

    c. *-Vk* intransitive

    muriaŋgo gu-aram-ik-i   (kuturona na imboza).

    3door   3-open-VK-FV from   by 9wind

    'The door opened (because of/from the wind).'

3.4 Oblique INSTRUMENTS

Oblique INSTRUMENT theta-roles are licensed only by the passive in Luragooli (9a). INSTRUMENTS are not permitted with plain intransitives (9b) or *-Vk* intransitives (9c).[10]

(9) a. passive

    imbwa y-um-iny-w-i   (na itahoro).

    9dog   9-dry-CAUS-PASS-FV PRT 9towel

    'The dog was dried (with a towel).'

---

[10] In general, instrumental subjects are not permitted in Luragooli. The active transitive version of (9a) with *itahoro* 'the towel' as the subject would be ungrammatical.

b. plain intransitive

imbwa y-um-i　　(*na　itahoro).

9dog　9-dry-FV　PRT 9towel

'The dog dried (*with a towel).'

c. *-Vk* intransitive

imbwa y-um-ik-i　　(*na　itahoro).

9dog　9-dry-VK-FV　PRT 9towel

'The dog dried (*with a towel).'

Thus, in terms of oblique theta-roles, the plain intransitive and the *-Vk* intransitive pattern together, separately from the Luragooli passive. This is summarized in Table 2.

Table 2: Theta-role properties of the intransitive constructions

	Passive	Plain intransitive	*-Vk* intransitive
Oblique agents	yes	no	no
Oblique causers	no	yes	yes
Oblique instruments	yes	no	no

The Luragooli patterns in Table 2 largely parallel properties of anticausative versus passive constructions in other languages. The anticausative forms do not permit oblique AGENTS or INSTRUMENTS, but are compatible with oblique CAUSERS.

4 Lexical aspect

In §3 we demonstrated how both plain and *-Vk* intransitives are distinct from the passive. In this section, we will show how *-Vk* intransitives are distinct from the plain intransitive and the passive. The data in §4 concerns lexical aspect, or Aktionsart. We use four pieces of evidence to show that *-Vk* intransitives differ from the other two intransitive forms with respect to lexical aspect. Our evidence involves interaction with negation, complementation under 'want', progressive aspect, and continuations. Our data suggest the provisional generalization in (10), which we will revise in §4.4.

(10) Telicity Restriction

-Vk only attaches to telic predicates. (*to be revised*)

This generalization is consistent with cross-linguistic findings on anticausatives; an interaction with telicity is also reported for anticausatives in other languages (Labelle 1992; Folli 2002; Folli & Harley 2005). Marked anticausatives tend to entail a telic reading of the event denoted by the predicate in Greek (Alexiadou & Anagnostopoulou 2004), Italian (Folli 2002), and French (Zribi-Hertz 1987). For example, in Italian, the marked

anticausative (with the reflexive *si*) cannot occur with a 'for'-temporal phrase (11b), while the unmarked intransitive form can (11a).[11]

(11) Italian

    a. Il   cioccolato è fuso   per pochi secondi/in pochi secondi.
       the chocolate is melted for few   seconds/in few   seconds

    b. Il   cioccolato si   è fuso   *per pochi secondi/in pochi secondi.
       the chocolate REFL is melted for   few   seconds/in few   seconds.

       'The chocolate melted for a few seconds/in a few seconds.' (Schäfer 2009)

We are not aware of a convincing explanation for why such a correlation between anticausatives and telicity should exist. It is not our aim to explain this correlation. Instead, we will merely show that such a pattern is consistent with what we find in Luragooli as well.

Finally, we note that the next two pieces of evidence involving complementation under 'want' and negation are, as far as we know, specific to Luragooli (or perhaps Bantu languages more generally). The negation test is inspired by Dubinsky & Simango's (1996) work on Chichewa. It remains to be seen what the results of such tests are in other languages.

4.1 Complements of 'want'

We first observe a contrast in interpretation when embedding the three intransitives under a verb like *kwenya* 'to want.'[12] We find that with the passive (12a) and plain intransitive (12b) the object of wanting can only be the beginning of the event, not the result state.[13] Conversely, with the *-Vk* intransitive in (12c), the thing that is wanted can only be the result state of the embedded verb. Thus, in a context where the door is already closed, it is infelicitous to use either (12a) or (12b). We take this as evidence that *-Vk* imposes a telicity restriction, i.e. requires a telic predicate; only with the *-Vk* form is the result state entailed.

(12) Context: The door is closed.

    a. passive
      # n-eny-a   murianggo gu-han-w-ɛ.
        1sg-want-FV 3door   3-close-PASS-FV

      'I want the door to be closed.'

---

[11] We refer the reader to the large body of work on Italian anticausatives, in particular Folli (2002), for a full explanation of the data.

[12] The form of the embedded verb in this context is subjunctive, indicated by the final vowel /ɛ/.

[13] The term "beginning of the event" is possibly not quite accurate. For passive and plain intransitive complements of 'want', the object of wanting is perhaps best described as "anything that is not the result state," which includes the beginning, but may also include the middle of the event as well.

   b. plain intransitive

      # n-eny-a   murianggo gu-han-ɛ.

        1sg-want-ꜰᴠ 3door   3-close-ꜰᴠ

      'I want the door to close.'

   c. *-Vk* intransitive

      n-eny-a   murianggo gu-han-ek-ɛ.

      1sg-want-ꜰᴠ 3door   3-close-Vᴋ-ꜰᴠ

      'I want the door closed.'[14]

The plain intransitive and the passive again pattern similarly in that the object of wanting is the movement of the door: 'I want the event of door-closing.' These forms cannot target the result state. In contrast, in (12c), the object of wanting can be either the event of door-closing or the result state: 'I want the state of the door to be closed.' This second reading is not available in (12a) and (13b).

Lastly, we note that at first glance, the data in (12) might be taken to indicate that the *-Vk* form is a stative, as argued in Dubinsky & Simango (1996). However, we observe that (12c) can have the same reading as (12b). That is, the object of wanting can be the event of closing. In other words, *-Vk* intransitives can still be interpreted as eventive. Furthermore, recall that both the plain and *-Vk* intransitive forms permit ᴄᴀᴜsᴇʀ theta-roles, which should be impossible with stative verbs.[15]

4.2 Negation

Our second piece of evidence that *-Vk* intransitives differ from the other two intransitive constructions comes from which parts of the event can be targeted by negation. We find that *-Vk* intransitives only permit the end of the event, i.e. the result state, to be negated, while both passives and plain intransitives permit either the beginning or end of the event to be negated. While less obvious, we think this can also be taken as evidence for a telicity restriction in the *-Vk* form. If the end of the event is entailed by the assertion, then it can be targeted by negation.

Given the context below in (13) in which the door has not moved at all, both the passive (13a) and plain intransitive (13b) are felicitous. Conversely, the *-Vk* intransitive (13c) is infelicitous. Example (13c) is only felicitous if the door moved, but didn't finish closing.

(13)   Context: The door hasn't moved at all.

   a. passive

      murianggo gu-han-w-i   daave.

      3door   3-close-ᴘᴀss-ꜰᴠ ɴᴇɢ

      'The door wasn't closed.'

---

[14] A reviewer asks whether (12c) can be translated as 'I want the door to be closed.' We think this translation is misleading for two reasons: a) it either suggests a passive reading of this sentence, or, b) it suggests a stative reading.

[15] It is unlikely that the *-Vk* form can be treated as an adjective. In Luragooli, a deverbal adjectival form would trigger a different set of agreement (concord) markers than verbal agreement.

b. plain intransitive
murianggo gu-han-i   daave.
3door   3-close-fv neg

'The door didn't close.'

c. *-Vk* intransitive
murianggo gu-han-ek-i   daave.
   3door   3-close-Vk-fv neg

'The door didn't close.'

Out of context, both (13a) and (13b) are ambiguous. They mean that either 'the door didn't start to close' or 'the door didn't finish closing.' That is, (13a) and (13b) can have a reading that the event of the door starting to close – the beginning of the event – didn't occur. However, the *-Vk* intransitive in (13c) is only compatible with a scenario in which the door moved, but didn't get all the way closed. Example (13c) only has the reading that the state of the door being closed didn't occur.[16] We suggest that this follows if the result state is entailed in the *-Vk* form, and so can be targeted by negation. Since there is no such entailment with either the plain or passive form, the result state is not a possible target for negation.

4.3 Progressive aspect

Telic predicates require the culmination of the event that they denote. As a result, we should expect to see an interaction with progressive (grammatical) aspect, since the progressive aspect asserts that the event is on-going, i.e., incomplete, with respect to a reference time. In Luragooli, both the passive (14a) and plain intransitive (14b) forms are compatible with the progressive.[17] In contrast, *-Vk* verb forms are ungrammatical in combination with the progressive aspect, as shown in (14c).

(14) a. passive
mpira gu-toony-w-ang-a.
3ball   3-drop-pass-prog-fv

'The ball was being dropped.'

b. plain intransitive
mpira gu-toony-ang-a.
3ball   3-drop-prog-fv

'The ball is dropping.'

---

[16] This reading is also compatible with the passive and plain forms in Luragooli. This differs from what Dubinsky & Simango (1996) report for Chichewa. We find our Luragooli data curious. It is unclear to us why the start of the event is not a possible target for negation with the *-Vk* form.

[17] A similar set of facts is apparently reported for Greek in Mavromanolaki (2002), as cited in Alexiadou, Anagnostopoulou & Schäfer (2015), although we have not been able to locate this source.

    c. *-Vk* intransitive

        *mpira gu-toony-ik-ang-a.

        3ball   3-drop-Vk-PROG-FV

        ('The ball is being dropped.'/'The ball is dropping.')

The ungrammaticality of (14c) follows if *-Vk* requires that the event culminate — that is, if *-Vk* must combine with a telic predicate, as proposed in (10).

4.4 Continuations

Our last piece of evidence on the interaction of *-Vk* and telicity concerns overt continuations. Related to the negation diagnostic above, we examine the felicity of continuations that deny the result state of an intransitive verb form. We find that continuations of both the passive (15a) and plain intransitive (15b) forms are felicitous if the result state is denied. However, the result state of a *-Vk* intransitive cannot be felicitously denied (15c). This supports the generalization in (10) in that only *-Vk* forms entail that the event culminate. As a result, it is infelicitous to later assert that the event did not culminate.

(15)   a. passive

        maguta ga-diny-iz-w-i   (netare ga-ker-e   ma-doto).

        6butter 6-harden-CAUS-PASS-FV but   6-be.still-FV 6-soft

        'The butter was hardened (but it's still soft).'

     b. plain intransitive

        maguta ga-diny-i   (netare ga-ker-e   ma-doto).

        6butter 6-harden-FV but   6-be.still-FV 6-soft

        'The butter hardened (but it's still soft).'

     c. *-Vk* intransitive

        maguta ga-diny-ik-i   (#netare ga-ker-e   ma-doto).

        6butter 6-harden-Vk-FV but   6-be.still-FV 6-soft

        'The butter hardened (#but it's still soft).'

In (15a) and (15b) we get a reading in which the butter has hardened somewhat, but still remains soft. However, (15c) is infelicitous if it is later asserted that the butter hasn't completed the hardening process. This follows if *-Vk* is required to attach only to telic predicates that denote a culminated event (i.e., telic predicates).

Thus, for contexts targeting lexical aspect, the *-Vk* form patterns distinctly from the passive and plain forms. These data suggest that *-Vk* requires that the event of the verb culminate, supporting the telicity generalization in (10). We summarize the aspectual properties of the Luragooli passive, plain intransitive, and *-Vk* intransitive in Table 3.

However, there are a number of counterexamples in Luragooli to the telicity generalization in (10). Not all verbs pattern similarly with respect to the four tests above. For instance, the *-Vk* form of *kwoma* 'to dry' fails the four diagnostics in Table 3. Given a *-Vk* form of *kwoma* 'to dry,' the object of wanting cannot be the result state (16a) of the

Table 3: Lexical aspect properties of passive, plain intransitive and *-Vk* intransitive

	Passive	Plain intransitive	*-Vk* intransitive
Negation	entire event	entire event	result state
'want'	entire event	entire event	result state
Progressive	grammatical	grammatical	ungrammatical
Continuations	can deny result state	can deny result state	cannot deny result state

event described by the verb. Negation can target the beginning of the event as well as the result state (16b). The *-Vk* form is compatible with progressive aspect (16c). Finally, a continuation that denies that result state is felicitous (16d).

(16)   a.   complement of *want*
           n-eny-a   imbwa y-um-ik-e.
           1sg-want-ꜰᴠ 9dog   9-dry-Vᴋ-ꜰᴠ
           'I want the dog to be dry.'
           Consultant's comment: Strange if the dog is already dry.

       b.   negation
           imbwa y-um-ek-i   daave.
           9dog   9-dry-Vᴋ-ꜰᴠ ɴᴇɢ
           'The dog didn't dry.'
           Consultant's comment: OK if the dog doesn't dry at all, or only dries halfway.

       c.   progressive
           imbwa y-um-ek-a.
           9dog   9-dry-Vᴋ-ꜰᴠ
           'The dog is drying.'

       d.   continuation
           imbwa y-um-ik-i   (netare i-ker-e   i-nzilu hadi).
           9dog   9-dry-Vᴋ-ꜰᴠ but   9-be.still-ꜰᴠ 9-wet   some.of
           'The dog dried (but it's still a little wet).'

That said, these exceptions do not necessarily argue against a treatment of *-Vk* as an anticausative marker, since such variation is consistent with what is observed cross-linguistically. Schäfer (2008) argues convincingly that the telicity restrictions for Greek, Italian, and French fail for a number of lexical items, and are not consistent across languages. For instance, in Italian, some marked anticausatives permit modification by 'for'-temporal adverbial phrases, and so are not necessarily telic.

(17) Italian (Schäfer 2008)
La temperatura si è alterata per due ore.
the temperature REFL is altered for two hours
'The temperature altered for two hours.'

While there is an overall tendency for anticausatives to co-occur with a telic read-ing of the predicate, telicity is not an absolute requirement for morphologically marked anticausatives. It is still an open question as to why some telicity diagnostics fail with certain Luragooli *-Vk* verbs. Our hypothesis, adopted from Schäfer (2008), is that there is something inherent about the semantics of the verb root itself that leads to the failure of a particular diagnostic. Further in-depth examination of lexical classes in Luragooli (along the lines of Levin & Rappaport Hovav 1995 and Haspelmath 1993) are needed to tease apart these differences.

5 Subclasses of Class I verbs

In §3 and 4 we provided evidence that Luragooli *-Vk* intransitives generally pattern dis-tinctly from the passive and the plain intransitive forms in terms of theta-roles and lex-ical aspect. The accumulated evidence led us to conclude that *-Vk* is the anticausative marker in Luragooli. Our conclusion was based largely on a comparison with cross-lin-guistic observations. In this section, we detail some "anomalous" uses of *-Vk* that fall out-side of what is typically associated with an anticausative alternation cross-linguistically.

We begin by designating two additional subclasses of Class I, that is, verbs which require *-Vk* to form non-passive intransitives: Class Ia and Class Ib.[18] These classes are differentiated based on semantic criteria.[19]

- **Class Ia**: Verbs that (loosely) denote an epistemic state, i.e., that license a mental experiencer argument.
- **Class Ib**: Verbs that have an affected argument. (We will return shortly to what we mean by "affected.")

Examples of verbs in these classes are shown in Table 4.

The Class Ia *-Vk* intransitives are productively formed with any verb that takes an experiencer subject. They pattern separately from the passive in not being able to occur with an oblique "demoted" subject (18b).[20] Passives, however, are acceptable with an EXPERIENCER subject that is expressed obliquely (18a).

---

[18] We thank a reviewer for helping us with the overall classification of the verbs.

[19] As far as we know, there is one exceptional verb, *kunwa* 'to drink.' The (true) passive of this verb is expressed with the *-Vk* form *kunwahuka* 'to be drunk.' This verb must be listed as an idiosyncratic exception.

[20] However, an oblique argument is sometimes licensed in the presence of *-Vk* with the addition of the recip-rocal *-an*. Such facts have also been reported for Chichewa and Swahili (Dubinsky & Simango 1996; Seidl & Dimitriadis 2003).

Table 4: Non-canonical anticausative verb classes

	Transitive	Intransitive with *-Vk*	Passive
Class Ia	*kuhola* 'to hear'	*kuholeka* 'to be heard'	*kuholwa* 'to be heard'
	kurora 'to see'	*kuroreka* 'to be seen'	*kurorwa* 'to be seen'
	kudiira 'to touch'	*kudiirika* 'to be touched'	*kudiirwa* 'to be touched'
	kumena 'to taste/lick'	*kumeneka* 'to be tasted/licked'	*kumenwa* 'to be tasted/licked'
Class Ib	*kuhola* 'to punch'[a]	*kuholeka* 'to be punched'	*kuholwa* 'to be punched'
	kurasa 'to throw'	*kurasika* 'to be thrown'	*kuraswa* 'to be thrown'
	kuroomba 'to make'	*kuroombika* 'to be made'	*kuroombwa* 'to be made'
	kulia 'to eat'	*kuliika* 'to be eaten'	*kuliwa* 'to be eaten'
	kunyanya 'to chew'	*kunyanyeka* 'to be chewed'	*kunyanywa* 'to be chewed'

[a] In Luragooli, 'to hear' and 'to punch' are entirely homophonous (*kuhola*), with no tonal differences.

(18) a. passive

iroli   i-ror-w-e   na Sira.
9truck 9-see-PASS-FV by 1Sira

'The truck was seen by Sira.'

b. *-Vk* intransitive

iroli   i-ror-ek-e   (*na Sira).
9truck 9-see-VK-FV   by 1Sira

'The truck was seen (by Sira).'

The only commonality that we can identify among Class Ib verbs is a notion of "affectedness." Class Ib transitive verbs all involve an affected object argument.[21] Things that are 'punched,' 'thrown,' 'made,' 'eaten,' and 'chewed' are affected in a broad sense.

---

[21] Dubinsky & Simango (1996) make a similar claim for *-Vk* in Chichewa.

However, a verb like *kwomba* 'to sing' does not have a form with *-Vk* (**kwombeka*), presumably because songs are not affected by the action of singing.[22]

The Class Ia and Class Ib verbs are a prima facie problem for our analysis of *-Vk* as an anticausative marker; these classes of verbs are not generally reported to have anticausative forms in other languages. Moreover, it is unclear how the diagnostics concerning thematic roles and lexical aspect are applicable to the Class Ia verbs, some of which seem to be inherently stative and non-agentive/non-causative. A potential way to incorporate these verbs into the more general analysis of anticausativization in Luragooli would be to appeal to Beavers's (2011) criteria for affected objects.[23] Objects of the Class Ib verbs can be thought of as being "physically impinged on" to some extent. We could possibly extend this to the experiencer verbs in Class Ia by assuming that experiencer subjects are also (mentally) impinged on. Thus, the descriptive generalization is that *-Vk* attaches to any verb that takes an affected ("impinged") argument, in the sense of Beavers (2011). This generalization subsumes canonical anticausative verbs (e.g. *break* and *melt*) as well, since these verbs also involve affected arguments: the PATIENT. We find this a promising avenue for further research, but we must leave it open for now.

6 Conclusion

In this paper, we have shown that the Luragooli morpheme *-Vk* has a wide distribution. While coinciding nicely with what we expect from an anticausative morpheme, as documented in §3 concerning theta-roles, and §4 concerning lexical aspect, §5 has shown that *-Vk*'s range extends beyond what are canonically seen as anticausative environments. Further investigation of the semantics of Classes Ia and Ib should provide a clearer picture as to what governs the distribution of *-Vk*.[24] Nonetheless, we do not view this exceptional data as an insurmountable obstacle to our proposal. Even in Romance and Germanic languages, the 'anticausative' morpheme does not solely mark anticausatives: it is also the reflexive morpheme. Having an anticausative marker that does double-duty with other functions is therefore not cross-linguistically unusual. Still, the Luragooli data suggest that more in-depth cross-linguistic research would be beneficial to our understanding of anticausatives in general, since the majority of in-depth work on anticausatives has been done for western European languages.

---

[22] The Class Ib verbs might all be classified as change of state verbs, although it requires us to loosen the definition of change of state considerably. See Dubinsky & Simango (1996) for discussion of change of state and *-Vk* in Chichewa.

[23] We thank an anonymous reviewer for pointing out the relevance of this work to us.

[24] *-Vk* forms have been reported to mean 'V-able' in Chichewa (Simango 2009) and Kikongo (Fernando 2013). This reading does not seem to be present with *-Vk* for our consultant, although further investigation is required to settle the matter. We further note that treating *-Vk* as a marker of a middle voice (e.g., 'This cheese cuts easily') is not straightforwardly possible.

Acknowledgments

Abbreviations

Luragooli has 20 noun classes. Following Bantuist convention, we mark noun classes via numerals at the beginning of nouns and verbs.

CAUS	causative	PASS	passive
DEM	demonstrative	PRT	particle
FV	final vowel	PLACT	plural act
NEG	negative	REFL	reflexive
PROG	progressive		

References

Alexiadou, Artemis & Elena Anagnostopoulou. 2004. Voice morphology and the causative-inchoative alternation: Evidence for a non-unified structural analysis of unaccusatives. In Artemis Alexiadou, Elena Anagnostopoulou & Martin Everaert (eds.), *The Unaccusativity Hypothesis*, 114–136. Oxford: Oxford University Press.

Alexiadou, Artemis, Elena Anagnostopoulou & Florian Schäfer. 2006. The properties of Anti-Causatives crosslinguistically. In Mara Frascarelli (ed.), *Phases of interpretation*, 187–211. Berlin: Mouton de Gruyter.

Alexiadou, Artemis, Elena Anagnostopoulou & Florian Schäfer. 2015. *External arguments in transitivity alternations: A layering approach.* Oxford: Oxford University Press.

Beavers, John. 2011. On affectedness. *Natural Language and Linguistic Theory* 29. 339–370.

Bowler, Margit & John Gluckman. 2015a. A compilation of verbal alternations in Luragooli. http://www.jgluckman.com/uploads/9/4/6/4/9464781/appendix_gluckman_bowleracal2015.pdf.

Bowler, Margit & John Gluckman. 2015b. Existential quantification in Logoori: Distribution and semantics of *ku*. (Paper presented at the World Congress of African Linguistics 8, University of Kyoto, 21-24 August 2015.)

Dubinsky, Stanley & Silvester R. Simango. 1996. Passive and stative in Chichewa: Evidence for modular distinctions in grammar. *Language* 72(4). 749–781.

Fernando, Mbiavanga. 2013. *The causative and anticausative alternation in Kikongo (Kizombo)*. Stellenbosch: Stellenbosch University. (Doctoral dissertation.)

Folli, Raffaella. 2002. *Constructing telicity in English and Italian*. Oxford: University of Oxford. (Doctoral dissertation.)

Folli, Raffaella & Heidi Harley. 2005. Flavors of v. In Paula Kempchinsky & Roumyana Slabakova (eds.), *Aspectual inquiries*, 95–120. Dordrecht: Springer.

Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity* (Studies in Language Companion Series 23), 87–120. Amsterdam: John Benjamins.

Labelle, Marie. 1992. Change of state and valency. *Journal of Linguistics* 28. 375–414.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical semantic interface*. Cambridge, MA: MIT Press.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the world*. 18th edn. Dallas, TX: SIL International. http://www.ethnologue.com.

Mavromanolaki, Georgia. 2002. *Unaccusative verbs, transitivity alternations and aspect*. Crete: University of Crete. (Master's thesis.)

Mchombo, Sam. 1993. A formal analysis of the stative construction in Bantu. *Journal of African Languages and Linguistics* 14(5). 5–28.

Reinhart, Tanya. 1996. *Syntactic effects of lexical operations: Reflexives and unaccusatives* (OTS Working Papers in Linguistics). Utrecht: Utrecht Institute of Linguistics, Utrecht University. http://www.let.uu.nl/~tanya.reinhart/personal/, accessed 2016-07-21.

Samuels, Alex & Mary Paster. 2015. Verbal tone in Logoori. (Paper presented at the Workshop on Luyia Bantu Languages at ACAL 46, Eugene, Oregon, 26-28 March, 2015.)

Schäfer, Florian. 2008. *The syntax of (anti-)causatives* (Linguistik Aktuell 126). Amsterdam: John Benjamins.

Schäfer, Florian. 2009. The causative alternation. *Language and Linguistics Compass* 3(2). 641–681.

Seidl, Amanda & Alexis Dimitriadis. 2003. Statives and reciprocal morphology in Swahili. In Patrick Sauzet & Anne Zribi-Hertz (eds.), *Typologie des langues d'Afrique et universaux de la grammaire*, 239–284. Paris: L'Harmattan.

Simango, Silvester R. 2009. Causative disguised as stative: The affix *il/ek* in ciCewa. *South African Linguistics and Applied Language Studies* 27(2). 121–134.

Smith, Carlota. 1970. Jespersen's 'move and change' class and causative verbs in English. In Mohammad A. Jazayery (ed.), *Linguistic and literary studies in honor of Archibald A. Hill*, 101–109. Berlin: Mouton de Gruyter.

Zribi-Hertz, Anne. 1987. La réflexivité ergative en français moderne. *Le Français Moderne* 55. 23–54.

Chapter 16

The locative applicative and the semantics of verb class in Kinyarwanda

Kyle Jerro

University of Texas at Austin

This paper investigates the interaction of verb class and the locative applicative in Kinyarwanda (Bantu; Rwanda). Previous analyses of applicative morphology have focused almost exclusively on the syntax of the applied object, assuming that applicativization adds a new object with a transparent thematic role (e.g. Kisseberth & Abasheikh 1977; Baker 1988; Bresnan & Moshi 1990; Alsina & Mchombo 1993; McGinnis 2001; Jeong 2007; Jerro 2015, *inter alia*). I show instead that the interpretation of the applied object is contingent upon the meaning of the verb, with the applied object having a PATH, SOURCE, or GOAL semantic role with motion verbs from different classes. The general LOCATIVE role discussed in previous work appears with non-motion verbs. I outline a typology of the interaction of the locative applicative with four different verb types and provide a semantic analysis of applicativization as a paradigmatic constraint on the lexical entailments of the applicativized variant of a particular verb.

1 Introduction

The applicative has been traditionally analyzed as a valency-increasing morpheme which adds a new object and an associated thematic role to the argument structure of a given verb (see Dixon & Aikhenvald 1997 and Peterson 2007 for a typological overview of valency-changing morphology cross-linguistically). An example from Kinyarwanda (Bantu; Rwanda) is given in (1b), with the applicative morpheme *-ir*:[1]

(1) a. Umu-gabo a-ra-ndik-a   in-kuru.
         1-man   1S-PRES-write-IMP 9-story

         'The man is writing the story.'

     b. Umu-gabo a-ra-ndik-ir-a   umw-ana in-kuru.
         1-man   1S-PRES-write-APPL-IMP 1-child 9-story

         'The man is writing the child the story.'

---

[1] All Kinyarwanda data in this paper come from elicitations conducted by the author.

Whereas the transitive verb *kwa-ndika* 'write' in (1a) licenses a subject and one object, the applicativized variant in (1b) has two post-verbal NPs.[2] Applicative morphology often licenses objects that are assigned one of various different thematic role types, such as BENEFICIARY, REASON/MOTIVE, and LOCATIVE. In the case of (1b), the applied object *umwana* 'child' is a BENEFICIARY.

A majority of the literature on applicative morphology has focused on the syntactic properties of the applied object, examining whether the thematic object and the applied object have equal access to objecthood diagnostics like passivization and object marking on the verb. These approaches assume that applicativization transparently adds a new object participant with a specific thematic role to the verb's argument structure. On this view, the applicative morpheme lexically or syntactically licenses a new object that may or may not block — contingent upon the language — the thematic object from having its default object properties (see §2 for an overview of the literature on object symmetry).

However, work outside Bantu has shown that the meaning of particular verb classes affects argument realization patterns (Fillmore 1970; Levin 1993; Rappaport Hovav & Levin 2008; Beavers 2011, *inter alia*). The hypothesis for languages with applicatives is that verb meaning should affect which thematic roles can assigned to the applied object. In this paper I show that verb class indeed affects the argument realization of the applicative morpheme, and I outline a four-way typology based on the nature of location denoted by the base verb (see Sibanda 2016 (this volume) for evidence for the interaction between verb class and applicative morphology in Ndebele). Furthermore, I suggest that applicativization is not a productive operation of argument addition. Instead, I analyze applicativization as a paradigmatic constraint which requires that the lexical entailments associated with a particular argument in the predicate be stricter than those in the corresponding non-applicativized predicate.

In the next section, I discuss previous approaches to applicative morphology. In §3 I show that there is variation in the use of the locative applicative in Kinyarwanda, and in §4 I show that these uses are tied to four separate verb classes. In §5 I outline an analysis based on the lexical entailments of the verb which accounts for both the traditional use of an applicative as well as the semantic uses described in this paper. In §6 I conclude the discussion.

2 Previous approaches

Previous work on applicative morphology has looked almost exclusively at the syntactic nature of the applied object in relation to the thematic object (i.e. the object licensed by the verb). The mainstay of research on applied objects has looked at the syntax of the applied object, analyzing applicativization as an operation that adds an object to the argument structure of the verb (Baker 1988; Bresnan & Moshi 1990; Alsina 1992; Alsina

---

[2] Bantu applicative morphemes usually immediately follow the verb stem, and the vowel is often conditioned by the preceding stem vowel. For Kinyarwanda, when the vowel in the preceding syllable is mid (i.e. [e] or [o]), the allomorph is *-er*, else it is *-ir*. Because *-ir* is the more general form, I use it as the citation for the duration of this paper.

& Mchombo 1993; Marantz 1993; Pylkkänen 2008; McGinnis 2001; McGinnis & Gerdts 2003; Baker & Collins 2006; Zeller 2006; Zeller & Ngoboka 2006; Peterson 2007; Jeong 2007; Baker, Safir & Sikuku 2012; Jerro 2015).

However, these approaches do not capture the full empirical range of uses of applicative morphology. For example, cases exist in which the applicative affects the meaning of the thematic object instead of adding a new syntactic object (Marten 2003; Creissels 2004; Cann & Mabugu 2007; Bond 2009). For example, Marten (2003) notes a use of the applicative in Swahili that indicates a pragmatically noteworthy property of the verbal object, as in (2b) where there is a pragmatically salient property that is absent with the non-applicativized variant.

(2) Swahili

    a. Juma a-li-va-a   kanzu.
       Juma 1S-PST-wear-FV kanzu

       'Juma was wearing a Kanzu.'

    b. Juma a-li-val-i-a   nguo   rasmi.
       Juma 1S-PST-wear-APPL-FV clothes official

       'Juma was dressed up in official/formal clothes.'

    c. #Juma a-li-val-i-a   kanzu.
       Juma   1S-PST-wear-APPL-FV kanzu

       Intended: 'Juma was wearing a Kanzu.' (Marten 2003: 9)

In (2b), there is no additional object that is not present with the non-applicativized verb in (2a); the only difference between the two is that (2b) provides additional pragmatic information about the object. Specifically, Marten argues for what he terms *concept strengthening*, where the applicative is used to make a narrower claim about the object of the sentence than in the non-applied sentence. This use of the applicative is a semantic/pragmatic use which lies outside of the standard analysis of applicativization as an object-adding operation.

A separate literature on the typology of directed motion has argued that applicatives in certain Bantu languages can be used to license GOALS (Schaefer 1985; Sitoe 1996). For example, consider the data in (3a) and (3b) from Setswana (Bantu; Botswana); in (3a), there is a locative phrase which describes a general location. In (3b), the presence of the applicative *-ẹl* indicates a GOAL reading of the locative that is not present in (3a).[3]

(3) Setswana

    a. mò-símàné ó-kíbítl-à   fá-tlàsé   gá-dì-tlhàrè.
       1-boy   1S-run.heavily-IMP NEARBY-under LOC-8-tree

       'The boy is running with heavy footfall under the trees.'

---

[3] The glosses and English translations of the data in (3) are copied from the original examples in Schaefer (1985).

b. mò-símàné ó-kíbítl-ȩ̀l-à   kwá-tlàsé   gá-thàbà.
   1-boy   1S-run.heavily-to-IMP DISTANT-under LOC-mountain

'The boy is running with heavy footfall to under the mountain.' (Schaefer 1985: Tables VI-VII)

Contra analyses where the locative applicative is assumed to add an object that is assigned a LOCATIVE thematic role, data such as that in (3) suggest that the semantic role of the locative applicative is not always a general LOCATIVE.

Few studies have investigated the effect of the semantics of verb class on the realization of applicative morphology. An exception is Rugemalira (1993), who gives a detailed account of the interaction of locatives with 500 different verbs in Runyambo (Bantu; Tanzania).[4] Rugemalira describes a four-way typology of locatives with different verbs: verbs that require an applicative to license a location, verbs that disallow the applicative with a locative phrase, verbs where the applicative changes the interpretation of the location, and verbs where the applicative and locative prefix are in complementary distribution.[5] The interaction with verb meaning has not been central in work since Rugemalira's dissertation. In this paper, I bring this perspective back to the fore, exploring the interaction of verb meaning and locative applicatives in Kinyarwanda.

3 Typology of locative meanings

In this section, I describe four kinds of locative meanings that may be added to a verb by the applicative: LOCATIVE, GOAL, PATH, and SOURCE. I employ different morphosyntactic tests to motivate the syntactic and semantic differences among the uses. I assume a typology of motion where a complete motion event involves an agent moving from a SOURCE, along a PATH, and ending at a GOAL (cf. Talmy 1975; Slobin 1996; Zlatev & Yangklang 2004; Beavers, Levin & Tham 2010, *inter alia*).[6] Of course, a particular motion event may not syntactically license all of these elements simultaneously, and — as I show below — different verbs in Kinyarwanda categorize syntactically and/or semantically for different components of the motion event.

The first category is verbs where the applicative adds a general LOCATIVE role, i.e. the location where the event took place.[7] For the verb *ku-vuga* 'to talk', the applicative is obligatory in (4b) for licensing the argument with the LOCATIVE role.

---

[4] See also Cann & Mabugu (2007) for discussion of the interaction of verb class and applicative in Shona and Sibanda (2016, this volume) for a related discussion of applicatives in Ndebele.

[5] Unlike the present study, Rugemalira rejects the notion that there are semantically-defined verb classes such as 'motion verbs'. The data below from Kinyarwanda do suggest unity of verbs across classes.

[6] My use of the term *path* here is what Asher & Sablayrolles (1995) refer to as *strict internal path*, i.e. the portion of a path which does not include the SOURCE and GOAL.

[7] This applicative form contrasts with the locative applicative described in Kimenyi (1980); Zeller (2006); Zeller & Ngoboka (2006), i.e. *–ho*. For all the speakers I have consulted, *-ir* is the locative applicative, while *–ho* is one of a class of locational clitics (cf. §4).

(4) a. Yohani a-ri   ku-vug-a.
         John   1S-be inf-talk-imp
         'John is talking.'

     b. Yohani a-ri   ku-vug-ir-a   mu n-zu.
         John   1S-be inf-talk-appl-imp 18   9-house
         'John is talking in the house.'

In (4a), the verb *ku-vuga* 'to talk' is intransitive, while in (4b), there is a new argument that is licensed by the applicative morpheme. Note that locative applicative sentences differ from other sentences with applicatives in that the Locative applied object often requires a locative prefix *mu*.[8] Other applied objects in Kinyarwanda, such as Beneficiaries in (1b), do not require any prefixes. Crucially, the locative prefixes found throughout this paper are class markers and not prepositions, contrary to the Kinyarwanda orthographic convention of writing the prefix separately (as demonstrated in Appendix A). Thus for the duration of the paper I treat phrases which are preceded by locatives such as *ku* and *mu* as arguments and not obliques.

In the second type, the applicative adds a Goal to the event described by the verb. This appears with verbs such as *kw-iruka* 'to run', *gu-tembera* 'to go about', *ku-jya* 'to go', and *gu-simbuka* 'to jump'.

(5) a. Yohani a-ri   kw-iruk-a.
         John   1S-be inf-run-imp
         'John is running.'

     b. Yohani a-ri   kw-iruk-ir-a   kw' i-soko.
         John   1S-be inf-run-appl-imp 17   5-market
         'John is running to the market.'

In (5b), the new location licensed by the applicative is not a general description of where the event took place, but rather the Goal of the running event.

Third, the applied object may be a Path, as in (6b). Unlike the verb *ku-vuga* 'to talk' in (4), the non-applicativized variant of the verb *kw-injira* 'to enter' permits a Locative object in the non-applicativized variant. Other verbs that pattern like *kw-injira* 'to enter' are *gu-sohoka* 'to exit', *ku-manuka* 'to descend', *kuzamoka* 'to ascend', and *ku-rira* 'to climb'.

(6) a. N-di   kw-injir-a   mu n-zu.
         1sg-be inf-enter-imp 18   9-house
         'I am entering the house.'

     b. N-di   kw-injir-ir-a   mu n-zu   mu mu-ryango.
         1sg-be inf-enter-appl-imp 18   9-house 18   3-door
         'I am entering the house through the door.'

---

[8] Though for some verbs, the locative prefix is omitted. I assume that whether the applied object is marked with a locative varies on a verb-by-verb basis. In fact, Rugemalira (2004) shows that there is considerable variation across and within languages as to whether the locative prefix is required.

Here, the applied object describes the PATH through which the motion event occurs. Note that in (6b), the applicative is obligatory.

Finally, the applied object may be a SOURCE, as in (7) where the applicative attaches to the verb *kw-ambuka* 'to cross'.

(7) a. Karemera y-∅-ambuts-e   in-yanja.
     Karemera 1S-PST-cross-PERF 9-ocean

     'Karemera crossed the ocean.'

   b. Karemera y-∅-ambuk-iy-e   i   Mombasa (mu) n-yanja.
     Karemera 1S-PST-cross-APPL-PERF 23 Mombasa 18   9-ocean

     'Karemera crossed the ocean from Mombasa.'

In this example, the applied object is obligatorily interpreted as the SOURCE of the motion event.[9]

3.1 Evidence for the typology

3.1.1 Interpretive differences

One indication of the differences between the applicativized and non-applicativized variants is the interpretive difference of the locational phrase in the applicativized and non-applicativized sentences. For example, consider the following context: Karemera is cooking, and he is talking about needing to run back to the store to get some things he forgot. I leave the room, but when I get back, and he has gone. In this context, I could ask the question in (8), to which an interlocutor could respond with (9a).

(8) [Ikibazo]: Karemera y-a-gi-ye   he?
     problem: Karemera 3S-PST-go-PERF where

     'Question: Where did Karemera go?'

(9) a. Y-iruk-iy-e   kw' isoko.
     1S-run-APPL-PERF 17   store

     'He ran to the store.'

   b. #Y-irutse-e   kw' isoko.
     1S-run-PERF 17   store

     ('He ran to the store.') (on intended reading)

In this context it is infelicitous to use (9b), where the locative is understood as describing the general location of the running event (e.g. a context where someone is running inside of a store), and not the GOAL of the subject's motion. The sentence in (9a), on the other hand, describes the GOAL to which the running event is directed.

Positional verbs also have a GOAL object with the locative applicative, where the applied object is understood as the place where the subject is intending to sit. Locative

---

[9] The perfective morpheme *-(y)e* often has phonological ramifications for the final consonant of the stem. Here, the infinitive *kw-ambuka* 'to cross' changes to *-ambuts* when the perfective suffix is present.

phrases can be used with both variants, but the interpretations are crucially distinct. The non-applicativized variant in (10a) is a general LOCATIVE, describing where the subject is sitting, while in (10b) it describes the sub-location on which the subject is sitting (by accident).[10]

(10) a. N-icay-e   ku ma-zi.
          1sGS-sit-PERF 17 6-water
          'I sat in the water.' (e.g. in a lake or a pool)

      b. N-icar-iy-e   ama-zi.
          1sGS-sit-APPL-PERF 6-water
          'I sat in the water.' (e.g. a puddle of water on a bench after it rained)

The data in this section show that while in certain cases a locative can appear with both a base verb and applicativized verb (though this property is verb-specific; see Appendix A), the interpretation of the location differs between the two.

3.1.2 Locational clitics

Another diagnostic for making precise the different locations that are selected for by different verbs with and without locative applicatives is the locative clitic. Kinyarwanda has three locative clitics that replace locative phrases (for discussion on cognate clitics in other languages, see Diercks (2011) for Lubukusu and Simango (2012) for Chicheŵa). Crucially, (and distinct from the function of locative clitics in Chicheŵa), the locative clitics are in complementary distribution with the locative object, behaving similarly to a pronoun.

To date there is no detailed semantic discussion of the meanings of the locative clitics in Kinyarwanda, but the intuitive definitions in Table 1 are suitable for the current discussion. I assume that these three clitics correspond to the class 16, 17, and 18 locative class prefixes, cf. Appendix A.

Table 1: Kinyarwanda locative clitics

Clitic	Meaning	Class
=ho	at or on something	16
=yo	at or to a place	17
=mwo/mo	inside of something	18

The use of a locative clitic is conditioned by two factors. First, the clitic replaces a locative phrase that is selected for by the verb (or applicative) and behaves as a syntactic object. Second, the semantics of the clitic must be compatible with the specific kind of

---

[10] Consultants have noted that the applicativized positional verbs indicate a GOAL that is construed as being accidentally sat upon. I treat these verbs as a subclass of the manner of motion verbs (see below), but with this class there is an implicature that arrival at the GOAL is accidental.

motion conveyed in the sentence. For example, in (11a) the locative clitic *=yo* is not licit since there is no GOAL selected for by the non-applicativized version of the verb *kw-iruka* 'to run'.

(11) KWIRUKA : to run

    a. *N-iruts-e=yo.
        1SG-run-PERF=17.LOC

        ('I ran there.')   *GOAL

    b. N-iruts-e=mwo.
        1SG-run-PERF=18.LOC

        'I ran inside of somewhere.' (e.g. a house)   General Location

(12) KWIRUK-IR-A : to run to

    a. N-iruk-iy-e=yo.
        1SG-run-APPL-PERF=17.LOC

        'I ran (to) there.'   GOAL

    b. N-iruk-iy-e=mwo.
        1SG-run-APPL-PERF=18.LOC

        'I ran into there.'   GOAL

In (12a), however, the locative clitic is permissible because the applicativized verb selects for a GOAL locative object (cf. (9), above). The clitic *=mwo* is permissible for both applicativized and non-applicativized variants of the verb *kw-iruka* 'to run' but, crucially, with different interpretations. With the non-applicativized verb in (11b), the clitic is a general location inside which the event is taking place (e.g. inside a house). In (12b), on the other hand, the clitic is the location into which the subject moves (i.e. the GOAL).

The locative clitic is sensitive to whether a verb permits a locative object and the kind of location that that object describes. The data in (11) and (12) show that the verb *kw-iruka* 'to run' optionally allows a general LOCATIVE, but with the applicative, the locative object is a GOAL.

Another example is with the verb *kw-injira* 'to enter', where the referent of the locative clitic differs in interpretation between the applicativized and non-applicativized verbs. With the bare verb, *=mwo* refers to the GOAL (13a), while with the applicativized variant the clitic refers to the PATH through which the event took place (14a). The clitic *=yo* presents a similar pattern; in (13b), the clitic refers to the GOAL, while in (14b), it refers to the PATH.

(13) KWINJIRA : to enter

    a. N-di   kw-injir-a=mwo.
        1SG-BE INF-enter-IMP=18.LOC

        'I am entering.' (e.g. the house)

    b. N-di   kw-injir-a=yo.
          1SG-BE INF-enter-IMP=17.LOC

          'I am entering.' (e.g. a country)

(14) KWINJIR-IR-A : to enter through

    a. N-di   kw-injir-ir-a=mwo.
          1SG-BE INF-enter-APPL-IMP=18.LOC

          'I am entering through somewhere.' (e.g. a window)

    b. N-di   kw-injir-ir-a=yo.
          1SG-BE INF-enter-APPL-IMP=17.LOC

          'I am entering through somewhere.' (e.g. Canada en route to America)

The data in (13) and (14) allow us to draw two conclusions regarding the argument structure of the verb *kw-injira* 'to enter'. First, in its non-applied form in (13), the verb selects for a grammatical object which is the GOAL. Second, in the applicativized variant in (14), the applied object is a PATH. Crucially, the semantic role of the locative arguments that are selected by both applicativized and non-applicativized variants of *kw-injira* 'to enter' are distinct from those selected by the verb *kw-iruka* 'to run' in (11) and (12).

4 Verb class interactions

In the previous section I showed that verbs differ in the kinds of locative arguments they take, with four general classes: verbs where the applied object is a general LOCATIVE, a GOAL, a PATH, or a SOURCE. In this section I show that the four interpretations of the locative applicative described in the previous section do not appear arbitrarily, but rather the interpretation of the applied argument is contingent upon the semantic class of the verb.

    The GOAL applied object is reserved for verbs of manner of motion, such as *kw-iruka* 'to run', *gu-tembera* 'to go about', and *gu-simbuka* 'to jump'. Evidence of the underlying GOAL with these verbs comes from the fact that the GOAL can be licensed by the verb without an applicative. Many manner of motion verbs are ambiguous between a static location reading and a change of location reading. For example, *gu-simbuka* 'to jump' in (15) can be coerced into having a GOAL reading independently of the applicative.

(15) Yohani y-a-simbuts-e   mu ma-zi.
      John   1-PST-jump-PERF 18 6-water

      'John jumped while in the water.'

      'John jumped into the water.'

The applicativized variant in (16), however, requires that the locative is a GOAL:

(16) Yohani y-a-simbuk-iy-e   mu ma-zi.
      John   1-PST-jump-APPL-PERF 18 6-water

      'John jumped into the water.'

      *('John jumped while in the water.')

The ability for the locative in (15) to be interpreted as a GOAL is a pattern attested in several unrelated languages where manner of motion verbs which convey or entail displacement are coercible to have a GOAL reading (Nikitina 2008; Tham, Beavers & Levin 2012; Bassa Vanrell 2013).[11]

   When the applicative licenses a PATH , it is with so-called path verbs, such as *kw-injira* 'to enter', *gu-sohoka* 'to exit', *ku-manuka* 'to descend', *kuzamuka* 'to ascend', and *ku-rira* 'to climb'. The SOURCE applied object is restricted to the verb of traversal *kw-ambuka* 'to cross'.

   With verbs that encode no location in their meaning, the applicative licenses a general location, and the applicative morpheme is obligatory for licensing a LOCATIVE with such verbs. For example, consider the verb *ku-vuga* 'to talk'; this verb does not license a location, as shown by the inability of the non-applicativized verb to appear with locative clitics, as in (17b).

(17) a. Habimana a-ri   ku-vug-a.
          Habimana 1S-BE INF-talk-IMP

          'Habmiana is talking.'

      b. Habimana a-ri   ku-vug-a(=*ho/yo/mo).
          Habimana 1-COP INF-talk-IMP=LOC

          'Habimana is talking (there).

The applicativized variant, however, does in fact permit locative clitics, as in (18). The use of different locatives is contingent upon context; in this example, *=m(w)o* would mean that the subject is speaking inside of a location (e.g. his house), while *=yo* means that he is speaking at a general location (e.g. a park).[12]

(18) Habimana a-ri   ku-vug-ir-a(=m(w)o/yo).
      Habimana 1S-BE INF-talk-APPL-IMP=LOC

      'Habimana is talking there.'

---

[11] An anonymous reviewer asks if there are other possibilities besides these patterns, specifically questioning whether the applied object with a verb like *kw-iruka* 'to run' could be a SOURCE and not a GOAL, i.e. to mean something like 'run away from X'. Consultants rejected this reading. What consultants do allow, however, is that an applicativized manner of motion predicate can have a 'toward' interpretation. For example, a sentence like that in (16) could mean 'John jumped toward the water' instead of 'John jumped into the water'. I assume that the GOAL in these cases is prospective; the subject is entailed to move in the direction of the GOAL , but it is not necessarily the case that the subject arrives, which subsumes both the 'to' and 'towards' readings.

[12] The clitic *=ho* is also permissible with the applicativized variant of *ku-vuga* 'to talk', but it does not have a literal locational interpretation. Rather, it means that the subject is using something to talk, such as his cellphone.

In this section, I have shown that the different meanings encoded by the locative clitic are conditioned by the verb to which the applicative attaches. Table 2 summarizes the thematic role of applied objects which are present with the different verbs. Note that this typology reflects all of the logically possible components of a motion event (SOURCE, PATH, AND GOAL), though which component of the motion event is brought out as the applied object crucially depends on the class of verb. When there is no motion in the meaning of the verb, the default interpretation is that the LOCATIVE is a general location.

Table 2: Verb classes and corresponding applied object meaning

Role of the Applied Object	Verb Type	Example
GOAL	manner of motion	*kw-iruka* 'to run'
PATH	change of location	*kw-injira* 'to enter'
SOURCE	traversal	*kw-ambuka* 'to cross'
LOCATIVE	no location encoded by verb	*ku-vuga* 'to talk'

5 Towards an analysis of locative applicatives

The previous two sections have shown that the semantic role of a locative applied object is contingent upon the meaning of the verb. Most previous approaches assume that the applicative adds a new applied object with a corresponding LOCATIVE thematic role and thus do not have an obvious means for capturing the various locative semantic roles found with different classes of motion verbs.

I suggest instead an analysis of applicative morphology as a paradigmatic constraint which requires that the applicativized variant of a given verb has monotonically stricter set of lexical entailments than the non-applicativized variant, as proposed for various argument alternations in Ackerman & Moore (2001) and Beavers (2010).[13] There are two ways that this semantic narrowing is present in sentences with a locative applicative. The first is that the applicative can add a wholesale new syntactic object absent from the meaning of the verb, as is the case with verbs that do not license a LOCATIVE in their non-applied form, such as *ku-vuga* 'to talk' in (18) above. Second, the applicative can realize an argument that is selected for semantically but not realized syntactically by the non-applicativized verb. In the former case, the entailments of the applicativized variant are narrower by virtue of specifying a particular location where the event takes place; in the latter case, the entailments are narrower by virtue of naming a specific location that is semantically entailed to exist (but not syntactically licensed) in the non-applicativized variant.

---

[13] By *monotonic* I mean that a new meaning is added without removing any prior meaning in the base predicate (Koontz-Garboden 2007; 2012).

Thus the applicativized variant should include all of the information in the non-applicativized variant with additional semantic information pertaining to an argument of the predicate. In order to make this distinction precise, I assume a neo-Davidsonian style semantics, where each participant is linked to the event by a specific thematic role (AGENT, THEME, PATH, GOAL, etc.). Due to restrictions of length, I do not give a fully articulated formalization of the mapping between syntax and semantics, but crucially, any entity linked to a thematic role is a syntactically realized argument. For example, the notation *ag(john', e)* states that *john* is the AGENT of the event *e*.

For verbs that do not have a LOCATIVE in the non-applicativized variant, a LOCATIVE is encoded by the applied object, narrowing the truth-conditional content by describing a location at which the event took place. As shown above in (18), the verb *ku-vuga* 'to talk' does not select for a LOCATIVE. From this, I assume the denotation in (19) for *kuvuga* 'to talk'.

(19) $\exists e.[talk'(e) \,\&\, ag(john', e)]$

This denotation states that there is a talking event and that John is the AGENT of that event. The applicativized variant licenses a locative object, which specifies a location at which the event takes place.

(20) a. Yohani a-ri   ku-vug-ir-a   mu nzu.
          John   1S-be inf-talk-appl-imp 18 house
          'John is talking in the house.'
      b. $\exists e.[talk'(e) \,\&\, ag(john', e) \,\&\, loc(house', e)]$

In (20a) — repeated from (4b) — the sentence has the same truth-conditions as in (19), but specifies an additional locative participant. For considerations of space, I do not provide a fully-articulated analysis of how the applicativized variant is derived from the non-applicativized variant.[14] The crucial point here is that the lexical entailments of the applicativized predicate are narrower than those of the non-applicativized predicate by virtue of the additional locative participant.

Verbs which denote a location in their denotation (i.e. verbs of directed motion, as discussed above) do not add a general locative applied object. In these sentences, the applicative is used to bring out a locative participant present in the meaning of the non-applicativized verb, which has the effect of narrowing the truth-conditional content by naming a specific location. For example, the verb *kw-injira* 'to enter' denotes a PATH, though this is not realized syntactically with the base verb, as shown in (21) where the PATH is existentially bound.[15]

---

[14] See Jerro (2016b), Chapter 3, for a more fully articulated analysis of locative applicatives.

[15] Technically, all verbs of directed motion entail the presence of a SOURCE, PATH, and GOAL. Why a particular component of motion is preferred with distinct verbs is left for future research. In the present discussion, I assume that verbs of different classes lexically specify a given component of a motion event that is brought out by applicativization.

(21) a. Yohani a-ri kw-injir-a   mu nzu.
          John   1-ʙᴇ ɪɴꜰ-enter-ɪᴍᴘ 18 house
          'John entered the house.'

      b. $\exists e \exists l.[enter'(e)\ \&\ ag(john', e)\ \&\ th(house', e)\ \&\ path(l, e)]$

With the applicativized variant of *kw-injira* 'to enter', the ᴘᴀᴛʜ participant which is existentially bound in (21) is instead licensed as a syntactic argument, as in (22) where the ᴘᴀᴛʜ is realized overtly as the applied object.

(22) a. Yohani a-ri kw-injir-ir-a   mu muryango mu nzu.
          John   1-ʙᴇ ɪɴꜰ-enter-ᴀᴘᴘʟ-ɪᴍᴘ 18 door   18 house
          'John entered the house through the door.'

      b. $\exists e.[enter'(e)\ \&\ ag(john', e)\ \&\ th(house', e)\ \&\ path(door', e)]$

The analysis presented so far has shown that an applicativized variant requires stricter truth-conditional content on an argument, which is satisfied by either adding an new locative object or syntactically licensing a participant that is only semantically entailed by the meaning of the verb. This analysis subsumes the object-adding function that has been the focus of the mainstay of research on applicatives and additionally provides a framework of analysis for discussing the applied objects found with particular motion verbs. A further possibility in this semantically-oriented analysis is that the applicativized variant need not necessarily license an additional object, but may change the semantic nature of an existing thematic object of the verb, provided there is a stricter semantic meaning in the applicativized variant, akin to the paradigmatic argument alternations discussed in Ackerman & Moore (2001) and Beavers (2010). I turn now to a brief discussion of this use.

First, I propose that the thematic roles of ɢᴏᴀʟ and ʀᴇᴄɪᴘɪᴇɴᴛ are in the appropriate relation of restricted truth conditions; namely, a ʀᴇᴄɪᴘɪᴇɴᴛ has all the entailments of a ɢᴏᴀʟ, but with the additional meaning of prospective change of possession. Consider the definitions of ɢᴏᴀʟ and ʀᴇᴄɪᴘɪᴇɴᴛ in (23).

(23) a. ɢᴏᴀʟ: a place to which motion is directed

      b. ʀᴇᴄɪᴘɪᴇɴᴛ: a place to which motion is directed + prospective change of possession

With these definitions in mind, the theory developed so far predicts that with verbs that license a ɢᴏᴀʟ, it should be possible to satisfy the output condition by 'narrowing' the ɢᴏᴀʟ to ʀᴇᴄɪᴘɪᴇɴᴛ — without modifying the argument structure of the verb. Consider the verb *gu-tera* 'to throw', which is ditransitive in its non-applied form and has a ɢᴏᴀʟ indirect object.[16]

---

[16] I use the term *indirect object* to describe the ɢᴏᴀʟ/ʀᴇᴄɪᴘɪᴇɴᴛ object of a double object construction (see Beavers 2011 for comparable terminology).

(24) Karemera y-a-tey-e   i-buye Nkusi.
      Karemera 1-PST-throw-PERF 5-rock Nkusi

      'Karemera threw the rock at Nkusi.'

The sentence in (24) has the reading that Karemera is throwing a rock directly at Nkusi, possibly trying to harm him and, crucially, without the intention of giving Nkusi possession of the rock. The denotation of this sentence is as follows:

(25) $\exists e.[throw'(e) \ \& \ ag(karemera', e) \ \& \ th(rock', e) \ \& \ goal(nkusi', e)]$

Given the relationship of GOALS and RECIPIENTS assumed above, the constraint that the applicativized variant has stricter truth conditions is satisfied by the change of the GOAL participant in (25) to a RECIPIENT, given (23). For the verb *gu-tera* 'to throw', this is precisely the meaning of the applicativized variant, as shown in (26).

(26) Karemera y-a-ter-ey-e   i-buye Nkusi.
      Karemera 1-PST-throw-APPL-PERF 5-ball Nkusi

      'Karemera threw the rock to Nkusi.'

Crucially, the required reading in (26) is one where Karemera is attempting to give Nkusi possession of the rock and not that Karemera is throwing the rock at Nkusi. In this sentence, Nkusi is not just the GOAL of the throwing event, but also the RECIPIENT of a prospective change of possession. The denotation of the sentence in (26) is that in (27).

(27) $\exists e.[throw'(e) \ \& \ ag(karemera', e) \ \& \ th(rock', e) \ \& \ rec(nkusi', e)]$

Evidence for the meaning difference between the two comes from the fact that prospective catching can only be modified when the applied object has been narrowed to a RECIPIENT, as in (28a), where the conjunction *ariko* 'but' is used to contrastingly deny the reception of the ball. In (28b), on the other hand, where there is no applicative, the object is not a RECIPIENT; thus it is infelicitous to modify any notion of Nkusi attempting to catch the rock.

(28) a. Karemera y-a-ter-ey-e   i-buye Nkusi, ariko Nkusi
         Karemera 1S-PST-throw-APPL-PERF 5-rock Nkusi, but   Nkusi
         nti-y-a-ri-fash-e.
         NEG-1S-PST-5O-catch-PERF

         'Karemera threw the rock to Nkusi, but Nkusi didn't catch it.'

      b. #Karemera y-a-tey-e   i-buye Nkusi, ariko Nkusi
         Karemera 1SG-PST-throw-PERF 5-rock Nkusi but   Nkusi
         nti-y-a-ri-fash-e.
         NEG-1S-PST-5O-catch-PERF

         'Karemera threw the rock at Nkusi, but Nkusi didn't catch it.'

Further evidence is that in the presence of another applied object (such as LOCATIVE), only the GOAL reading is possible with the applicativized variant of *gu-tera* 'to throw'; here, the applicative is used to license a LOCATIVE object, leaving the lexical entailments of the indirect object unchanged. Consider the example in (29). Given that the applicative licenses the LOCATIVE, it is expected that in (29), *Nkusi* is the GOAL and not the RECIPIENT since the applicative is not being used to narrow the entailments associated with the GOAL.

(29)   Karemera y-a-ter-ey-e   i-buye Nkusi mu nzu.
       Karemera 1S-PST-throw-APPL-PERF 5-rock Nkusi in   9.house
       'Karemera threw the rock at Nkusi in the house.'

Thus in (29) the reading is that the rock is thrown *at* Nkusi, and not that it is thrown *to* him.

6 Conclusion

In this paper, I have shown that the semantic role of an applied object may be contingent upon the meaning of the verb, and I have described a case in Kinyarwanda where the use of the applicative does not license a new argument at all, but rather modifies the semantic role of an existing argument. In order to capture these facts, I proposed an analysis of applicativization as sensitive to a paradigmatic output condition.

Specifically, in §3 I showed that there exist three classes of motion predicates where the applied object is assigned the role of SOURCE, PATH, or GOAL, respectively, and in §4 I showed that these applied objects appear with verbs of traversal, path verbs, and manner of motion verbs, respectively. In §5 I provided a preliminary account of applicativization as a paradigmatic output condition on the applicativized variant of a given verb where the predicate of the applicativized verb must have stricter lexical entailments associated with a particular argument than the non-applicativized verb. This analysis captures the typology of predicates presented in §§3-4 and makes the further prediction that certain verbs may not add a new argument at all under applicativization. I have not attempted an exhaustive account of applicativization in Kinyarwanda, but rather I have shown that applicativization cannot be analyzed simply as an operation which adds a whole new argument with an associated semantic role. Instead, I have proposed a framework for discussing applicatives which provides a more empirically predictive analysis of the uses of applicative morphology in Bantu.

Appendix A: Prepositions or class prefixes?

Bantu languages are well known for their gender class agreement, a class of prefixes that indicate singular and plural as well as other semantic distinctions such as animacy. Relevant to the discussion above is that there are classes reserved in many Bantu languages

for locative expressions. For example, Chicheŵa has *pa–*, *ku–*, and *mu–* (traceable to Proto-Bantu), which are referred to as classes 16, 17, and 18 in the literature (Bresnan & Kanerva 1989; Bresnan 1994; Bresnan & Mchombo 1995; Maho 1999). While Kinyarwanda orthographic conventions (which are adopted above) separate *ku* and *mu* from the following noun, an empirical question arises as to the status of these locatives. I show in this appendix that the locatives *ku* and *mu* in Kinyarwanda are in fact locative class markers and that nouns marked with locative class prefixes are arguments and not prepositional phrases (see Jerro 2013 and Jerro & Wechsler 2015 for general discussion of agreement in Kinyarwanda).

First, locatives can appear as the subject of a passive, a position reserved for arguments. In (30), the locative phrase *mw' ishymaba* is the subject of a passivized verb, triggering subject agreement (cf. Bresnan & Kanerva 1989; Bresnan 1994). It is important to note that the subject agreement marker is from class 16 (the class for inherent locatives, such as *aha-ntu* 'a place' and *ah-irengeye* 'a high place'), and in Kinyarwanda and various other languages such as Kesukuma (Batibo 1985: 245), any noun marked from class 16, 17, 18, or 23 (i.e. locative classes) triggers a subject marker from class 16 (Maho 1999).

(30) Mw' i-shyamba h-a-tem-e-w-e   igi-ti n' umu-higi.
       18   5-forest   16S-PST-cut-APPL-PASS-PERF 7-tree by 1-hunter

      'In the forest was cut the tree by the hunter.'

Of crucial importance is that the subject triggers agreement on the verb, an agreement relation reserved for arguments.

Furthermore, locative phrases can be object-marked on verbs, as shown in (31b), where the class 16 object marker *ha-* replaces the locative phrase.

(31) a. N-a-bon-ey-e   umw-ana mw' i-shyamba.
          1SGS-PST-see-APPL-PERF 1-child 18   5-forest

         'I saw the child in the forest.'

      b. N-a-ha-bon-ey-e   umw-ana.
          1SGS-PST-16O-see-APPL-PERF 1-child

         'I saw the child there.'

The final piece of evidence that locative phrases are arguments is that they cannot appear productively across predicates, which would be expected if the locative prefixes were in fact prepositions that license oblique phrases. For example, in (32) the locative phrase *mu nzu* 'in the house' cannot be used with the verb *ku-vuga* 'to talk'.[17]

(32) Habimana a-ri   ku-vug-a   (*mu n-zu).
       Habimana 1S-BE INF-talk-IMP 18   9-house

      'Habimana is talking (*in the house).'

---

[17] Some verbs do select a locative argument, but I this is selection is on a case-by-case basis and not productive across verbs.

In order to have a locative phrase such as *mu nzu* 'in the house' with the verb *ku-vuga* 'to talk', the applicative is obligatory, as in (33).

(33) Habimana a-ri   ku-vug-ir-a   mu n-zu.
      Habimana 1S-BE INF-talk-APPL-IMP 18   9-house

      'Habimana is talking in the house.'

From these diagnostics, I conclude that Kinyarwanda locative phrases are class-marked arguments and not oblique phrases.

Acknowledgements

Abbreviations

1 *to* 23	gender class prefixes	LOC	locative
1SG	first-person singular	O	object marker
ASP	aspect	PERF	perfective
APPL	applicative	PST	past tense
BEN	benefactive applicative	PL	plural
FV	final vowel	PRES	present tense
IMP	imperfective	S	subject prefix
INF	infinitive	SG	singular

References

Ackerman, Farrell & John Moore. 2001. *Proto-properties and grammatical encoding: A correspondence theory of argument selection.* Stanford: CSLI Publications.

Alsina, Alex. 1992. On the argument structure of causatives. *Linguistic Inquiry* 23 (4). 517–555.

Alsina, Alex & Sam Mchombo. 1993. Object asymmetries and the Chicheŵa applicative construction. In Sam Mchombo (ed.), *Theoretical aspects of Bantu grammar*, 17–45. Stanford: CSLI Publications.

Asher, Nicholas & Pierre Sablayrolles. 1995. A typology and discourse semantics for motion verbs and spatial PPs in French. *Journal of Semantics* 12. 163–209.

Baker, Mark. 1988. Theta theory and the syntax of applicatives in Chicheŵa. *Natural Language and Linguistic Theory* 6. 353–389.

Baker, Mark & Chris Collins. 2006. Linkers and the internal structure of *v*P. *Natural Language and Linguistic Theory* 24. 307–354.

Baker, Mark, Ken Safir & Justine Sikuku. 2012. Sources of (a)symmetry in Bantu double object constructions. In Nathan Arnett & Ryan Bennett (eds.), *Proceedings of the 30th West Coast Conference on Formal Linguistics*, 54–64. Somerville, MA: Cascadilla Proceedings Project.

Bassa Vanrell, Maria del Mar. 2013. *Preposition typology with manner of motion verbs in Spanish.* Austin, TX: University of Texas at Austin. (Master's thesis.)

Batibo, Herman. 1985. *Le kesukuma (langue bantu de Tanzania): phonologie, morphologie.* Paris: Centre de Recherches, d'Échanges et de Documentation Universitaire, Kenya.

Beavers, John. 2010. The structure of lexical meaning: Why semantics really matters. *Language* 86. 821–864.

Beavers, John. 2011. An aspectual analysis of ditransitive verbs of caused possession in English. *Journal of Semantics* 28. 1–54.

Beavers, John, Beth Levin & Shiao Wei Tham. 2010. The typology of motion expressions revisited. *Journal of Linguistics* 46. 331–377.

Bond, Oliver. 2009. The locative applicative in Eleme. *Transactions of the Philological Society* 107 (1). 1–30.

Bresnan, Joan. 1994. Locative inversion and the architecture of universal grammar. *Language* 70(1). 72–131.

Bresnan, Joan & Jonni Kanerva. 1989. Locative inversion in Chicheŵa: A case study of factorization in grammar. *Linguistic Inquiry* 20 (1). 1–50.

Bresnan, Joan & Sam Mchombo. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13. 181–254.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21 (2). 147–185.

Cann, Ronnie & Patricia Mabugu. 2007. Constructional polysemy: The applicative construction in chiShona. *Metalinguistica* 19. 221–245.

Creissels, Denis. 2004. Non-canonical applicatives and focalization in Tswana.

Diercks, Michael. 2011. Incorporating location in argument structure: The Lubukusu locative clitic. In E. Bokamba, R Shosted & BT Ayalew (eds.), *Selected Proceedings of the 40th Annual Conference on African Linguistics: African Languages and Linguistics Today*, 65–79. Somerville, MA: Cascadilla Proceedings Project.

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 1997. A typology of argument-determined constructions. In Joan Bybee, John Haiman & Sandra Thompson (eds.), *Essays on language function and language type: Dedicated to T. Givón*, 71–113. John Benjamins: Amsterdam.

Fillmore, Charles J. 1970. The grammar of *hitting* and *breaking*. In Roderick Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 120–133. Waltham: Ginn.

Jeong, Youngmi. 2007. *Applicatives: Structure and interpretation from a Minimalist perspective*. Amsterdam: John Benjamins.

Jerro, Kyle. 2013. Argument structure and the typology of causatives in Kinyarwanda: Explaining the Causative-Instrumental syncretism. Austin, TX: University of Texas at Austin. (MA report.)

Jerro, Kyle. 2015. Revisiting object symmetry in Bantu. In Ruth Kramer, Elizabeth Zsiga & One Tlale Boyer (eds.), *The selected Proceedings of the 44th Annual Conference on African Linguistics*, 130–145. Somerville, MA: Cascadilla Proceedings Project.

Jerro, Kyle. 2016b. *The syntax and semantics of applicative morphology in Bantu*. Austin, TX: The University of Texas at Austin. (Doctoral dissertation.)

Jerro, Kyle & Stephen Wechsler. 2015. Quantifiers and person morphology: The case study of Bantu. In Jürg Fleischer, Elisabeth Rieken & Paul Widmer (eds.), *Agreement from a diachronic perspective*, 147–164. Berlin: Mouton de Gruyter.

Kimenyi, Alexandre. 1980. *A relational grammar of Kinyarwanda*. Berkeley: University of California Press.

Kisseberth, Charles & Mohammad Imam Abasheikh. 1977. The object relationship in Chi-Mwi: Ni, a Bantu language. In Peter Cole & Jerrold Sadock (eds.), *Syntax and semantics: Grammatical relations*, 179–218. New York: Academic Press.

Koontz-Garboden, Andrew. 2007. *States, changes of state, and the monotonicity hypothesis*. Palo Alto, CA: Stanford University. (Doctoral dissertation.)

Koontz-Garboden, Andrew. 2012. The monotonicity hypothesis. In Louise McNally & Violeta Demonte (eds.), *Telicity change and state: A cross-categorial view of event structure*, 139–161. Oxford: Oxford University Press.

Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Maho, Jouni. 1999. *A comparative study of Bantu Noun classes*. Göteborg: Acta Universitatis Gothoburgensis.

Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In Sam Mchombo (ed.), *Theoretical aspects of Bantu grammar*, 113–148. Stanford: CSLI Publications.

Marten, Lutz. 2003. The dynamics of Bantu applied verbs: An analysis at the Syntax-Pragmatics interface. In Kézié K. Lébikaza (ed.), *Acts du 3e congrès mondial de linguistique africaine lomé 2000*, 207–221. Köln: Köppe.

McGinnis, Martha. 2001. Variation in the phase structure of applicatives. *Linguistic Variation Yearbook* 1. 105–146.

McGinnis, Martha & Donna B. Gerdts. 2003. A Phase-Theoretic analysis of Kinyarwanda multiple applicatives. In *Proceedings of the 2003 Canadian Linguistic Association Annual Conference Department of Linguistics*, 154–165. Université du Québec à Montréal.

Nikitina, Tatiana. 2008. Pragmatic factors and variation in the expression of spatial goals: The case of *in* vs. *into*. In *Syntax and semantics of spatial p*. Amsterdam: John Benjamins.

Peterson, David A. 2007. *Applicative constructions*. Oxford: Oxford University Press.

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge: The MIT Press.

Rappaport Hovav, Malka & Beth Levin. 2008. The English dative alternation: A case for verb sensitivity. *Journal of Linguistics* 44. 129–167.

Rugemalira, Josephat M. 1993. *Runyambo verb extension and constraints on predicate structure*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation.)

Rugemalira, Josephat M. 2004. Locative arguments in Bantu. In *Proceedings of the 4th World Congress of African Linguistics*, 285–296. Cologne: Rüdiger Köppe Verlag.

Schaefer, Ronald. 1985. Motion in Tswana and its characteristic lexicalization. *Studies in African Linguistics* 16 (1). 57–87.

Sibanda, Galen. 2016. The Ndebele applicative construction. In Doris L. Payne, Sara Pacchiarotti & Mokaya Bosire (eds.), *The Ndebele applicative construction* (Contemporary African Linguisics 1), 309–333. Berlin: Language Science Press.

Simango, Silvester R. 2012. The semantics of locative clitics and locative applicatives in ciCewa. In Bruce Connell & Nicholas Rolle (eds.), *Selected Proceedings of the 41st Annual Conference on African Linguistics: African Languages and Linguistics Today*, 141–149. Somerville, MA: Cascadilla Proceedings Project.

Sitoe, Bento. 1996. The semantics and syntax of the Tsonga verbs *kuwà* 'fall' and *kuntlúlá* 'jump', and their relatives. *South African Journal of African Languages* 16. 144–148.

Slobin, Dan I. 1996. Two ways to travel: Verbs of motion in English and Spanish. In Masayoshi Shibatani & Sandra Thompson (eds.), *Grammatical constructions: Their form and meaning*, 195–219. Oxford: Oxford University Press.

Talmy, Leonard. 1975. Semantics and syntax of motion. In John P. Kimball (ed.), *Syntax and semantics*, vol. 4, 181–238. New York: Academic Press.

Tham, Shiao Wei, John Beavers & Beth Levin. 2012. Directional interpretations with locative prepositions. (Paper presented at "The Meaning of P" in Bochum, Germany.)

Zeller, Jochen. 2006. Agreement and the EPP in Kinyarwanda applicatives. In Patrick Brandt & Eric Fuß (eds.), *Form, structure, and grammar*, 275–295. Berlin: Akademie-Verlag.

Zeller, Jochen & Jean Paul Ngoboka. 2006. Kinyarwanda locative applicatives and the Minimal Link Condition. *Southern African Linguistics and Applied Language Studies* 24 (1). 101–124.

Zlatev, Jordan & Peerapat Yangklang. 2004. A third way to travel: The place of Thai in motion-event typology. In Sven Strömqvist & Ludo Verhoeven (eds.), *Relating events in narrative, vol 2: Typological and contextual perspectives*, 159–190. Mahwah, NJ: Lawrence Erlbaum.

The Ndebele applicative construction

Galen Sibanda
Michigan State University

Applicatives have been shown to be generally valence adding, and the Ndebele applicative construction is no exception. While change in argument structure is discussed in the article, the main focus is on the variation in thematic roles of the participants that the applicative introduces. The thematic roles associated with a given Ndebele verb are predictable from the semantic class of that verb, including those of participants introduced by applicative *-el*. A number of different semantic classes are considered. The discussion raises theoretical questions about verb meaning and that of the Ndebele applicative suffix *-el*. The precise meaning of *-el* can be best captured by appealing to the notion of underspecification. The thematic roles *-el* introduces in all semantic verb classes are REASON and LOCATION, and not always BENEFICIARY which has often been highlighted in other studies of applicatives.

1 Introduction

There have been many studies of applicative constructions, covering a wide range of languages around the world (e.g. Peterson 2007) including Bantu languages (e.g. Ngonyani 1996; Mabugu 2011; Jerro 2016 & this volume). Of particular interest in most of these studies has been how the argument structure of the verb is altered in applicative constructions. Applicatives have been shown to be generally valence adding, and the Ndebele applicative construction considered here is no exception.

While change in Ndebele argument structure is discussed in this article, specifically in §3, the main focus is on the variation in the thematic roles of participants that the applicative introduces. A number of verbs in different semantic classes are considered in §4 although neither the list of verbs nor semantic classes is exhaustive. As such, this work does not provide statistics, but the goal is to establish some generalizations about the thematic roles of participants introduced by the applicative suffix and to show how these relate to semantic classes of verbs. The semantic classification of verbs follows from the works of, for example, Chafe (1970); Dowty (1987; 1991); Foley & Van Valin (1984), and Payne (1997). Thematic roles appealed to are based on the works of Frawley (1992); Fillmore (1968; 1977) and Halliday (1970), among others. The discussion raises important questions about verb meaning and the precise meaning of the applicative suffix *-el* (normally followed by default final vowel *-a* or another suffix). §5 is the conclusion.

2 Background

The discussion in this chapter revolves around three key issues: thematic roles, semantic classes and applicative constructions. It is therefore necessary to provide some background to these before delving into the main investigation of the Ndebele applicative construction.

2.1 Thematic roles

Thematic role lists date back to the work of the Sanskrit grammarian, Panini, in the 6th century BCE (Dowty 1991: 548; Srikumar 2013: 19–20), and structuralists such as Blake (1930). They were brought to prominence in linguistic theory by Gruber (1965); Fillmore (1968; 1977), and Jackendoff (1972; 1976). Commonly known as *semantic roles*, thematic roles have also been referred to by other names. For instance, Panini called them *karakas* (Dowty 1991: 548; Srikumar 2013; Cardona 1974: 19-20); Fillmore (1968) called them semantic *cases* or *deep cases*; Gruber (1965) and Jackendoff (1972) used the term *thematic relations*; for Stowell (1981) they were *theta-grids*; and for yet others *theta roles* (Chomsky 1981; Marantz 1984).

While the significance of thematic roles in linguistic theory cannot be refuted, particularly in studies concerned with the syntax-semantics interface, identifying or labeling thematic roles has had challenges, traces of which may still be evident in this article. Some of the important issues are discussed in the works of, for example, Dowty (1991), Jackendoff (1987), Newmeyer (2010: 689) and Kittilä, Katja & Ylikoski (2011: 6). The number of proposed roles varies from a few to an almost limitless list. For instance, in his localist approach, Anderson (1971) proposed only three roles (SOURCE, LOCATION and GOAL), and all non-local values then derive from these. In his original work, Fillmore (1968) claimed that thematic roles (what he called *cases*) formed a finite set including AGENTIVE, INSTRUMENTAL, DATIVE, FACTITIVE, LOCATIVE and OBJECTIVE, although he made it clear this was not a complete list. With more research the number later grew to over two thousand roles, also referred to as *Frame Elements* (Fillmore 1985). Blake (1930) listed 87 temporal or locative roles and 26 other roles. To aggravate the situation, as Dowty (1991: 548-549) observes, "new candidates for thematic roles are being proposed all the time, e.g. FIGURE and GROUND in Talmy 1985a, NEUTRAL in Rozwadowska 1988, LANDMARK in Jackendoff 1982, even SUBJECT in Baker 1985." Perhaps the most extreme case which, however, seems to have been ignored by many, is the HPSG one where each verb assigns its own peculiar thematic roles, different from the roles of any other verb (Pollard & Sag 1994). In HPSG the verb *love* would, for example, assign two thematic roles: LOVER and LOVEE. As Dowty (1989) notes, in this approach, there would be no thematic role types but only individual thematic roles, and important semantic generalizations are lost. Payne (1997: 52) hammers the same point in stating that "an infinitely long list of sematic roles is as useless as no list at all".

A related problem is the issue of delimiting boundaries between roles. Depending on theoretical approach, some roles are further subdivided into more specific roles; differ-

ent names are sometimes used for the same role concept; and definitions of some roles overlap. For instance, Dowty (1991: 553) illustrates "role fragmentation" with AGENT, which he says has been divided into AGENT and ACTOR (Jackendoff 1983), AGENT and EF-FECTOR (Van Valin 1990), VOLITIVE, EFFECTIVE, INITIATIVE, and AGENTIVE (Cruse 1973), while Lakoff (1977) proposes up to fourteen different characteristics. As another instance, the role called DIRECTION ('towards/away from') is sometimes used as a cover term for SOURCE and GOAL. Anderson (1977) uses THEME for what is now widely taken to be a PA-TIENT, and others use PATIENT and THEME interchangeably. The PATIENT versus THEME issue partly reflects a definitional problem (Dowty 1991: 548-549; Löbner 2002: 113). Palmer, Gildea & Xue (2010: 5) explain that while it is difficult to draw clear boundaries between PATIENTS and THEMES, the commonly held view is that a PATIENT undergoes a change of state whereas a THEME simply changes location.

Another problem is that some participants have been claimed to take more than one role. Following Culicover & Wilkins (1986) and Talmy (1985b), Jackendoff (1987: 395) proposes two tiers where AGENT and PATIENT occupy the "action tier" while other roles dealing with motion and location (e.g. SOURCE, THEME, GOAL) occupy the "thematic tier", such that a single participant could have both PATIENT and GOAL roles but on different tiers, or both PATIENT and THEME, etc.

Observing all these problems, Dowty (1991) moves away from positing many roles and proposes defining Prototypical AGENTS and PATIENTS such that each one covers an array of different finer types. He argues that thematic roles should not be viewed as semantic primitives or discrete categories, but must be defined in terms of entailments so that they are seen as prototypes where there may be different degrees of membership. Thus some roles will be more agent-like or patient-like depending on the number of AGENT or PATIENT Proto-role properties they fulfill. This would seem to solve the problems of "role fragmentation" and boundaries. However, his proposal has also been criticized (Levin & Rappaport Hovav 1996).

Another criticism regarding thematic role lists, especially when they first gained pop-ularity, was that they were often unstructured. As a result they generally could not capture important commonalities and differences across roles. To address this issue, a number of hierarchies, many of which make reference to animacy, frequency in the world's languages, and subjects and objects, have since been proposed (Fillmore 1968; Saint-Dizier & Viegas 1995: 12; Kiyosawa & Gerdts 2010: 334), Bresnan & Kanerva (1989), though they often differ in their details. Of relevance to this study, Bentley (1994: vii) mentions that thematic role hierarchies have sometimes been used to explain differences in the behavior of objects, as well as the relative prominence of arguments in events. Mchombo (2004: 129), for example, has argued for the thematic hierarchy in (3) which attempts to explain the morphosyntactic behavior of different applied objects. (See also Ngonyani & Githinji 2006.)

(1) AGENT > BENEFICIARY > GOAL/EXPERIENCER > INSTRUMENT > PATIENT/THEME > LOCATION > MALEFACTIVE > CIRCUMSTANTIAL

The ordering of roles in this hierarchy can explain why a BENEFICIARY and a CIRCUM-STANTIAL, for example, take different object positions in a Bantu sentence. Generally, roles on the left of the hierarchy, often associated with animate objects, surface closer to the verb while those to the right are more peripheral and are usually associated with inanimate objects. It will be shown in the the remainder of this paper that Ndebele does not depart much from this.

Despite the criticisms noted above, thematic roles remain necessary in capturing important semantic and syntax-semantics generalization, including the behavior of applicative constructions. However, clear definitions of semantic roles are necessary before getting into detailed discussion. The definitions of roles introduced by the Ndebele applicative *-el* and others used in this study (some of which are combinations of definitions) and their sources are provided in Table 1.[1]

2.2 Semantic verb classes

Verb classes structure the lexicon. One of the most influential studies is Levin's (1993) study of English verbs, which is based on syntactic alternations. Her investigation shows correlations between some aspects of the semantics and the syntactic behavior of English verbs. Similar studies have been done for Spanish (Vázquez et al. 2000) and German (Schumacher 1986; Schulte im Walde & Brew 2002). Other approaches to identifying semantic verb classes include elements of Lexical Conceptual Structure (Gruber 1965; Jackendoff 1983; 1990) and sematic roles (Chafe 1970; Cook 1979; Longacre 1976; Foley & Van Valin 1984; Van Valin 1993). No attempt will be made in this paper to identify Ndebele verb semantic classes using such methods, but this study utilizes some classes proposed elsewhere in the literature.

2.3 Applicative construction

Studies of Bantu applicative constructions have arrived at some interesting conclusions on a number of issues such as object symmetry, the BENEFICIARY thematic role, animacy and thematic hierarchies.

Most studies of applicative constructions in Bantu languages have concentrated on syntactic properties, especially regarding the behavior of the applied object (AO) versus that of the base or "logical object" (LO). Of particular interests has been the comparison between symmetrical and asymmetrical type languages often involving the subjection of the objects to various syntactic tests such as passivization, object agreement and word order have been applied to distinguish them (Bresnan & Moshi 1990; Ngonyani &

---

[1] One semantic notion associated with the Ndebele applicative construction, but largely left out of the discussion, is 'in the presence of' or 'witnessed by.' Sometimes this notion has been mistaken for a BENEFICIARY or MALAFICIARY, but the (un)fortunatness of a situation is sometimes simply implied by the verb root itself. An example is *Umntwana uyangigulela*, literally 'The child is sick "for me"' where 'me' is clearly not BENEFICIARY as no one can be sick on behalf of another person. The sentence should be interpreted as 'The child is sick in my presence'. That this is to my detriment can be inferred from the verb *gula* 'be sick' itself, not from the applicative suffix *-el*.

Table 1: Definitions of thematic roles

Thematic Role	Definition	Reference
AGENT (A)	animate and volitional perceived instigator/initiator of the action or event	Fillmore (1968); Payne (1997: 49)
BENEFICIARY (B)	participant for whose benefit the action expressed by the verb is performed	Saint-Dizier & Viegas (1995: 11); Palmer, Gildea & Xue (2010: 4)
EMITTER (EM)	entity that produces or emits a sound, smoke, fumes, gas, dust, etc.	
EXPERIENCER (E)	(animate) participant that is informed of something or that experiences perception, feeling or some psychological state expressed by the predicate (e.g. first argument of *love*, second argument of *annoy*)	Saint-Dizier & Viegas (1995: 11); Löbner (2002: 113)
GOAL (G)	entity towards which a movement is directed (e.g. second argument of *reach*, *arrive*), or the motivation of an action	Saint-Dizier & Viegas (1995: 11)
LOCATION (L)	place in which the action or state described by the predicate takes place (e.g. second argument of *fall*)	Saint-Dizier & Viegas (1995: 11)
MALEFICIARY (M)	participant to whose detriment the action expressed by the verb is performed	(Kittilä & Zúñiga 2010: 5)
PATIENT (P)	participant undergoing the action and that is affected by it – usually undergoes a physical, visible change in state (e.g., second argument of *kill*, *eat*)	Saint-Dizier & Viegas (1995: 11); Palmer, Gildea & Xue (2010: 4); Payne (1997: 51)
REASON (R)	motivational source of a predication or event	Frawley (1992: 225)
STIMULUS (STI)	causer of an emotional reaction	Palmer, Gildea & Xue (2010: 13); Dowty (1991)
THEME (T)	entity that is moving or changing location, condition, or state or being in a given state or position (e.g. the second argument of *give*, the argument of *walk*, *die*)	Saint-Dizier & Viegas (1995: 11); Palmer, Gildea & Xue (2010: 4)

Githinji 2006; Pylkkänen 2000; Machobane 1989; Alsina & Mchombo 1993; Baker 1988; Harford 1993; Ngonyani 1996). While both object types generally display "true" or "primary object" syntactic properties in a symmetrical type language, only one of the two objects behaves like a "true" or "primary" object in an asymmetrical language. Only a few studies have paid more attention to the semantics of the applicative suffix.

One of the main semantic questions has concerned the issue of BENEFACTIVE versus other thematic roles. In particular, studies of applicative constructions have claimed that the BENEFICIARY role is the one most commonly associated with the applied object cross-linguistically (e.g., Peterson 2007; Polinsky 2008; Kittilä & Zúñiga 2010). Peterson (2007: 40) concludes that "if a language has a construction which could be characterized as an applicative it is most common that the semantic role of the applicative object will be that of a RECIPIENT and/or BENEFICIARY/MALEFICIARY." While acknowledging the presence of other roles of the AO such as RECIPIENT, MALEFICIARY, REASON and INSTRUMENT, Schadeberg (2003), De Kind & Bostoen (2012: 101) and Marten & Kula (2014: 1) also view the BENEFICIARY as the most widespread and productive role associated with the applied object in Bantu.

Also regarding semantics, a number of studies have been concerned with tracing the original or underlying meaning associated with the applicative suffix. For instance, Cann & Mabugu (2007: 3) argue that in Shona "all the primary meanings associated with [the] applicative suffix can be derived from an underspecified generalized GOAL relation." De Kind & Bostoen (2012) also argue for an underlying GOAL function of the applicative in ciLubà. According to Trithart (1983), earlier scholars such as Endemann (1876); Van Eeden (1956) and Kähler-Meyer (1966) proposed an original LOCATIVE use of the applicative, a view also taken by Schadeberg (2003: 74). However, Trithart (1983: 75) suggests an original BENEFACTIVE function of the applicative in Bantu and in Niger-Congo languages in general.

While all these observations might be true to some extent, the different claims may be due to the fact that these studies have not systematically analyzed the behavior of verbs from different semantic classes in the applicative construction. Although semantic classes are sometimes mentioned, in most studies there has been no clear demonstration that the conclusions have been arrived at after examining verbs from different semantic classes, rather than just choosing commonly used or random verbs. As will be illustrated with Ndebele examples below, the applicative construction behaves differently with verbs from different semantic classes. It will be shown that with many classes, the BENEFICIARY does not feature at all. Thus, blanket statements like those cited above about BENEFICIARY in applicative constructions are inaccurate, and they may be valid only with regards to particular semantic verb classes.

3 General properties of the Ndebele applicative construction

As already alluded to, the Ndebele Applicative is marked by the derivational suffix *-el* plus the verbal default final vowel *-a* or another derivational or inflectional suffix. It is valence adding, as illustrated in (2) with a divalent verb root *phek* 'cook' which normally takes only one object.

(2) Divalent verb root

    a. U-mama   u-ø-phek-a   i-lambazi.   A P
        1a-mother 1a-TNS-cook-a 5-porridge

        'Mother is cooking porridge.'

    b. U-mama   u-ø-phek-**el**-a   um-ntwana i-lambazi.   A B P
        1a-mother 1a-TNS-cook-APP-a 1-child   5-porridge

        'Mother is cooking the child porridge.'

    c. U-mama   u-ø-phek-**el**-a   i-lambazi pha-ndle.   A P L
        1a-mother 1a-TNS-cook-APP-a 5-porridge 16-outside

        'Mother is cooking the porridge outside.'

    d. U-mama   u-ø-phek-**el**-a   i-lambazi in-dlala.   A P R
        1a-mother 1a-TNS-cook-APP-a 5-porridge 9-hunger

        'Mother is cooking the porridge due to hunger.'

    e. *U-mama u-ø-phek-**el**-a   i-lambazi.   A P
        1a-mother 1a-TNS-cook-APP-a 5-porridge

        'Mother is cooking the porridge for/at ...'

    f. U-mama   u-ø-phek-a   i-lambazi (pha-ndle).   A P (L)
        1a-mother 1a-TNS-cook-a 5-porridge (16-outside)

        'Mother is cooking the porridge (outside).'

As seen in (2a), a root such as *phek* 'cook' only requires an AGENT (A) and a PATIENT (P) as arguments. However, suffixing *-el*, as in (2b), introduces a new argument with a BENEFICIARY (B) participant role and the verb ends up with two objects. Note that the BENEFICIARY intepretation can be replaced by a MALEFICIARY one if a different root such as *w* 'fall' or *chem* 'urinate' is used, or if the pragmatic situation yields a negative interpretation. The applicative versions of these two roots would most likely be interpreted as *wel* 'fall on' or *chemel* 'urinate on', not 'fall for (on behalf of)' or 'urinate for (on behalf of)'. Thus wherever reference is made to a BENEFICIARY role, there is need to keep in mind that a MALEFICIARY interpretation would most likely apply in the same situation if a different verb encoding misfortune is used or if the same verb is used in a negative pragmatic context; hence it is not necessary to discuss the MALEFICIARY interpretation further in this article.

The same applicative can introduce a LOCATIVE (L) sematic role, as in (2c). While an unapplicativized form of the verb can optionally take the LOCATIVE, as in (2f) so that the sentence carries roughly the same meaning, the LOCATIVE in (2c) behaves as a core argument since its omission makes the sentence incomplete or ungrammatical. It should be noted that while the BENEFICIARY must immediately follow the verb, the LOCATIVE comes after the PATIENT, perhaps reflecting a thematic hierarchy: A > B > P > L.

The core argument introduced by the applicative can also be REASON (R), as in (6d). Like an applied LOCATIVE, an applied REASON behaves as a core argument since it cannot be left out without making the sentence ungrammatical.

Example (2e) illustrates that suffixing the applicative *-el* without introducing a third argument (whose role may be BENEFICIARY, LOCATIVE or REASON) is ungrammatical. The only exception would be a case of "definite null instantiation", i.e. an identifiable referent expressed by zero (Fillmore 1986; Fillmore & Kay 1999), where the second of the two overt arguments of the applicativized verb would be a BENEFICIARY, LOCATIVE or REASON, not a PATIENT.

The facts are similar with regards to intransitive monovalent verb roots such as *khal* 'cry', a sound emission root, illustrated in (3).

(3) Monovalent verb root with EMITTER (EM) role
     (Similar roots: *dum* 'make a sound, thunder', *lil* 'cry, moan', *bhong* 'roar' *bhons* 'low (of cattle)' *bubul* 'groan', *bovul* 'bellow', *klabalal* 'shout loudly')

   a. U-sane lu-ya-khal-a.   EM
       11-baby 11-TNS-cry-a

       'The baby is crying.'

   b. U-sane lu-ø-khal-**el**-a   u-chago.   EM R
       11-baby 11-TNS-cry-APP-a 11-milk

       'The baby is crying for milk.'

   c. U-sane lu-ø-khal-**el**-a   pha-ndle.   EM L
       11-baby 11-TNS-cry-APP-a 16-outside

       'The baby is crying outside.'

   d. *U-sane lu-ø-khal-**el**-a.   EM
       11-baby 11-TNS-cry-APP-a

       'The baby is crying for/at …'

Example (3a) shows that *khal* 'cry' has only one argument, realized as the subject and with an EMITTER (EM) semantic role. When the verb is applicativized, a new argument is introduced and its semantic role can be REASON as in (3b) or LOCATION as in (3c). Without a second argument the applicativized verb is ungrammatical, as shown in (3d).

An additional argument is required even when a three-place verb root, such as *ph* 'give' is applicativized, as exemplified in (4).

(4) Trivalent verb root

    a. U-Sihle u-ø-pha   u-sane uku-dla.   A G T
       1a-Sihle 1a-TNS-give 11-baby 15-food

       'Sihle is giving the baby food.'

    b. U-Sihle u-ø-ph-**el**-a   u-mama   u-sane uku-dla.   A B G T
       1a-Sihle 1a-TNS-give-APP-a 1a-mother 11-baby 15-food

       'Sihle is giving the baby food for mother.'

    c. U-Sihle u-ø-ph-**el**-a   u-sane uku-dla pha-ndle.   A G T L
       1a-Sihle 1a-TNS-give-APP-a 11-baby 15-food 16-outside

       'Sihle is giving the baby food outside.'

    d. U-Sihle u-ø-ph-**el**-a   u-sane uku-dla in-dlala.   A G T R
       1a-Sihle 1a-TNS-give-APP-a 11-baby 15-food 9-hunger

       'Sihle is giving the baby food for hunger.'

    e. *U-Sihle u-ø-ph-**el**-a   u-sane uku-dla.   A G T
       1a-Sihle 1a-TNS-give-APP-a 11-baby 15-food

       'Sihle is giving the baby food for/at …'

The three arguments of the verb root *ph* 'give' are normally associated with the roles AGENT, GOAL (G) and THEME (T), as in (4a). When the applicative *-el* is suffixed, the additional argument may be BENEFICIARY (8b), LOCATION (4c), or REASON (4d). Leaving out the fourth argument is unacceptable (4e). Note that although nouns are used for REASON in (2–4), these can be replaced by a phrase or clause beginning with *ukuthi… /ukuba… /ukuze …* 'because …; so that ….' In fact, out of context REASON is usually expressed more clearly with such a phrase. Ignoring the EMITTER which might, perhaps, be ranked high like the AGENT, the thematic hierarchy drawn from the examples above and based on the relative prominence of arguments in events (the more prominent occurring in the subject position, or if they are objects, closer to the applicativized verb stem) can now be hypothesized as A > B > G > T/P > L/R.

4 The applicative with verbs in different semantic classes

It is clear from the example sets in preceding sections that suffixing *-el* always entails introducing a new argument to the clause. However, the semantic role of the new argument varies from verb to verb. This section looks at what is predictable in this semantic variation.

4.1 Verbs of motion

Verbs of motion are considered first. Frawley (1992: 171) notes that motion involves either positional change or the displacement of some entity and that a complete semantic characterization of motion events "requires the specification of eight semantic properties

in addition to displacement itself." These properties are captured in the meaning of the roles THEME, SOURCE, GOAL, PATH including direction, SITE and medium, INSTRUMENT or conveyance, MANNER, and AGENT. In (5) we see what happens when applicative *-el* is suffixed to monovalent voluntary motion verb roots. (Although an instrument can indeed be associated with motion in Ndebele, it is not introduced by the applicative suffix but by a *nga-* 'by/with' phrase.)

(5) Monovalent verb root + voluntary motion *gijim* 'run'
     (Similar roots: *hamb* 'walk, move', *tshitsh* 'walk fast', *phaph* 'fly', *khas* 'crawl', *eq* 'jump', *ntshez* 'swim')

   a. U-Themba u-ya-gijim-a.   T
       1a-Themba 1a-TNS-run-a

       'Themba is running.'

   b. U-Themba u-ø-gijim-**el**-a   u-nina.   T B
       1a-Themba 1a-TNS-run-APP-a 1a-mother

       'Themba is running for his mother.'

   c. U-Themba u-ø-gijim-**el**-a   u-nina.   T G
       1a-Themba 1a-TNS-run-APP-a 1a-mother

       'Themba is running to his mother.'

   d. U-Themba u-ø-gijim-**el**-a   e-nkundleni.   T G
       1a-Themba 1a-TNS-run-APP-a LOC-stadium

       'Themba is running to the stadium.'

   e. U-Themba u-ø-gijim-**el**-a   u-nina.   T R
       1a-Themba 1a-TNS-run-APP-a 1a-mother

       'Themba is running because of his mother.'

   f. U-Themba u-ø-gijim-**el**-a   e-nkundleni.   T L
       1a-Themba 1a-TNS-run-APP-a LOC-stadium

       'Themba is running in the stadium.'

Verbs of voluntary motion with monovalent verb roots show that the applicative may introduce the BENEFICIARY (5b), GOAL (5c-d), REASON (5e) or LOCATION (5f). In (5) *Themba* is treated as a THEME, not an AGENT because he is definitely the entity that is moving or changing location. While he is animate and can act with volition like an AGENT, it is not clear if he is the initiator of the running and whether or not he is actually acting volitionally. In (5d) and (5f) the prefix *e-* is traditionally treated as a locative marker but (5d) shows that the location can also be a GOAL that an object moves towards.

Verbs of involuntary motion with monovalent verb roots, exemplified in (6), show that the applicative introduces the same thematic roles as in (5) except for the BENEFICIARY, as none of the actions implied by each of the verbs in this subclass can be done on behalf of another person or thing.

(6) Monovalent verb root + involuntary motion *balek* 'flee'
(Similar roots: *-w* 'fall, drop', *gelez* 'flow', *ntshaz* 'squirt')

   a. U-Themba u-ya-baleka-a.   T
      1a-Themba 1a-TNS-flee-a

      'Themba is fleeing.'

   b. U-Themba u-ø-balek-**el**-a   in-yoka.   T R
      1a-Themba 1a-TNS-flee-APP-a 9-snake

      'Themba is fleeing from the snake.'

   c. U-Themba u-ø-balek-**el**-a   e-ndlini.   T G
      1a-Themba 1a-TNS-flee-APP-a LOC-house

      'Themba is fleeing into the house.'

   d. U-Themba u-ø-balek-**el**-a   e-ndlini.   T L
      1a-Themba 1a-TNS-run-APP-a LOC-house

      'Themba is fleeing in the house.'

Themba is again here treated as a THEME for the same reasons as in (5), and the locative prefix can still introduce a GOAL (6c).

Possible thematic roles added by the applicative are even fewer with verbs of motion whose roots are divalent, as seen in (7) where only BENEFICIARY and REASON are permissible.

(7) Verb of motion with divalent verb root *y* 'go to'
(Similar roots: *z* 'come', *suk* 'depart, leave', *fik* 'arrive')

   a. U-Themba u-ø-ya   e-sitolo.   T G
      1a-Themba 1a-TNS-go-a LOC-store

      'Themba is going to the store.'

   b. U-Themba u-ø-y-**el**-a[2]   u-mama e-sitolo.   T B G
      1a-Themba 1a-TNS-go-APP-a 1a-mother LOC-store

      'Themba is going to the store for mother.'

   c. U-Themba u-ø-y-**el**-a   u-mama e-sitolo / e-sitolo u-mama.
      1a-Themba 1a-TNS-go-APP-a 1a-mother LOC-store / LOC-store 1a-mother

                                                     T R G / G R

      'Themba is going to the store for (because of) mother.'

   d. U-Themba u-ø-y-**el**-a   uku-sebenza e-sitolo /e-sitolo uku-sebenza.
      1a-Themba 1a-TNS-go-APP-a 15-work   LOC-store /LOC-store 15-work

                                                       T R G / G R

      'Themba is going to the store (in order) to work.'

Each of these motion verbs with divalent roots already has a GOAL or SOURCE role as a base argument. The root *suk* 'depart' is one example with a SOURCE rather than a GOAL.

---

[2] I do not include the 'defecate on/at' metaphorical meaning of *yela*.

4.2 Verbs of surface contact through motion

Another interesting verb class is that of surface contact through motion. Some verbs in this class have more than one base argument frame as their roots may subcategorized for a noun or locative object (besides a phrase beginning with *ukuze/ukuthi* 'so that'). For most verbs in this category, *-el* introduces a BENEFICIARY, GOAL, REASON or LOCATION if the object of the unapplicativized verb is expressed as a noun, as illustrated in (8).

(8) Divalent or trivalent With noun object
     (Similar roots: *esula* 'wipe', *sunduz* 'push', *fuq* 'push', *dons* 'pull', *nind* 'smear', *gcob* 'smear')

  a. U-Musa u-ø-thanyel-a   izi-bi.   A T
     1a-Musa 1a-TNS-sweep-a 8-trash

     'Musa is sweeping the trash.'

  b. U-Musa u-ø-thanyel-**el**-a   u-nina   izi-bi.   A B T
     1a-Musa 1a-TNS-sweep-APP-a 1a-mother 8-trash

     'Musa is sweeping the trash for her mother.'

  c. U-Musa u-ø-thanyel-**el**-a   u-nina   izi-bi.   A G T
     1a-Musa 1a-TNS-sweep-APP-a 1a-mother 8-trash

     'Musa is sweeping the trash to her mother.'

  d. U-Musa u-ø-thanyel-**el**-a   u-nina   izi-bi.   A R T
     1a-Musa 1a-TNS-sweep-APP-a 1a-mother 8-trash

     'Musa is sweeping the trash because of her mother.'

  e. U-Musa u-thanyel-**el**-a   izi-bi   pha-ndle.   A T L
     cl.1a-Musa cl.1a-TNS-sweep-APP-a 8-trash 15-outside

     'Musa is sweeping the trash outside.'

  f. U-Musa u-thanyel-**el**-a   izi-bi   pha-ndle.   A T G
     cl.1a-Musa cl.1a-TNS-sweep-APP-a 8-trash 15-outside

     'Musa is sweeping the trash (to the) outside.'

Without a derivational suffix such as *−el*, only the verb roots *thanyel* 'sweep', *esul* 'wipe' and *hlikihl* 'wipe, scrab' may subcategorize for a locative object, as shown in (9a). When on a verb root that has a root-determined locative argument, *-el* can only introduce an applied argument with the role of BENEFICIARY or REASON (9b-c).

(9) Divalent verb roots with locative object
     (Similar roots: *esula* 'wipe' and *hlikihl* 'wipe, scrab'; null instantiation: *nind* 'smear' & *gcoba* 'smear')

  a. U-Musa u-thanyel-a   pha-nsi.   A L
     1a-Musa 1a-TNS-sweep-a 16-down

     'Musa is sweeping the floor.'

b. U-Musa u-ø-thanyel-**el**-a u-nina pha-nsi. A B L
   1a-Musa 1a-TNS-sweep-APP-a 1a-mother 16-down

   'Musa is sweeping the floor for her mother.'

c. U-Musa u-ø-thanyel-**el**-a uku-lala pha-nsi. A R L
   1a-Musa 1a-TNS-sweep-APP-a 15-sleep 16-down

   'Musa is sweeping the floor (in order) to sleep.'

Note that roots such as *nind* 'smear' and *gcob* 'smear' can take a null-instantiated locative object because they are actually three place verbs. Null-instantiation is possible due to the fact that these verb roots are also acceptable with divalent argument frames, although the omitted object, usually with a THEME semantic role, can be recovered through logical reasoning. Their normal behavior when *-el* is suffixed is illustrated in (10) where *-el* introduces the BENEFICIARY, REASON or LOCATION. The GOAL role is determined by the base verb root itself.

(10) Trivalent root

a. U-Musa u-ø-gcob-a u-sana ama-futha. A G T
   1a-Musa 1a-TNS-smear-a 11-baby 6-oil

   'Musa is smearing the oil onto the baby.' /
   'Musa is smearing the baby with oil.'

b. U-Musa u-ø-gcoba-**el**-a u-nina u-sana ama-futha. A B G T
   1a-Musa 1a-TNS-smear-APP-a 1a-mother 11-baby 6-oil

   'Musa is smearing the oil onto the baby for mother.' /
   'Musa is smearing the baby with oil for mother.'

c. U-Musa u-ø-gcoba-**el**-a u-nina u-sana ama-futha. A R G T
   1a-Musa 1a-TNS-smear-APP-a 1a-mother 11-baby 6-oil

   'Musa is smearing the oil onto the baby because of mother.' /
   'Musa is smearing the baby with oil because of mother.'

d. U-Musa u-ø-gcob-**el**-a u-sana ama-futha pha-ndle. A G T L
   1a-Musa 1a-TNS-smear-APP-a 11-baby 6-oil 15-outside

   'Musa is smearing the oil onto the baby outside.' /
   'Musa is smearing the baby with oil outside.'

Where there is a GOAL, as in (10a), it is also possible to drop it if that GOAL and the AGENT are co-referential (i.e. *UMusa ugcoba uMusa amafutha* is normally expressed as *UMusa ugcoba amafutha* 'Musa is smearing the oil onto herself'/'Musa is smearing herself with oil'). However, applicativization is odd in the absence of the GOAL.

4.3 Verbs of surface contact

Verbs of surface contact are similar to those of surface contact through motion except that with the former the applicative suffix generally does not introduce the GOAL sematic

role. For most roots in this class, including *mukul* 'slap', *wakal* 'slap' and *tshay* 'hit', the GOAL is excluded because neither the AGENT nor PATIENT moves towards any specific entity. Examples are provided in (11) with the verb root *tshay* 'slap'.

(11) Surface-contact root *tshay* 'slap'
(Similar roots: *tshay* 'hit', *khab* 'kick', *ang* 'kiss', *qabul/qabuj/qabuz* 'kiss', *mukul, makal/wakal/waqaz* 'slap', *bhansul* 'slap lightly', *thint* 'touch')

    a. U-Themba u-ø-tshay-a in-yoka.   A P
       1a-Themba 1a-TNS-hit-a 9-snake
       'Themba is hitting a snake.'

    b. U-Themba u-ø-tshay-*el*-a   u-Musa in-yoka.   A B P
       1a-Themba 1a-TNS-hit-APP-a 1a-Musa 9-snake
       'Themba is hitting the snake for Musa.'

    c. U-Themba u-ø-tshay-*el*-a   in-yoka umu-thi.   A P R
       1a-Themba 1a-TNS-hit-APP-a 9-snake 4-medicine
       'Themba is hitting the snake for medicine.'

    d. U-Themba u-ø-tshay-*el*-a   in-yoka pha-ndle.   A P L
       1a-Themba 1a-TNS-hit-APP-a 9-snake 16-outside
       'Themba is hitting the snake outside.'

For a few roots such as *khab* 'kick' and *waqaz* 'slap', the PATIENT object may be treated as a THEME if it is viewed as a moving entity. As seen with the root *khab* 'kick' in (12e), the locative object introduced by *-el* is then treated as a GOAL. Examples (12d) and (12e) actually show that there are two separate argument frames: AGENT-PATIENT-LOCATIVE, and AGENT-THEME-GOAL. It seems we get a THEME and GOAL reading in (12e) because such verbs have dual membership. They also become members of the class of verbs of surface contact through motion if more force is exerted on the object and the object yields. In other words, whether or not they have a GOAL role is dependent on the amount of force exerted and the weakness of the entity to which force is being applied.

(12) Exceptions (Dual membership), e.g. *khab* 'kick'
(Similar roots: *waqaz* 'slap', *khahlel* 'kick', *gqubul/gqikil* 'head butt', *hlankal/muhluz* 'slap')

    a. U-Themba u-ø-khab-a   um-duli.   A P
       1a-Themba 1a-TNS-kick-a 4-wall
       'Themba is kicking the wall.'

    b. U-Themba u-ø-khab-*el*-a   u-Musa um-duli.   A B P
       1a-Themba 1a-TNS-kick-APP-a 1a-Musa 4-wall
       'Themba is kicking the wall for Musa.'

    c. U-Themba u-ø-khab-*el*-a   um-duli i-mali.   A P R
       1a-Themba 1a-TNS-kick-APP-a 4-wall   9-money
       'Themba is kicking the wall for money.'

    d. U-Themba u-ø-khab-*el*-a   um-duli pha-ndle.   A P L

           1a-Themba 1a-TNS-hit-APP-a 4-wall   16-outside

           'Themba is kicking the wall outside.'

    e. U-Themba u-ø-khab-*el*-a   um-duli pha-ndle.   A T G

           1a-Themba 1a-TNS-hit-APP-a 4-wall   16-outside

           'Themba is kicking the wall (to the) outside.'

With roots that inherently imply little force, such as *bhansul* 'slap lightly (at the back)', *thint* 'touch' and *-anga* 'kiss', there is never a THEME and GOAL reading.

4.4 Involuntary processes (no BENEFICIARY)

Verbs of involuntary processes can be divided into two subgroups: those that involve motion and those that do not. Verbs in both subcategories take a PATIENT subject. Where motion is involved, the subject could also be viewed as a THEME. However, there is no complete change of location since the subject does not totally leave the point of origin. On the basis of there being no complete change in location from the point of origin, I treat the subject as a PATIENT. In fact, the subject can be seen more as coming out changed from the actions of a CAUSER than from movement. For instance, what is at the fore in (13a) is that something is causing the tree to grow (changing it from small to big), not to move (from point A to B). To the verbs of involuntary processes that encode motion, the applicative supplies a GOAL, LOCATION or REASON as in (13b), (13c) and (13d), respectively.

(13)   +Motion

       (Similar root: *ncibilik* 'melt')

      a. Isi-hlahla si-ya-khula.   P

         7-tree   7-TNS-grow-a

         'The tree is growing.'

      b. Isi-hlahla si-khul-*el*-a   ko-makhelwane.   P G

         7-tree   7-TNS-grow-APP-a LOC-neighbor

         'The tree is growing towards/into the neighbors.'

      c. Isi-hlahla si-khul-*el*-a   ko-makhelwane.   P L

         7-tree   7-TNS-grow-APP-a LOC-neighbor

         'The tree is growing at the neighbors.'

      d. Isi-hlahla si-khulela   uku-thi si-dl-iw-e.   P R

         7-tree   7-TNS-grow-APP-a 15-that 7-eat-PASS-TNS

         'The tree is growing so that it will be eaten.'

Although the root *ncibilik* 'melt' belongs to this class where the single argument undergoes a change of state, it has dual class membership as it can also be treated like a THEME-GOAL verb of motion (contrasting with the PATIENT-GOAL frame similar to 13b).

This is, however, only possible when motion, not change of state, is at the fore and after applicativazation, unlike with the verb root *y* 'go to', for example, where the THEME-GOAL reading occurs before applicativization. An example of the THEME-GOAL semantic frame is in (14) where *phansi* can refer to the ground or floor.

(14)   Verb of motion: root *ncibilik* 'melt'
U-ngqwaqwane w-a-ncibilik-*el*-a   pha-nsi.   T G (or M)
1a-ice   1a-TNS-melt-APP-a 16-down

'The ice melted onto the ground/floor.'

Without motion, verbs of involuntary processes introduce the same roles as (13) except for the GOAL. This is illustrated (15).

(15)   −Motion
(Similar root: *f* 'die; break; break down')

a. In-hlama i-y-om-a.   P
9-dough 9-TNS-dry-a

'The dough dries.'

b. In-hlama y-om-*el*-a   pha-ndle.   P L
9-dough 9-TNS-dry-APP-a 15-outside

'The dough dries outside.'

c. In-hlama y-om-*el*-a   ukuthi i-langa li-ya-tshisa.   P R
9-dough 9-TNS-dry-APP-a that   5-sun   5-TNS-hot

'The dough dries because the sun is hot.'

While the literature supports BENEFICIARY as the most common role associated with applicative constructions, it is clear from (13) and (15) that *-el* does not introduce this role to verbs of involuntary process. The BENEFICIARY/MALEFICIARY role may arguably only be inferred in very specific circumstances that also involve something good or bad happening to a LOCATION or GOAL as a result of the process. That is, a BENEFI-CACTIVE/MALEFICACTIVE reading can be inferred for verbs of this class only if they also have membership in another class, as in (14).

4.5 State verbs (no BENEFICIARY)

Example (16) shows that state verbs have a PATIENT subject and behave exactly like those represented by (15) when the applicative *-el* is suffixed. The single argument of the verb here is a PATIENT rather than a THEME because there is no clear movement or change of location, but the participant in these stative verbs may change state, for example, from hot to cold in (16b-c) even if 'hot' is not mentioned.

(16) State root *qanda* 'be cold'
(Similar roots: *phil* 'be alive', *khudumal* 'be warm', *bil* 'boil')

    a. Ama-nzi a-ya-qand-a.   P
       6-water 6-TNS-cold-a

       'The water is cold.'

    b. Ama-nzi a-qand-*el*-a   e-mbiz-eni.   P L
       6-water 6-TNS-cold-APP-a LOC-pot-LOC

       'The water is/becomes cold in the pot.'

    c. Ama-nzi a-qand-*el*-a   uku-thi a-se-friji-ni.   P R
       6-water 6-TNS-cold-a 15-that 6-LOC-refrigerator-LOC

       'The water is/becomes cold because it is in the refrigerator.'

As can be seen, when *-el* is suffixed to the root, the thematic role of the new participant can only be LOCATION (16b) or REASON (16c), not BENEFICIARY. In fact Ndebele verbs exemplified by (16), although not unambiguously verbs of involuntary process, can be better classified with those in (15) since they also often involve a change of state when *-el* is suffixed to the root.

4.6 Verbs of feeling (no BENEFICIARY)

The last class considered contains verbs of feeling that normally surface with two arguments with the sematic roles of EXPERIENCER (E) and THEME (or STIMULUS STI).

(17) Verbs of feeling: Verb root *esab* 'be scared/afraid'
(Similar root: *-enyany* 'be disgusted (by something)')

    a. u-Themba w-e-sab-a   in-yoka.   E T
       1a-Themba 1a-TNS-afraid-a 9-snake

       'Themba was afraid/scared of the snake.'

    b. u-Themba w-e-sab-*el*-a   in-nyoka uku-luma.   E T R
       1a-Themba 1a-TNS-afraid-APP-a 9-snake   15-bite

       'Themba got scared of the snake because it bites.'

    c. u-Themba w-e-sab-*el*-a   in-nyoka e-gusw-ini.   E T L
       1a-Themba 1a-TNS-afraid-APP-a 9-snake   LOC-forest-LOC

       'Themba got scared of the snake in the forest.'

Example (17) shows that for verbs of feeling, the argument introduced by the applicative may have a REASON or LOCATION semantic role. Again here the BENEFICIARY is completely excluded.

4.7 Summary: The applicative with verbs in different semantic classes

Table 2 presents argument frames associated with applicative constructions formed from roots of different semantic classes. Thematic roles in the argument fames are presented in plain type, and those associated with the argument introduced by applicative *-el* are in bold and vary across B, G, L and R.

Table 2: Argument frames by verb class

Verb class	Thematic role frames

1. motion	a. Monovalent root + voluntary motion
	     T**B** T**G** T**L** T**R**
	 b. Monovalent root + involuntary motion
	     T**G** T**L** T**R**
	 c. Divalent root
	     TB**G** TR**G**/TG**R**

2. surface contact
     through motion
	a. Divalent root with applied noun object
	     A**B**T A**G**T A**R**T AT**L** AT**G**
	 b. Divalent root with applied locative object
	     A**B**L A**R**L
	 c. Trivalent root
	     A**B**GT A**R**GT AGT**L**

3. surface contact (no motion)	A**B**P A**P**R A**P**L
4. involuntary processes	P**G** P**L** P**R**
5. state	P**L** P**R**
6. feeling	E**T**R E**T**L
7. sound emission	EM.**R** EM.**L**

As can be seen, *-el* can introduce REASON and LOCATION thematic roles in all the classes we have seen above. REASON can also be introduced in all subclasses, which translates to all verbs. Note the absence of LOCATION when a verb of motion has a divalent root (case 1c in Table 2). The BENEFICIARY role does not feature at all in the last four classes in Table 2. However, although the first two classes involve motion and the last four do not, it would be premature to conclude that the BENEFICIARY thematic role occurs only in classes where motion is involved since this study has not exhaustively covered all semantic verb classes.

5 Conclusion

The discussion has shown that there is clear variation across semantic classes and sometimes also a little variation within classes in terms of thematic roles added by the applicative. While the variation across sematic classes can be captured by identifying the different thematic roles assigned to arguments in each class and those that applicative *-el* introduces, the variations within a given semantic class may be due to differences in transitivity, the number of participants associated with the roots, and whether or not the action associated with the verb is voluntary. Also, some classes overlap, resulting in some verbs not being the best representatives of their classes. In spite of the variation within classes, thematic roles of participants associated with a given Ndebele verb are generally predictable from the semantic class of the root including those of participants introduced by applicative *-el*. It has been shown that REASON and LOCATION are the thematic roles *-el* introduces in all semantic classes, and not BENEFICIARY counter to many other studies of applicatives (e.g., Schadeberg 2003; Peterson 2007; Polinsky 2008; Kittilä & Zúñiga 2010; De Kind & Bostoen 2012; and Marten & Kula 2014). The Ndebele applicative was also shown to introduce arguments with roles often thought of as either participants (for example, BENEFICIARY, SOURCE, GOAL) or nonparticipants (LOCATION, REASON).

An issue that arises from the discussion is whether there is a precise or "single" meaning of the applicative suffix *-el*. Due to English influence, it is tempting to conclude that the applicative *-el* is a polysemous suffix[3] that takes various "prepositional" meanings such as 'for', 'at', 'in', 'on', and so on, which serve as cues for thematic roles. For example English *at* is a cue for LOCATION and *for* signals BENEFICIARY or REASON. However, a close examination of the use of *-el* suggests that it is an underspecified suffix[4] whose "prepositional" information, such as 'for', 'at', 'in' and 'on' is largely determined by the verb class. It appears that *-el* encodes a very general relationship such as 'extra argument', and further semantic specificity is available from the verb itself if we know its semantic class. For example, in Ndebele native speakers already know which semantic classes are compatible with BENEFICIARY and REASON roles, so there is no need to specify any specific preposition-like meaning inhering "in" the applicative morpheme. If semantic class information is known then there is no need for detailed information in the applicative.

In short, the discussion above has shown that the Ndebele Applicative Construction suffixes the applicative *-el* to the verb which adds a new argument, and the thematic role of the argument is constrained by the semantic class of the verb and context. The results of this study can subsequently be used as a test tool for evaluating class membership of additional verbs, since verbs of the same class take similar thematic roles. If the applicative argument triggers an irregular thematic role then the verb does not belong to the expected class or, at best, has dual membership.

---

[3] For an argument in support of polysemy, though accepting underspecification to some extent, see Mabugu's (2011) analysis of the Shona applicative construction.

[4] Marten (2002) also advances an argument for underspecification, although his analysis does not focus on semantic classes, but is driven by the Relevance-Theoretic notion of concept strengthening.

Acknowledgments

Abbreviations

As in Bantuist tradition, numbers indicate noun class prefixes or agreement markers (e.g. 1a = noun class 1a prefix or agreement marker).

A	AGENT	LOC	LOCATIVE AFFIX
APP	APPLICATIVE	M	MALEFICIARY
AO	APPLIED OBJECT	P	PATIENT
B	BENEFICIARY	PUR	PURPOSE
C	CAUSER	R	REASON
E	EXPERIENCER	S	SOURCE
EM	EMITTER	STI	STIMULUS
G	GOAL	T	THEME
L	LOCATION	TNS	TENSE
LO	LOGICAL OBJECT	-a	default final vowel for verbs

References

Alsina, Alex & Sam Mchombo. 1993. Object asymmetries and the Chicheŵa applicative construction. In Sam Mchombo (ed.), *Theoretical aspects of Bantu grammar*, 17–45. Stanford: CSLI Publications.

Anderson, Stephen R. 1971. The role of deep structure in semantic interpretation. *Foundations of Language* 6. 387–96.

Anderson, Stephen R. 1977. Comments on the paper by Wasow. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 361–378. New York: Academic Press.

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16. 373–415.

Baker, Mark. 1988. Theta theory and the syntax of applicatives in Chichewa. *Natural Language and Linguistic Theory* 6. 353–380.

Blake, Frank R. 1930. A semantic analysis of case. In James T. Hatfield, Werner Leopold & A. J. Friedrich Zieglschmid (eds.), *Curme volume of linguistic studies* (Language Monographs 7), 34–49. Baltimore, MD: Linguistic Society of America.

Bresnan, Joan & Jonni Kanerva. 1989. Locative inversion in Chicheŵa: A case study of factorization in grammar. *Linguistic Inquiry* 20 (1). 1–50.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21. 147–185.

Cann, Ronnie & Patricia Mabugu. 2007. Constructional polysemy: The applicative construction in chiShona. *Metalinguistica* 19. 221–245.

Cardona, George. 1974. Panini's karakas: Agency, animation and identity. *Journal of Indian Philosophy* 2. 231–306.

Chafe, Wallace L. 1970. *Meaning and the structure of language.* Chicago: University of Chicago Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Cook, Walter A. 1979. *Case grammar: Development of the Matrix Model (1979-1978).* Washington, DC: Georgetown University Press.

Cruse, D. Alan. 1973. Some thoughts on agentivity. *Journal of Linguistis* 9. 1–204.

Culicover, Peter & Wendy Wilkins. 1986. Control, PRO and the projection principle. *Language* 62. 120–153.

De Kind, Jasper & Koen Bostoen. 2012. The applicative in ciLubà grammar and discourse: A semantic goal analysis. *Southern African Linguistics and Applied Language Studies* 30(1). 101–124.

Dowty, David. 1987. Thematic proto roles, subject selection and lexical semantic defaults. (Paper presented at the 1987 Linguistic Society of America Annual Meeting.)

Dowty, David. 1989. On the semantic content of the notion of thematic role. In G. Cherchia, B. Partee & R. Turner (eds.), *Properties, types and meaning, vol. 2,* 69–129. Dordrecht: Kluwer Academic Publishing.

Dowty, David. 1991. Thematic proto roles and argument selection. *Language* 67(3). 547–619.

Endemann, Karl. 1876. *Versuch einer Grammatik des Sotho.* Berlin: W. Hertz.

Fillmore, Charles J. 1968. The case for case. In E. Bach & R. T. Harms (eds.), *Universals in linguistic theory*, 1–88. New York: Holt, Rinehart & Winston.

Fillmore, Charles J. 1977. The case for case reopened. In Peter Cole & Jerrold Sadock (eds.), *Syntax and semantics 8: Grammatical relations*, 59–81. New York: Academic Press.

Fillmore, Charles J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* VI(2). 222–254.

Fillmore, Charles J. 1986. Pragmatically controlled zero anaphora. In Vassiliki Nikiforidou, Mary VanClay, Mary Niepokuj & Deborah Feder (eds.), *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, 95–107.

Fillmore, Charles J. & Paul Kay. 1999. *Construction grammar coursebook, Ling X20.* Berkeley: University of California.

Foley, William A. & Jr. Van Valin Robert D. 1984. *Functional syntax and universal grammar.* Cambridge: Cambridge University Press.

Frawley, William. 1992. *Linguistic semantics.* Hillsdale, NJ: Lawrence Erlbaum.

Gruber, Jeffrey. 1965. *Studies in lexical relations.* Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation.)

Halliday, M. A. K. 1970. Language structure and language function. In John Lyons (ed.), *New horizons in linguistics*, 140–65. Harmondsworth: Penguin.

Harford, Carolyn. 1993. The applicative in chiShona and lexical mapping theory. In Sam Mchombo (ed.), *Theoretical aspects of Bantu grammar 1,* 93–111. Stanford, CA: CSLI Publications.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar.* Cambridge, MA: MIT Press.

Jackendoff, Ray S. 1976. Toward an explanatory semantic representation. *Linguistic Inquiry* 7(1). 89–150.

Jackendoff, Ray S. 1982. Commentary on Bowerman, "Evaluating competing linguistic models". *Quaderni di Semantica* 5. 67–71.

Jackendoff, Ray S. 1983. *Semantics and cognition.* Cambridge, MA: The MIT Press.

Jackendoff, Ray S. 1987. The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18(3). 369–411.

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

Jerro, Kyle. 2016. *The syntax and semantics of applicative morphology in Bantu.* Austin, TX.

Kähler-Meyer, E. 1966. Die örtliche Funktion der Applikativendung in Bantusprachen. In J. Lukas (ed.), *Neue afrikanistische Studien: Festschrift für A. Klingenheben*, 126–136. Hamburg: Deutsches Institut für Afrika-Forschung.

Kittilä, Seppo, Vasti Katja & Jussi Ylikoski. 2011. Introduction to case, animacy, and semantic roles. In Seppo Kittilä, Vasti Katja & Jussi Ylikoski (eds.), *Case, animacy, and semantic roles*, 1–28. Amsterdam: John Benjamins.

Kittilä, Seppo & Fernando Zúñiga. 2010. Introduction. Benefaction and malefaction from a cross-linguistic perspective. In Fernando. Zúñiga & Seppo Kiittilä (eds.), *Benefactives and malefactives: Typological perspectives and case studies*, 1–28. Amsterdam /Philadelphia: John Benjamins.

Kiyosawa, Kaoru & Donna B. Gerdts. 2010. Benefactive and malefactive uses of Salish applicatives. In Seppo Kittilä & Fernando Zúñiga (eds.), *Benefactives and malefactives: Typological perspectives and case studies*, 147–184. Amsterdam: John Benjamins.

Lakoff, George. 1977. Linguistic gestalts. *Chicago Linguistic Society* 13. 236–87.

Levin, Beth & Malka Rappaport Hovav. 1996. Lexical semantics and syntactic structure. In Shalom Lappin (ed.), *The handbook of contemporary semantic theory*, 487–507. Malden, MA: Blackwell.

Löbner, Sebastian. 2002. *Understanding semantics.* New York: Arnold Publishers & Oxford University Press.

Longacre, Robert E. 1976. *An anatomy of speech notions.* Lisse: Peter de Ridder Press.

Mabugu, Patricia. 2011. *Explaining the ChiShona applicative construction: A case of polysemy.* Saarbrücken: Lambert Academic Publishing.

Machobane, Malillo. 1989. *Some restrictions of transitivising morphemes in Sesotho.* Montreal: McGill University. (Doctoral dissertation).

Marantz, Alec. 1984. *On the nature of grammatical relations.* Cambridge, MA: MIT Press.

Marten, Lutz. 2002. *At the syntax-pragmatics interface.* Oxford: Oxford University Press.

Marten, Lutz & Nancy C. Kula. 2014. Benefactive and substitutive applicatives in Bemba. *Journal of African Languages and Linguistics* 35(1). 1–44.

Mchombo, Sam. 2004. *The syntax of Chichewa*. Cambridge: Cambridge University Press.

Newmeyer, Frederick. 2010. On comparative concepts and descriptive categories: A reply to Haspelmath. *Language* 86(3). 688–695.

Ngonyani, Deo. 1996. *The morphosyntax of applicatives*. Los Angeles: UCLA. (Doctoral dissertation.)

Ngonyani, Deo & Peter Githinji. 2006. The asymmetric nature of Bantu applicative constructions. *Lingua* 116. 31–63.

Palmer, Martha, Daniel Gildea & Nianwen Xue. 2010. *Semantic role labeling*. San Rafael, California: Morgan & Claypool.

Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.

Peterson, David A. 2007. *Applicative constructions*. Oxford: Oxford University Press.

Polinsky, Maria. 2008. Applicative constructions. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures online*, Chapter 109. Munich: Max Planck Digital Library.

Pollard, Carl & Ivan Sag. 1994. *Head-driven phrase structure grammar*. Stanford: Center for the Study of Language & Information.

Pylkkänen, Liina. 2000. What applicative heads apply to. *University of Pennsylvania Working Papers in Linguistics* 7(1). http://repository.upenn.edu/pwpl/vol7/iss1/18.

Rozwadowska, Bozena. 1988. Thematic restrictions on derived nominal. In Wendy K. Wilkins (ed.), *Syntax and semantics vol. 21: Thematic relations*, 147–165. New York: Academic Press.

Saint-Dizier, Patric & Evelyne Viegas. 1995. An introduction to lexical semantics from a linguistic and a psycholinguistic perspective. In Patric Saint-Dizier & Evelyne Viegas (eds.), *Computational lexical semantics*, 1–29. Cambridge: Cambridge University Press.

Schadeberg, Thilo C. 2003. Derivations. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 71–89. London: Routledge.

Schulte im Walde, Sabine & Chris Brew. 2002. Inducing German semantic verb classes from purely syntactic subcategorisation information. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, July*, 223–230.

Schumacher, Helmut. 1986. *Verben in Feldern*. Berlin: de Gruyter.

Srikumar, Vivek. 2013. *The semantics of role labeling*. Urbana: University of Illinois at Urbana-Champaign. (Doctoral dissertation).

Stowell, Timothy. 1981. *Origins of phrase structure*. Cambridge, MA: MIT Dissertation.

Talmy, Leonard. 1985a. Figure and ground as thematic roles. (Paper presented at the 1985 Annual meeting of the Linguistic Society of America, Seattle.)

Talmy, Leonard. 1985b. Lexicalization patterns: Semantic structure in lexical form. In Timothy Shopen (ed.), *Language typology and syntactic description: Grammatical categories and the lexicon*, vol. III, 57–149. Cambridge: Cambridge University Press.

Trithart, Mary Lee. 1983. *The applied affix and transitivity: A historical study in Bantu*. Los Angeles: UCLA. (Doctoral dissertation).

Van Eeden, Bernardus I. C. 1956. *Zoeloe-grammatika*. Stellenbosch: Die Universitieitsuit-gewers en Boekhandelaars.

Van Valin, Jr., Robert D. 1990. Semantic parameters of split intransitivity. *Language* 66(2). 221–260.

Van Valin, Jr., Robert D. 1993. A synopsis of Role and Reference Grammar. In Jr. Van Valin Robert D. (ed.), *Advances in Role and Reference Grammar*, 1–164. Amsterdam: John Benjamins.

Vázquez, Gloria, Ana Fernández, Irene Castellón & M. Antonia Martí. 2000. *Clasificación verbal: Alternancias de diátesis* (Quaderns de Sintagma 3). Lleida: Universitat de Lleida.

Chapter 18

Differential object marking in Mozambican languages

Armindo Saúl Atelela Ngunga
University of Eduardo Mondlane

Fábio Bonfim Duarte
Federal University of Minas Gerais

Quesler Fagundes Camargos
Federal University of Rondônia

This article investigates differential object marking (DOM) in four Bantu languages spoken in Mozambique. In Shimakonde DOM is triggered when the direct object is an animate noun of classes 1, 2, 5, 7 and 9, whereas in Emakhuwa DOM can be understood as the consequence of a noun class hierarchy. In Changana and Rhonga DOM is regulated not by animacy, but by definiteness and specificity in that only specific and definite DPs can trigger object agreement in simple and complex transitive predicates. The paper also explores the grammatical status of the object prefixes that occur in the verb. In all four languages the object prefixes behave more like referential agreement than pronoun incorporation because the object prefixes match in person, number and gender/class with the DP object. In these languages, a syntactic adjacency condition is present between the DP objects and the verb in the transitive verb structure, which indicates that the DP object really occurs in an internal argument position. As for double object constructions, the proposal is that Changana and Rhonga are asymmetrical object languages since GOAL arguments carry primary object properties, while THEME arguments do not. This analysis is confirmed by the facts that (i) GOALS must precede THEMES in unmarked ditransitive sentences; (ii) verb agreement only occurs with the GOALS; and (iii) only GOALS can be passivized. In double object constructions, it is postulated that the reason the THEME object is never cross-referenced on the verb is because it is not the closest DP in the search domain of the little *v*. We then hypothesize that the applied GOAL object is generated in a higher position than the THEME object.

Armindo Atelela Ngunga, Fábio Bonfim Duarte & Quesler Fagundes Camargos

1 Introduction

This paper has three main objectives. First, it seeks to examine the grammatical status of the object marker[1] in four Bantu languages that are spoken in Mozambique. In Changana and Rhonga the object marker can optionally co-occur with the DP object (i.e. a lexical or free pronominal expression of the object) in the same clause. However, in Shimakonde the object marker is triggered in the verb morphology if the direct object is an animate noun of classes 1, 2, 5, 7 and 9. Emakhuwa differs slightly in this regard owing to the fact that the object marker is always obligatory whenever the DP expression of the object is of classes 1 or 2, regardless of whether it is animate or not.

Since the 1980s, there has been intense debate over whether object markers in Bantu languages are instances of referential agreement or simply pronominal arguments that are incorporated onto the verbal complex. Bresnan & Mchombo (1987), for instance, propose that object marking in Chichewa is best understood as an incorporated pronoun and not as a referential agreement marker. They assume that when an overt lexical DP co-occurs with the object marker, the latter signals that the DP object has been dislocated to a right or left-peripheral position. Their analysis is based on three arguments: (i) object marking on the verb stem is related to alterations of the basic order of the DP object; (ii) DP objects that are referred to by an object marker in the verbal complex are prosodically separated from the verb; and (iii) focused elements cannot be referred to with an object marker. By contrast, Baker (2008) and Riedel (2009) argue that the object marking in languages such as Chichewa is indeed a manifestation of referential agreement on an active vP. Strong evidence in favor of Baker's analysis comes from the fact that object markers cannot appear on verbs that are in passive or reciprocal voice (see §4.1). Additionally, Downing (2014) contends that the distribution of object markers in Chichewa fails to consistently satisfy all three of the diagnostics for anaphoric status, i.e. as proposed by Bresnan & Mchombo (1987).

Given these theoretical debates in the literature, the second purpose of this paper is to present evidence that the object markers in the four Mozambican languages mentioned above are indeed best analyzed as instances of referential agreement. The third purpose is to bring further syntactic diagnostics to bear in order to demonstrate that the occurrence of the object markers on the verb reflects the fact that the verb agrees with definite and specific objects in a very local syntactic domain. In simple and complex transitive constructions, definite and specific objects move systematically to Spec-vP in order to establish an agree operation with the little v, whereas indefinite objects remain inside the VP. Details of this proposal are developed in §4.

The article is organized into five sections. §2 outlines the theoretical assumptions on which the analysis is based. §3 and §4 present the relevant data that serves to advance the theoretical proposal. We contend that the definiteness scale is relevant in Changana and Rhonga, whereas, apart from the classes 1 and 2 constraint, it is the animacy hierarchy

---

[1] In this paper, we use the term *object marker* to refer to an agreement relation that is established between a definite/specific direct object and a (di)transitive verb. Furthermore, the expression *internal argument* will be used to refer to a direct object that is projected as a complement of a transitive verb.

that determines the differential object agreement in Shimakonde. As for Emakhuwa, we propose that it is classes 1 and 2 that constrain the appearance of the object marker on the verb. §5 concludes the paper.

2 Theoretical background

Research over the last decades across various languages has shown that verbal agreement with a DP object (=internal argument) must usually present certain conditions in order to license differential object agreement. By *differential object agreement* we mean that agreement on the verb may serve as a grammatical device to encode certain semantic differences. In line with this, we follow Comrie's (1981) and Croft's (1988, 1990) assumptions that specificity, animacy and person-number features play a major role regarding the activation of differential object agreement across languages. Within the typological literature (Givón 1976; Comrie 1981; Croft 1988; 1990; Bentley 1994), it has been assumed that the relevant semantic features that trigger object agreement on the verb stem are the ones that occupy a higher position in the hierarchies in (1).

(1) Relevant hierarchies for licensing object agreement

    a. Definiteness Hierarchy: definite > specific > indefinite > non-specific

    b. Animacy Hierarchy: human > animate > inanimate

Differential object agreement constrained by the specificity of the object is attested, for instance, in Palauan. According to Woolford (2000: 218), a definite and specific object in this language triggers verbal agreement, whereas an indefinite object does not (2).

(2) Palauan (Georgopoulos 1991: 30)

    a. te-'illebed-ii   a bilis a rengalek.
        3PL-PRF.hit-3SG dog   children
        'The children hit the dog.'

    b. te-'illebed   a bílis a rengalek.
        3PL-PRF.hit dog   children
        'The children hit a dog/some dogs.'

In addition to verbal agreement, other strategies are also found across languages to convey semantic differences among objects. Danon (2002), for instance, shows that in Hebrew the case particle *et* is obligatorily triggered whenever the object is definite, as in (3).

(3) Hebrew (Danon 2002: 1)

    a. Dan kara **et**   **ha**-sefer.
        Dan read ACC DEF-book
        'Dan reads the book.'

b. *Dan kara **ha**-sefer.
   Dan   read DEF-book
   ('Dan reads a book.')

Sentence (3b) is ungrammatical because the definite object *ha-sefer* 'the book' must be preceded by the case marker *et*. Nonetheless, if the object is indefinite, this case marker cannot appear. Compare (4a-b):

(4) Hebrew (Danon 2002: 1)

   a. Dan kara sefer.
      Dan read book

      'Dan reads a book.'

   b. *Dan kara **et**   sefer.
      Dan   read ACC book

      ('Dan reads a book.')

In sum, the data presented for Palaun and Hebrew clearly indicate that differential object marking, henceforth DOM, is directly connected to the semantic reading of the object in some languages of the world. The strategies of DOM vary from language to language, but the purpose is the same, encoding semantic contrasts such as those outlined in (1). With this theoretical background in mind, the purpose of §3 and §4 is to investigate how DOM operates in Emakhuwa and Shimakonde on the one hand, and in Changana and Rhonga on the other hand.

3 DOM realization in Shimakonde and Emakhuwa

DOM in Emakhuwa[2] is obligatorily triggered in the verb morphology if the direct object is an animate noun of either class 1, if singular as shown in (5a), or class 2, if plural as indicated in (6a). According to our consultants, classes 1 and 2 basically comprise human and animate nouns:

(5) Emakhuwa (human object)

   a. mulopwana o-o-**mu**-on-a   **m'miravo**.
      1.man   1.SM-PST-1.OM-see-FV 1.boy

      'The man has seen the boy.'

   b. *mulopwana o-o-on-a   m'miravo.
      1.man   1.SM-PST-see-FV 1.boy

      ('The man has seen the boy.')

---

[2] The data for this study were collected during fieldwork in Mozambique, where these languages are spoken as mother tongues. In Bantu language studies, noun classes are named by Arabic numerals (hence, 1 means class 1 (and not first person), for instance).

(6) Emakhuwa (animate object)

    a. mulopwana o-o-**wo**-on-a   **apaakha**.
       1.man   1.SM-PST-2.OM-see-FV 2.cat

       'The man has seen the cats.'

    b. *mulopwana o-o-on-a   apaakha.
       1.man   1.SM-PST-see-FV 2.cat

       ('The man has seen the cats.')

On the other hand, animate DPs that belong to other nominal classes never trigger the agreement object prefix on the verb, as shown in (7a), (7c) and (7e) and in the ungrammaticality of (7b), (7d) and (7f) with the object prefix present:

(7) Emakhuwa (animate object)

    a. mulopwana o-h-on-a   nikhule.
       1.man   1.SM-PST-see-FV 3.rat

       'The man has seen the/a rat.'

    b. *mulopwana o-o-**mu**-on-a   nikhule.
       1.man   1.SM-PST-3.OM-see-FV 3.rat

       ('The man has seen the rat.')

    c. mulopwana o-h-on-a   njojo.
       1.man   1.SM-PST-see-FV 5.giraffe

       'The man has seen the/a giraffe.'

    d. *mulopwana o-o-**ni**-on-a   njojo.
       1.man   1.SM-PST-5.OM-see-FV 5.giraffe

       ('The man has seen the giraffe.')

    e. mulopwana o-o-on-a   ekhuluwe.
       1.man   1.SM-PST-see-FV 7.pig

       'The man has seen the/a pig.'

    f. *mulopwana o-o-**yi**-on-a   ekhuluwe.
       1.man   1.SM-PST-7.OM-see-FV 7.pig

       ('The man has seen the pig.')

The ungrammaticality of (7b), (7d) and (7f) is due to the fact that only object markers of classes 1 and 2 are allowed to occur on the verb stem. Furthermore, examples (7a), (7c) and (7e) without the object prefix on the verb stem allow either a definite or indefinite reading of the object. This suggests that the presence versus absence of the object marker does not contribute to the definiteness reading of the referent of the DP object. In conclusion, DOM in Emakhuwa is based on a noun class hierarchy, as follows:

(8)   CL 1/2 > CLs ≠ 1/2 (i.e. classes that are different from 1/2)

Shimakonde differs from Emakhuwa due to the fact that Shimakonde object marking is not limited to classes 1 and 2, which generally include [+HUMAN, +ANIMATE] nouns, but is also extended to animate nouns of other classes, such as singular classes 5, 7, and 9 and their plural counterparts. For the most part, the singular forms in these classes take the nominal class prefixes {li-}, {shi-} and {(i)N-} and the plural forms take {ma-}, {vi-}, {di-}, respectively. The only exception is nouns of class 10 that can take either the prefix {di-} or the prefix {va-}, as shown in the end of Table 2. It is important to point out that in these classes there are both [-ANIMATE, -HUMAN] and [+ANIMATE, -HUMAN] nouns. Compare the data in Tables 1 and 2.

Table 1: Shimakonde: [-HUMAN, -ANIMATE] nouns of classes 5, 7, and 9

Class 5 (Singular)	English	Class 6 (Plural)	Class 2 (Plural)
li-dodo	'leg'	*ma-dodo*	*va-dodo*
li-pela	'pear'	*ma-pela*	*va-pela*
li-pote	'pot, bowel'	*ma-pote*	*va-pote*
li-pipa	'drummer'	*ma-pipa*	*va-pipa*
li-kalale	'basket'	*ma-kalale*	*va-kalale*
ly-atu	'ear'	*ma-atu*	*va-tu*

Class 7 (Singular)	English	Class 8 (Plural)	Class 2 (Plural)
shi-julu	'hat'	*vi-julu*	*vajulu*
shi-latu	'shoe'	*vi-latu*	*valatu*
sh-elo	'sieve'	*vy-elo*	*velo*
shi-ja	'thigh'	*vi-ja*	*va-ja*

Class 9 (Singular)	English	Class 10 (Plural)	Class 2 (Plural)
i-kanywa	'mouth'	*di-kanywa*	*va-kanywa*
im-bedo	'ax'	*di-mbedo*	*va-mbedo*
i-pete	'ring'	*di-pete*	*va-pete*
i-kiti	'chair'	*di-kiti*	*va-kiti*
ing'-owu	'banana'	*di-ng'owu*	*va-ng'owu*
ing'-ope	'face'	*di-ng'ope*	*va-ng'ope*

In Shimakonde, any animate object DP, regardless of the class prefix that the noun itself has, can be referenced in the verb by *m(V)-* 'singular class 1' and *va-* 'plural class 2'. The morphological distribution of the object prefixes {*mu- ~ m- ~ n-*} and {*va-*} that occur on the verb can be understood by the examples in (9) and (10). It is important to point out that these animate object markers occur in the verb, regardless of the class of the noun. If the object prefix does not appear in the verbal complex when the object DP is animate, the sentence becomes ungrammatical, as shown in (9) and (10).

Table 2: Shimakonde: [-HUMAN, +ANIMATE] nouns of classes 5, 7, and 9

Class 5 (Singular)	English	Class 6 (Plural)	Class 2 (Plural)
ly-umu	'frog'	*ma-umi*	**va-umi*

Class 7 (Singular)	English	Class 8 (Plural)	Class 2 (Plural)
shi-n'gila	'scorpion'	*vi-n'gila*	**va-n'gila*
sh-uvi	'leopard'	*vy-uvi*	**va-uvi*
sh-uni	'bird'	*vy-uni*	**va-uni*

Class 9 (Singular)	English	Class 10 (Plural)	Class 2(Plural)
ng'avanga	'dog'	*di-ng'avanga*	*vang'vanga*
namembe	'fly'	*di-namembe*	*vanamembe*
imbudi	'goat'	*di-mbudi*	*vambudi*

(9) Shimakonde

    a. nangu ni-ndi-**n**-kody-a   munu   n'ng'ande.
       I   1SG.SM-PST-1.OM-find-FV 1.person inside.the.house

       'I found someone inside the house.'

    b. *nangu ni-ndi-kody-a   munu   n'ng'ande.
       I   1SG.SM-PST-find-FV 1.person inside.the.house

       ('I found someone inside the house.')

(10) Shimakonde

    a. nangu ni-ndi-**va**-kody-a   vanu   n'ng'ande.
       I   1SG.SM-PST-2.OM-find-FV 2.people inside.the.house

       'I found some people inside the house.'

    b. *nangu ni-ndi-kody-a   vanu   n'ng'ande.
       I   1SG.SM-PST-find-FV 2.people inside.the.house

       ('I found some people inside the house.')

Examples (9) and (10) illustrate these facts for nouns of classes 1 and 2. But in addition to objects of class 1 and class 2, the object markers are also extended to cross-reference animate nouns belonging to classes 5, 7 and 9. This thus signals that it is not noun class but animacy that regulates DOM in Shimakonde. Notice that there is a mismatch in the glossing in (11) as the object prefix {*n-*} that appears on the verb stems is of class 1, whereas the DP objects belong to classes 5 and 7, respectively. To explain this mismatch one might assume that {*mu- ~ m- ~ n-*} 'singular' and {*va-*} 'plural' have become specialized as differential object markers to indicate animacy in Shimakonde, regardless of whether the object belongs to the (singular) classes 1, 5, 7 or 9.

(11) Shimakonde

    a. nangu ni-ndi-**n**-nivat-a   lipuluputu.
       I   1SG.SM-PST-1.OM-step-FV 5.butterfly

     'I stepped on the butterfly (a definite one).'

    b. nangu ni-ndi-**n**-kang'an-a   shiboko na   upanga.
       I   1SG.SM-PST-1.OM-poke-FV 7.hippo with 14.machete

     'I have poked the hippo (a definite one) with a machete.'

    c. nangu ni-ndi-**n**-gang'ol-a   mbudi.
       I   1SG.SM-PST-1.OM-push-FV 9.goat

     'I pushed the goat (a definite one).'

    That animacy really matters to trigger DOM in Shimakonde becomes evident in contexts where the object is not an animate noun, but an inanimate one. In such contexts, the object prefix cannot appear in the verb morphology (12a-13a). Then, if the inanimate DP object triggers the agreement prefix, the sentence becomes ungrammatical (12b-13b):

(12) Shimakonde

    a. ngw-on-ile   lijanga.
      1SG.SM-see-PST 5.stone

     'I saw a stone.'

    b. *ngw-li-on-ile   lijanga.
      1SG.SM-5.OM-see-PST 5.stone

     ('I saw the stone (a definite one).')

(13) Shimakonde

    a. nangu ni-ndy-on-a   shijiko.
      I   1SG.SM-PST-see-FV 7.spoon

     'I saw a spoon.'

    b. *nangu ni-ndy-shi-on-a   shijiko.
      I   1SG.SM-PST-7.OM-see-FV 7.spoon

     ('I saw this spoon (a definite one).')

    Nonetheless, the object prefix {*n-*} is omitted when the animate object referent is indefinite and not given in the previous discourse. This indicates that the occurrence of this prefix is regulated not only by animacy but also by the definiteness property of DP objects that belong to nouns of classes 5, 7 and 9, as comparison of (11a-c) and (14a-c) shows.

(14) Shimakonde

    a. nangu ni-ndi-nivat-a   lipuluputu.
      I   1SG.SM-PST-step-FV 5.butterfly

     'I stepped on a butterfly.'

b. nangu ni-ndi-kang'an-a   shiboko na   upanga.
   I   1SG.SM-PST-poke-FV 7.hippo with 14.machete

   'I have poked a hippo with a machete.'

c. nangu ni-ndi-gang'ol-a   imbudi.
   I   1SG.SM-PST-push-FV 9.goat

   'I pushed a goat (an indefinite one).'

In conclusion, the examples in (14) demonstrate that when the object prefix is absent the referent of the DP object must be interpreted as indefinite, whereas the presence of the prefix entails that the object constitutes given information in discourse. An alternative way to encode definiteness of the referents of in- animate nouns of classes 5, 7 and 9 is by placing modifiers after the noun, such as the demonstrative {*a-*} 'this'. This strategy is illustrated with the demonstratives *a-li* and *a-shi* in (15).

(15)   Shimakonde

a. ngw-on-ile   lijanga a-li.
   1SG.SM-see-PST 5.stone this-5

   'I have seen this stone (a definite one).'

b. nangu ni-ndy-on-a   shijiko a-shi.
   I   1SG.SM-PST-see-VF 7.spoon this-7

   'I saw this spoon (a definite one).'

Based on these data, one may postulate that the differential object marking in Shimakonde is mainly regulated by a more complex set of criteria than in Emakhuwa. In Shimakonde the agreement prefix is obligatory if it refers to animate nouns of classes 1 and 2, whereas when it cross-references nouns of classes 5, 7 and 9, the object is both animate and definite. The Shimakonde data require us to postulate a slightly different generalization from the one proposed by Liphola (2001: 23), who says "verbs in Shimakonde only take OPs (object prefixes) of classes 1 and 2, which have a feature [+ANIMATE]". Instead, Liphola's analysis must be refined to state that the Shimakonde object prefix agrees with DP objects of class 1 and 2 and may agree with DP objects of classes 5, 7 and 9 that carry the semantic features [+ANIMATE, +HUMAN, +DEFINITE], as shown in Table 2.

4 DOM in Rhonga and Changana

Changana and Rhonga differ from Emakhuwa and Shimakonde in that it is not animacy, or animacy-plus-definiteness, or a specific noun class, but only definiteness that regulates the appearance of object marking on the verb. We will argue that the following grammatical constraint is operative in Changana and Rhonga, at least in simple transitive constructions:

(16) Changana and Rhonga DOM: Object agreement in the verb is possible but structurally optional if the object is definite/specific. It is forbidden if the object is indefinite/non-specific.

As a preliminary illustration of first part of this generalization, consider (17) below, which was taken from a narration of the pear story film (Chafe 1980).[3] In this narrative, the noun phrase *mufana* 'boy', which is introduced as the head of the relative clause in (17a), is cross-referenced as object on the verbs in (17c, e) by means of the class 1 object prefix {*mu-*}. Notice that the main function of the object agreement prefix in this discourse section is to keep the reader updated about the main topic of the discourse. In contrast to (17c) and (17e), no object marker occurs on the transitive verbs in (17b) and (17d), even though 'his hat' (17b) and 'that hat' (17d) are presumably definite/specific.

(17) Changana

    a. a   lew (luwayi) wa mufana luwiya a-nga-dib-a   ni   basek-e(ni).
       PART that   1   1.boy   who 1SG.SM-REL-fall-FV from bicycle-LOC
       'That boy has fallen from the bicycle'

    b. a-a-khohlw-e   xigqoko xa-kwe.
       1-T/A-forget-FV 7.hat   7-his
       'and (then) (he) forgot his hat (on the ground).'

    c. se... svi-**mu**$_i$-vitan-a.
       then... 8.SM-1.OM-call-FV
       'Then they called him'

    d. a ku va a-tek-a   xigqoko lexiya.
       for   1SM-pick-FV 7.hat   that
       'in order for him to pick up that hat.'

    e. swi-**mu**$_i$-nyik-et-a.
       8.SM-1.OM-hand-CONT-FV
       '(then) they handed him back (his hat).'

In order to demonstrate that the agreement prefix occurrence is really dependent on the definiteness of the object, we investigate in the next subsections how DOM operates in simple transitive and double object constructions. Another purpose is to present evidence in favor of the hypothesis that object markers are purely agreement morphemes.

4.1 Simple object constructions

In Rhonga and Changana, object agreement can occur with definite DPs, regardless of whether the object is animate/human or not. Compare the (a) and (b) examples in (18) through (22).

---

[3] The six-minute film made at the University of California at Berkeley in 1975, and described in Chafe (1980), can be downloaded at http://www.linguistics.ucsb.edu/faculty/chafe/pearfilm.htm.

(18) Rhonga (human object)

    a. mufana a-von-ile   wansati.
       1.boy   1.SM-see-PST 1.woman

      'The boy has seen a woman.'

    b. mufana a-**mu**-von-ile   **wansati**.
       1.boy   1.SM-1.OM-see-PST 1.woman

      'The boy has seen the woman (a specific and definite one).'

(19) Rhonga (animate object)

    a. mufana a-von-ile   xipixi.
       1.boy   1.SM-see-PST 7.cat

      'The boy has seen a cat.'

    b. mufana a-**xi**-von-ile   **xipixi**.
       1.boy   1.SM-7.OM-see-PST 7.cat

      'The boy has seen the cat (a specific and definite one).'

(20) Rhonga (inanimate object)

    a. mufana a-von-ile   zambana.
       1.boy   1.SM-see-PST 5.potato

      'The boy has seen a potato.'

    b. mufana a-**dri**-von-ile   **zambana**.
       1.boy   1.SM-5.OM-see-PST 5.potato

      'The boy has seen the potato (a specific and definite one).'

(21) Changana (human object)

    a. ni-tiv-a   mufana.
       1SG.SM-know-FV 1.boy

      'I know a boy.'

    b. na-**mu**-tiv-a   a[4]   **mufana**.
       1SG.SM-1.OM-know-FV PART 1.boy

      'I know the boy (a specific and definite one).'

(22) Changana (inanimate object)

    a. xin'wanani xi-rhandz-a madonsi.
       7.child   7.SM-like-FV 6.candy

      'The child likes candies.'

---

[4] According to Duarte (2011), the particle *a* occurs only in some dialectical varieties of Changana and Rhonga. When it precedes a DP object, its major role is to encode definiteness and to indicate that the referent of that object has already been given in previous discourse.

b. xin'wanani xa-**ma**-rhandz-a **madonsi**.
7.child   7.SM-6.OM-like-FV 6.candy

'The child likes the candies (a specific and definite type of candy).'

The data show that definiteness and specificity play a major role in regulating the occurrence of DOM in Rhonga and Changana simple transitive constructions. Moreover, the Rhonga and Changana data further support certain theoretical assumptions within the generative literature, according to which agreed-with objects are usually interpreted as definite, whereas unagreed-with objects are interpreted as indefinite. According to Baker (2008: 199), "a standard account of this sort of phenomenon since Diesing (1992) is to say that the verb phrase is the domain of existential closure. NPs that remain in the verb phrase are within the domain of existential closure, and get weak/indefinite/nonspecific readings, whereas NPs that escape the verb phrase get strong generic, specific, or (in article-less languages) definite readings." In sum, this theory predicts that, if an object remains in the VP, it is more likely to be interpreted as indefinite/nonspecific. In contrast, an object that is raised out of the VP may obtain a definite and specific reading, as the tree diagram in Figure 1 shows.
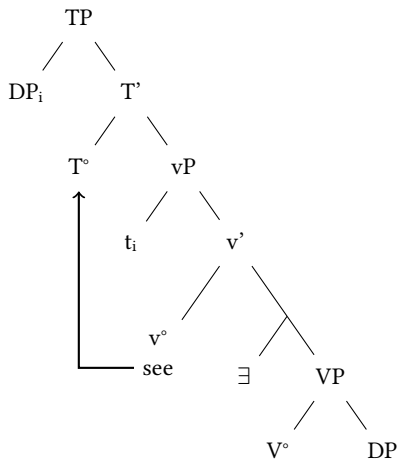
Figure 1: An object in situ allows an indefinite reading

Given these background assumptions, we will assume, hereafter, that Rhonga and Changana set "yes" to Baker's Directionality of Agreement Parameter (23). Notice that "F" can be read as the little v^o that heads the vP projection in Figure 2.

(23)   The Directionality of Agreement Parameter (Baker 2008: 155)
      F agrees with D/NP only if D/NP asymmetrically c-commands F.
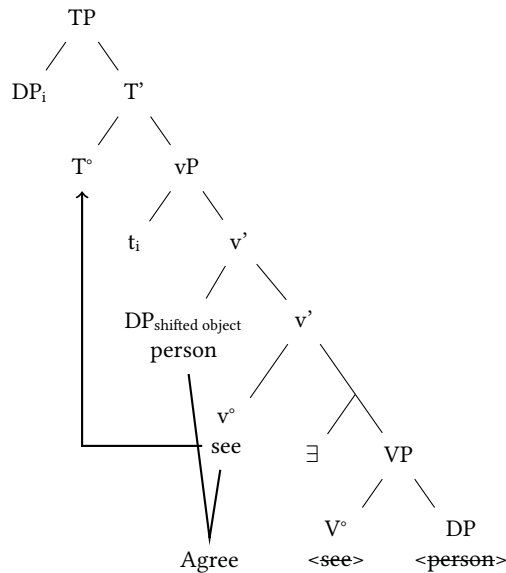
Figure 2: An object raised out of the VP allows a definite reading

This syntactic analysis entails that the head *v* searches upward for the theme object to agree with, not downward. This accounts for why DP objects must move to Spec-*v*P, or even to a higher position such as to a topic position.

4.2 Arguments for the referential agreement analysis

The data also provide evidence that the object prefixes behave more like referential agreement than pronoun incorporation, insofar as the object prefixes can co-occur with and match in person and number with the DP object in Emakhuwa, Shimakonde, Rhonga and Changana. All these languages allow a syntactic adjacency condition between the DP objects and the object prefixes in the transitive verb structure (though the DP need not occur in some discourse conditions, as in (17c) and (17e)), which must be interpreted as indicating that the DP object occurs in an internal argument position. This syntactic adjacency becomes apparent in the fact that DP objects that are referred to by the agreement prefix in the verbal complex are not prosodically separated from the verb, as would be expected if they were dislocated to an adjunct position. This syntactic co-occurrence possibility is taken here as evidence that the object prefixes are not incorporated pronouns, but are instances of referential agreement. This proposal is also in accordance with the assumption that pronominal inflection is clearly absent in languages that have agreement. In incorporated pronominal languages there is no DP inside the vP domain with which this inflection could agree (Jelinek 1989). This is the case in languages like Egyptian Arabic, where pronominal inflections and DP objects are mutually exclusive (Jelinek 1989), as is shown in (24).

(24) Egyptian Arabic

    a. šuft-uh.
       see-3SG.M
       'I saw him.'

    b. šuft il-walad.
       see the-boy
       'I saw the boy.'

Jelinek (1989) posits that the suffix -*uh* does not function as agreement, but as an incorporated pronoun. She assumes that, as an incorporated pronoun occupies an argument position, it receives a theta-role in the usual way that core arguments of the verb do. The same pattern does not emerge in the Bantu languages examined here, since the object and the agreement prefix can co-occur without causing ungrammaticality, as data above show. Another piece of evidence that the Bantu object markers are really agreement in nature comes from the fact that they can only attach to the lower verb in contexts of verbal cluster constructions, as the Rhonga data in (25) indicate:

(25) Rhonga

    a. mufana a-gam-ile   ku-**xi**-b-a   **xipixi**.
       1.boy   1.SM-finish-PST INF-7.OM-hit-FV 7.cat
       'The boy finished hitting the dog.'

    b. *mufana a-**xi**-gam-ile   ku-b-a   **xipixi**.
       1.boy   1.SM-7.OM-finish-PST INF-hit-FV 7.cat
       ('The boy finished hitting the dog.')

The fact that the object marker must occur in the embedded clause in (25) suggests that the lower verb must establish agreement first with the closest DP that is located in the little v^o c-command domain. The fact that the object markers stay close to the lower verb is expected if they are agreement. If they were object clitics or incorporated pronouns, they would be able to attach to higher verbs in the structure. This proposal serves to reinforce the hypothesis that the object prefix is not an incorporated pronoun, but agreement, since there should be no object agreement on v unless the agreed-with NP occurs in an argument position, that is, in the structural position of Spec, vP. Another piece of evidence in favor of this analysis comes from the fact that the object prefix cannot appear on the verb when it is in a passive voice, as in (26):

(26) Changana

    a. ngonyama yi-dlay-iw-ile   hi muhloti.
       9.lion   9.SM-kill-PASS-PST by 1.hunter
       'The lion was killed (by the hunter).'

    b. *ngonyama yi-***yi**-dlay-iw-ile   hi muhloti.
       9.lion   9.SM-9.OM-kill-PASS-PST by 1.hunter
       ('The lion was killed (by the hunter).')

The ungrammaticality of (26b) indicates that object agreement is sensitive to the locality condition and to the fact that the *v*P must be in the active voice, and not in the passive one. In (26a) with a passive *v*P, the DP object is raised from a position internal to the VP to the subject position, thereby blocking the occurrence of the object agreement prefix {*yi-*}.

The next section is devoted to an analysis of DOM in double object constructions in Rhonga and Changana. The objective is to explain why the theme object, even when it is definite and specific, never triggers agreement in contexts where the verb selects two internal objects.

4.3 Double object constructions

The double object constructions examined here all involve applicative constructions. In the literature, a distinction is proposed between *symmetrical* and *asymmetrical* object languages (Bresnan & Moshi 1990; Chimbutane 2002; Ngonyani & Githinji 2006). The two types of languages are diagnosed by syntactic tests involving: (a) object order, (b) passivization, (c) object marking and (d) relativization. In general, in Bantu languages it is observed that the GOAL or applied object (AO) can occur before the direct (THEME or base) object (DO) and vice-versa in symmetrical object languages, while the GOAL or applied object must precede the THEME or base object in asymmetrical object Bantu languages. Only the GOAL/applied object may be passivized in asymmetrical object languages, whereas both can be passivized in symmetrical object languages. Furthermore, either object may trigger the object prefix on the verb stem in symmetrical object languages, as opposed to asymmetrical object languages in which only the goal may be cross-referenced on the verb. Based on these three tests, Bresnan & Moshi (1990) propose the typology in Table 3 to differentiate both types of (Bantu) languages.

Table 3: Bantu asymmetrical and symmetrical object languages

	Symmetrical object languages	Asymmetrical object languages
Object order	(i) AO DO (ii) DO AO	(i) AO DO (ii) *DO AO
Passivization	(i) AO (ii) DO	(i) AO (ii) *DO
Object marking	(i) AO (ii) DO	(i) AO (ii) *DO

In line with the assumptions above, we address two questions about the Rhonga and Changana applicative constructions,[5] as follows: (i) Why does the verb never agree with

---

[5] We will not address the double object construction in Shimakonde and Emakhuwa owing to lack of enough empirical data.

the THEME object, but only with the GOAL/BENEFICIARY object? and (ii) Why is there just one agreement slot in the verbal morphological complex? Our proposal is that Changana and Rhonga are asymmetrical object languages since GOAL arguments carry primary object properties, while THEME arguments do not. This is confirmed by the following grammatical facts: (i) GOALS must precede THEMES in unmarked ditransitive sentences as in (27a); (ii) verb agreement must occur with the GOAL whenever it follows the THEME object, as in (27b); (iii) only GOALS can be passivized, as in (27c). By unmarked ditransitive sentences, we refer to those syntactic contexts in which the GOAL is projected in a higher position than the THEME object, thereby giving rise to the [Subject + Verb + GOAL + THEME] word order.[6] The bolding in (27) indicates the GOAL, be it in an (indirect) object (27a-b) position or in a subject (27c) position. Compare the examples below:

(27) Changana

    a. nahani a-svek-el-a **vapfumba** tihlampfi.
       1.aunt 1.SM-cook-APPL-FV 2.guest 10.fish

      'My aunt is cooking fish for the guests.'

    b. hahani a-**va**-svek-el-a tihlampfi **vapfumba**.
       1.aunt 1.SM-2.OM-cook-APPL-FV 10.fish 2.guest

      'My aunt is cooking fish for them, the guests (and not chicken).'

    c. **vapfumba** va-svek-el-iw-a tihlampfi (hi hahani).
       2.guest 2.SM-cook-APPL-PASS-FV 10.fish by 1.aunt

      'The guests are being cooked some fish (by my aunt).'

In example (27b), the object agreement prefix {*va-*} refers to the GOAL object and emphasizes that it corresponds to given information in the discourse. Therefore the GOAL object must be interpreted as definite and specific in such construction. As the THEME object carries new information in (27b), it must move around the GOAL object to a dedicated focus position in the left periphery of the *v*P. Evidence that the THEME object really represents the contrastive focus comes from the fact that sentence (27b) denotes the semantic interpretation that 'my aunt is cooking fish and not chicken for the guests'. This grammatical pattern is confirmed by the fact that the sentence becomes ungrammatical if the agreement prefix {*va-*} occurs on the verb stem and the GOAL object remains in its unmarked (i.e. immediately postverbal) position, thereby occupying a syntactic position before the THEME object, as in (28):

(28) Changana
    *hahani a-**va**-svek-el-a **vapfumba** tihlampfi.
    1.aunt 1.SM-2.OM-cook-APPL-FV 2.guest 10.fish

    (*'My aunt is cooking fish for them, the guests.')

---

[6] According to Chimbutane (2002: 111), "there is consensus that in such cases the NPs immediately after the verb are mapped onto the thematically higher roles. This suggests, though not without controversy, that the position adjacent to the verb belongs to the arguments ranked thematically higher".

Based on the fact that the GOAL (i) is generated in a high thematic position; (ii) controls object agreement; and (iii) is subject to passivization, one possibility is to posit that the GOAL argument is then introduced in a high syntactic position, while the THEME object is generated in a low syntactic position. Let's then propose that the GOAL object is first merged as a specifier of a high Applicative Phrase (ApplP), which allows it to be base-generated between the AGENT and the THEME objects, as depicted in Figure 3.
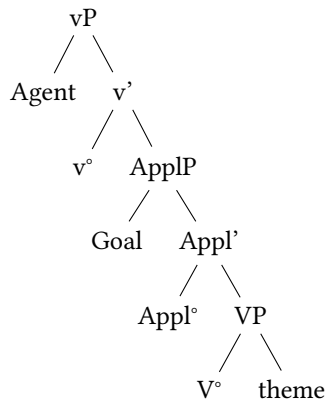
Figure 3: Double object construction structure in Changana

In double object constructions, at least in asymmetrical object languages such as Rhonga and Changana,[7] the THEME object is never cross-referenced on the verb. Let's then assume that the agreement between the GOAL and the little v takes place at the moment when the lexical verb performs successive cyclic movement from V-to-Appl and then from Appl-to-v. Additionally, let's postulate that this head movement operation is followed by the GOAL object shift from the specifier of ApplP to the specifier of vP. Under this view, the agreement between the applied object and the verb follows from the Directionality of Agreement Parameter, proposed in (23), according to which a head F (i.e. the little light verb) agrees with a DP only if that DP asymmetrically c-commands F. This proposal has the advantage of accounting for why only the applied (=GOAL/BENEFICIARY) argument controls the agreement on the verb complex, whereas the THEME argument cannot. In short, this analysis presupposes that both the verb and the applied object are in the same local domain at the moment that the Agree operation takes place. The derivation proposed here entails that the sentence in (29) has the syntactic derivation shown in Figure 4. In this derivation, moving the THEME object around the GOAL to Spec-vP would put it before the GOAL object and the verb, thereby yielding [THEME-GOAL-Verb] word order. A way to restore the superficial [Verb-THEME-GOAL] word order is to propose that the verb subsequently moves to the INFL node (Tense/Aspect/Mood domain) of the clause, as depicted by the derivation in Figure 4.

---

[7] We refer the reader to Chimbutane's (2002) arguments for Changana as an asymmetrical object language.

(29) Changana
       hahani a-va-svek-el-a   tihlampfi vapfumba.
       1.aunt 1.SM-2.OM-cook-APPL-FV 10.fish   2.guest
       'My aunt is cooking fish for them, the guests.'
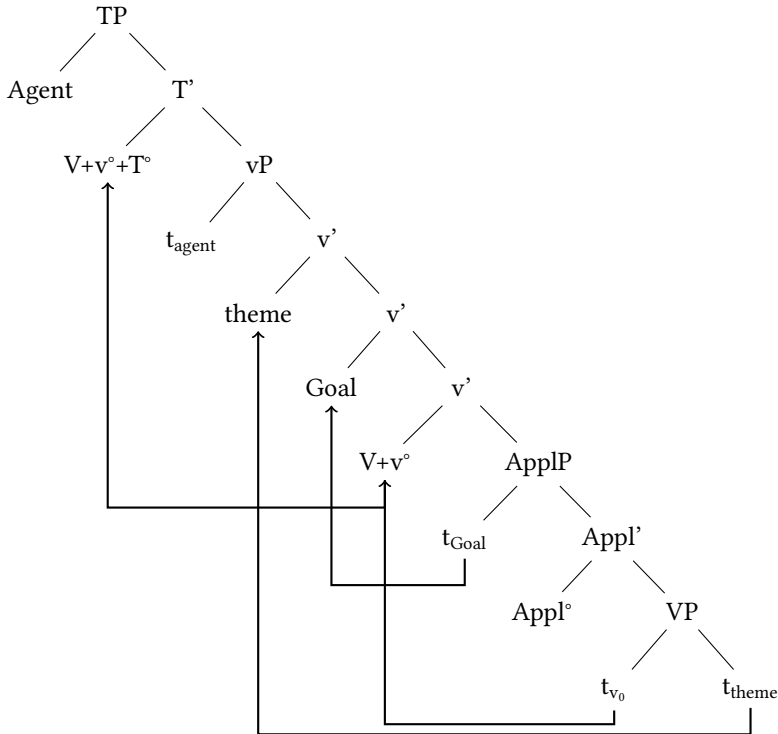
Figure 4: Syntactic derivation of the [SUBJECT+VERB+THEME+GOAL] word order

An immediate consequence of the syntactic derivation outlined here is that it reinforces the proposal advanced in §4.1, according to which the object prefix is not an incorporated pronoun, but simply referential agreement. Evidence comes from the fact that the verb agrees only with the GOAL object, which must be positioned internal to the *v*P domain and not dislocated to a right or left-peripheral position. Additionally, the fact that the DP GOAL occurs inside the vP/ApplP domain corresponds to it not being in an adjunction position, but in an argument position. Recall that in incorporated pronoun languages the DP object and the pronominal inflection cannot co-occur in the same domain, since they are in complementary distribution. In conclusion, as the data above show, such distribution does not occur in the Bantu languages examined here, since the DP object, regardless of whether it is GOAL or THEME, must remain within the domain of the *v*P for the verb to establish agreement with it.

5 Final remarks

In this paper, we have shown that object agreement in Emakhuwa can be interpreted as the realization of a noun class hierarchy, whereas in Shimakonde it is regulated by the animacy and definiteness hierarchy: human > definite animate > inanimate, with definite animate and human nouns controlling object agreement in the verbal complex. Object marking in Changana and Rhonga is regulated not by animacy, but only by definiteness and specificity. Only specific and definite DPs can trigger object agreement in simple transitive predicates.

Regarding double object constructions, the analysis has shown that there is just one slot per clause for object agreement. We have proposed that what regulates the occurrence of agreement in this context is locality, with the GOAL object merging in a higher position than the THEME object in the syntactic structure. This means that GOAL object is the closest candidate for the little *v* head to agree with. This in turn accounts for why agreement with THEME object is systematically forbidden in the double object constructions.

We have also presented empirical evidence in favor of the analysis that the object prefixes behave more like referential agreement than pronoun incorporation, namely: (i) object prefixes match in person and number with the DP object in Shimakonde, Rhonga and Changana; (ii) in these languages, the DP object that triggers the object prefix on the verb stem must appear within the *v*P, which is reflected in the fact that the object is not prosodically separated from the verb. The only exception is Emakhuwa where animacy/humanness overrides class/gender. These facts lead one to assume the DP object sits in an internal specifier position of *v*P. This then reinforces the hypothesis that the THEME object remains in a nuclear position within the v-VP and not in an adjunction position.

Abbreviations

1	1st person	LOC	locative
2	2nd person	M	masculine
3	3rd person	OM	object marker
ACC	accusative	PART	particle
AO	applied object	PASS	passive
APPL	applicative	PL	plural
CONT	contactive	PRF	perfective
DEF	definite	PST	past
DO	direct object	REL	relative marker
OM	double object marking	SG	singular
FV	final vowel	SM	subject marker
INF	infinitive	T/A	tense and aspect marker

References

Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.

Bentley, Mayrene Elizabeth. 1994. *The syntactic effects of animacy in Bantu languages*. Bloomington, IN: Indiana University. (Doctoral dissertation.)

Bresnan, Joan & Sam Mchombo. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63(4). 741–782.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21. 147–185.

Chafe, Wallace L. (ed.). 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (Advances in Discourse Processes III). Norwood, NJ: Ablex.

Chimbutane, Feliciano. 2002. *Grammatical functions in Changana: Types, properties and function alternations*. Canberra: The Australian National University. (Master's dissertation.)

Comrie, Bernard. 1981. *Language universals and linguistic typology*. Oxford: Basil Blackwell.

Croft, William. 1988. Agreement vs. Case marking and direct objects. In Michael Barlow & Charles Ferguson (eds.), *Agreement in natural language: Approaches, theories, descriptions*, 159–180. Chicago: University of Chicago Press.

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

Danon, Gabi. 2002. *Case and formal definiteness: The licensing of definite and indefinite noun phrases in Hebrew*. Telaviv: University of Telaviv. (Doctoral dissertation.)

Diesing, Molly. 1992. *Indefinites*. Cambridge: The MIT Press.

Downing, Laura J. 2014. Differential object marking in Chichewa. (Paper presented at the workshop of The Diachronic Typology of Differential Argument marking, University of Konstanz, April 5-6, 2014.)

Duarte, Fábio Bonfim. 2011. Tense encoding, agreement patterns, definiteness and relativization strategies in Changana. In Eyamba G. Bokamba, Ryan K. Shosted & Bezza Tesfaw Ayalew (eds.), *Selected Proceedings of the 40th Annual Conference on African Linguistics*, 8–94. http://www.lingref.com, document no. 2567. Somerville, MA: Cascadilla Proceedings Project.

Georgopoulos, Carol. 1991. *Syntactic variables: Resumptive pronouns and A' binding in Palauan*. Dordrecht: Kluwer Academic Publishers.

Givón, Talmy. 1976. Topic, pronoun and grammatical agreement. In Charles Li (ed.), *Subject and topic*, 149–188. New York: Academic Press.

Jelinek, Eloise. 1989. The case split and pronominal arguments in Choctaw. In La´szló Maracz & Pieter Muysken (eds.), *Configurationality: The typology of asymmetries*, 117–141. Dordrecht: Foris Publications.

Ngonyani, Deo & Peter Githinji. 2006. The asymmetric nature of Bantu applicative constructions. *Lingua* 116. 31–63.

Riedel, Kristina. 2009. *The syntax of object marking in Sambaa: A comparative Bantu perspective*. Leiden: University of Leiden (Doctoral dissertation.)

Woolford, Ellen. 2000. Object agreement in Palauan: Specificity, humanness, economy, and optimality. In Ileana Paul, Vivianne Phillips & Lisa Travis (eds.), *Formal issues in Austronesian linguistics,* 215–245. Dordrecht: Kluwer Publishers.

Chapter 19

The interaction of two focus marking strategies in Luganda

Jenneke van der Wal
University of Cambridge

Saudah Namyalo
Makerere University

The Bantu language Luganda (Uganda, JE15) has three morphosyntactic strategies to express focus on a nominal referent. This paper examines the detailed interpretational properties of two of these strategies: the preverbal focus construction and augmentless nouns. Based on various focus tests and effects of the combination of the two strategies, we propose that they express identificational focus and exclusive focus, respectively. This detailed and systematic testing refines our understanding of the expression of focus in Luganda and focus types in general.

1 Introduction: Bantu basics and focus

Luganda is one of approximately 500 Bantu languages (Nurse & Philippson 2003: 4), spoken as a native language and lingua franca in Uganda.[1] As is common in Bantu, the canonical word order is SVO, the language has extensive verbal morphology including subject and object indexing, and nouns are divided into noun classes based on their prefixes and their concord forms and agreement patterns in the phrase and clause.[2] In Luganda, noun stems are preceded by a noun class prefix as well as an augment or initial vowel: in (1), the noun *omusota* 'snake' is in class 3 and has the prefix *mu* plus the augment *o*.

(1) O-mu-sóta   gw-áá-kútte   e-mmése.
     3A-3PX-snake 3SM-PAST-catch.PERF 9A-9.rat

     'A/The snake caught a/the rat.'

---

[1] Data come from joint fieldwork by the authors in August 2014 in Mpigi, Masaka and Kampala, unless otherwise indicated.

[2] Noun classes are indicated in glosses by numbers; see further the abbreviations section at the end of the paper.

The augment is important in the expression of focus, which is the subject of this paper. Specifically, we address three research questions: (i) the morphosyntactic strategies used in Luganda to express focus; (ii) the type of focus that these strategies express; and (iii) how different types of focus can be identified in elicitation.

In the field of information structure, many different terms have been used to refer to *focus* and many subtypes of focus have been proposed.[3] When we refer to *focus* in this paper, we take a by now widely accepted semantic definition of focus proposed in Rooth's (1985; 1992; 1996) Alternative Semantics. This states that focus "indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions" (Krifka 2007: 6). A set of alternatives for the object 'rat' in (1) could be, for example, {owl, rabbit, fish,....}. The triggering of a set of alternatives is a unified core function of focus, and the various types of focus can be seen as the outcomes of additional pragmatic and semantic factors (Zimmermann & Onea 2011). Focus can thus be underspecified in its exact interpretation, the only semantically consistent part being the presence of a set of alternatives. The precise interpretation of the focus can differ, however, depending on the context in which it occurs, allowing *completive*, *replacing*, *selecting* etc. interpretations, which differ in their "communicative point" (Dik 1997: 281). Such potential interpretations are not the core point of the current paper, as we first want to establish the semantic type of focus of the Luganda focus strategies.

Semantic types of focus not only trigger a set of alternatives, but also operate on that set of alternatives. This can result in a scalar, exhaustive or exclusive reading, and can have a truth-conditional effect. Exhaustive and exclusive readings are important for the current paper. Exclusive focus means that there is at least *some* referent in the set of alternatives to which the predicate does not apply, e.g. the snake caught a rat but not a fish (and potentially not other prey either but we do not know). Exhaustive focus means that for *all* alternative referents the predicate does not hold, i.e. the predicate is true only for the focused referent, e.g. the snake caught a rat and nothing else. These semantic types of focus can thus be said to be more specific than merely triggering alternatives.

A separate type of focus is associated with a presupposition (for example of existence or exhaustivity). An example is the English *it*-cleft, where the relative clause forms the presupposition and the focus identifies the referent for which the proposition is true. This is what we mean by identificational focus, but it is crucially not necessarily exhaustive, contra to what É. Kiss (1998) claims (see also Onea & Beaver 2011; Byram Washburn 2013; Destruel et al. 2015). For further discussion of semantic and pragmatic types of focus, we refer the reader to Bazalgette (2015) and van der Wal (2016).

The rest of the paper is structured as follows. We briefly introduce the Immediate After Verb (IAV) effect in §2, before discussing the two main focus constructions under study: the Preverbal Focus Construction (PFC, §3) and the absence of the augment (§4). §5 shows the possible interpretations in the combination of the PFC and augmentless nouns. §6 concludes the paper.

---

[3] We do not consider the scope or size of focus in this paper (e.g. Lambrecht's 1994 distinction between argument vs. predicate vs. sentence focus).

2 Immediate After Verb focus position

Like some other Bantu languages (Aghem: Watters 1979; Zulu: Buell 2009; Makhuwa: van der Wal 2006; 2009; Matengo: Yoneda 2011), Luganda has a dedicated linear position for focus, the Immediate After Verb (IAV) position. Although the IAV position is not the primary concern of this paper, it is important to mention its effects here. The focus effect of the IAV position can be seen in restrictions on the placement of inherently focused content question words in (2) and (3), which need to be in IAV position (2b,3b): it is ungrammatical for a content question word like *ani* 'who' or *ki* 'what' to occur in other postverbal positions (2c,3c). When postverbal, focused arguments as well as adverbs must occur in the IAV position.

(2)   a.   A-fúmb-íir-á   á-b-áana   e-m-mére.
          1SM-cook-APPL-FS 2A-2PX-children 9A-9PX-food
          'She cooks food for the children.'

       b.   A-fúmb-ír-á   ání   e-m-mére?
          1SM-cook-APPL-FS who 9A-9PX-food
          'Who does she cook the food for?'

       c.   *A-fúmb-ír-á   é-m-mére   aní?
          1SM-cook-APPL-FS 9A-9PX-food who
          ('Who does she cook the food for?')

(3)   a.   O-mu-sómésa y-a-wá   á-b-áanà   e-m-mére.
          1A-1PX-teacher 1SM-PAST-give 2A-2PX-children 9A-9PX-food
          'The teacher gave the children food.'

       b.   O-mu-sómésa y-a-wá   kí   á-b-áanà?
          1A-1PX-teacher 1SM-PAST-give what 2A-2PX-children
          'What did the teacher give the children?'

       c.   *O-mu-sómésa ya-wá   à-b-áana   kí?
          1A-1PX-teacher 1SM.PAST-give 2A-2PX-children what
          ('What did the teacher give the children?')

Answers to content questions are also found in IAV position, with potential intervening non-focal elements in dislocated position, as in (4d). All examples (4b-d) are felicitous answers to (4a).

(4)   a.   A-fúmbyé   á-tyá e-m-púúta?
          1SM-cook.PERF 1-how 9A-9PX-Nile.perch
          'How has she cooked the Nile perch?'

       b.   Bulúngi.
          'Well.'

   c. A-gi-fúmbyé   búlúngi.
      1sm-9om-cook.perf well

      'She has cooked it well.'

   d. (E-m-púúta)   a-gi-fúmbyé   búlúngi.
      9a-9px-Nile.perch 1sm-9om-cook.perf well

      '(The Nile perch) She has cooked it well.'

As shown by Hyman & Katamba (1993), a Luganda noun without an augment [A] in an affirmative clause is in focus (see §4). If a [A] noun occurs postverbally, it can only appear in IAV, as shown in (5): it is grammatical as the first noun after the verb (5b) but not as the second after an augmented [+A] noun (5c).[4]

(5)   Hyman & Katamba (1993: 228–229)

   a. Y-a-gúl-ir-a   a-b-áana   e-bí-tábó.   [+A, +A]
      1sm-past-buy-appl 2a-2px-children 8a-8px-books

      'He bought the children books.'

   b. Yagúlírá __-báana ebitábó.   [-A, +A]

      'He bought THE CHILDREN books.'

   c. *Yagúlira a-báana __bitábó.   [+A, -A]

      ('He bought the children BOOKS.')

   d. Yagúlírá __-báana __-bitábó.   [-A, -A]

      'He bought THE CHILDREN BOOKS.'

For completeness we note that when both postverbal objects are [-A], as in (5d), the interpretation is VP focus, including the verb and both objects. This is visible in the appropriate situation for (6), which contrasts the action with a whole other action of doing something else, and the inappropriateness of the contrast with just one object rather than the whole action in (7).

(6) Y-á-gúlíddé   __-m-mótoka __-mu-píira.   [-A, -A]
     1sm-past-buy.perf 9px-car   3px-tyre

     'He BOUGHT A TYRE FOR THE CAR.' (i.e. this was the only work he did)

(7) #Y-á-gúl-íddé   __-m-mótoka __-mu-píira, sí   picipici.   [-A, -A]
     1sm-past-buy.perf 9px-car   3px-tyre   neg.cop motorbike

     'HE BOUGHT A TYRE FOR THE CAR, not for the motorcycle.'

---

[4] In examples (5–7) only, the absence of the augment is indicated by an underscore placed before the noun, and small caps in the translation indicate focus. [+A] indicates that the noun has the augment, while [-A] indicates it does not. The first linear [A] corresponds to the first linear noun in the clause and the second linear [A] to the second linear noun in the clause.

The interpretation of focus in IAV is not restricted to one type, but is underspecified: postverbal objects can be interpreted as exclusive focus or part of the comment (see §4), and in locative inversion the subject can occur in IAV too. In subject inversion, either the subject is in focus, as in the question in (8a) and the narrowly focused subject in (8b), or it is just detopicalised as part of a thetic sentence (said "out of the blue"), as in (9):

(8) a. Mu-no mú-súlá-mú   àni?
       18-DEM 18SM-sleep-18LOC who

       'Who sleeps in here?'

    b. Mu-no mú-súlá-mú   mu-lalû.
       18-DEM 18SM-sleep-18LOC 1PX-crazy

       'It's a mad person who sleeps here.'

(9) Ku-kyalo   kw-á-f-íír-á-kó   o-mu-sájja.
    17-7.village 17SM-PAST-die-APPL-FS-17LOC 1A-1PX-man

    'In the village died the man.'

3 The Preverbal Focus Construction

A second focus strategy, the Preverbal Focus Construction (PFC), places the focused referent in the preverbal domain with an agreeing morpheme -*e*[5] preceding the verb. Both arguments and adverbs can be focused in the PFC. In (10a), the object *emmese* 'rat' precedes the subject *omusota* 'snake' and triggers class 9 agreement on *e*. In (10b) the subject *omusota* 'snake' is in focus and triggers class 3 agreement on *e*. In (11), the adverb *eggulo* 'yesterday' appears sentence-initially and *e* agrees with it in class 11.

(10) a. E-m-mése o-mu-sóta   gy-e gw-a-kuttê.
        9A-9PX-rat 3A-3PX-snake 9-e 3SM-PAST-catch.PRF

        'It's a/the rat that the snake caught.'

     b. O-mu-sóta   gw-e gw-á-lya   e-m-mése.
        3A-3PX-snake 3-e 3SM-PAST-eat 9A-9PX-rat

        'It's a/the snake that ate the rat.'

(11) E-ggulo   lw-e y-a-limyé.
     11A-yesterday 11-e 1SM-PAST-dig.PERF

     'It's yesterday that he dug.'

Leaving aside questions concerning the syntactic structure of the PFC,[6] the main concern in this paper is what type of focus the PFC expresses. Our hypothesis is that the

---

[5] In the examples, we gloss the morpheme -*e* as 'e', since it is as yet unclear whether it is a focus marker or a relative marker.

[6] A specific question is how far the PFC has developed from a biclausal cleft to a monoclausal focus construction, since the PFC shows characteristics of both.

PFC expresses identificational focus. Identification is here taken to mean that there is a presupposition that a referent exists for which the proposition holds true, and that the speaker asserts that this referent is identified as corresponding to the focused element (possibly from among various alternatives). For example, in (10b) there is a presupposition that someone ate the rat, and the one who is identified as that someone is the snake. In what follows we test this hypothesised meaning of the Luganda PFC in four linguistic environments: content questions, indefinites, incomplete yes/no questions and with focus particles (see van der Wal 2016 for an overview of focus tests that we draw on).

3.1 Content questions

The identification reading is evident in the interpretation of content questions, such as 'who did you hit?' in (12) and (13). If there is a presupposition that somebody was hit, then the prediction is that the answer is drawn from a non-empty set, and hence that answering 'nobody' is not felicitous. This is true for the question formed with the PFC as in (12), confirming our hypothesis that the PFC contains a presupposition and expresses identificational focus.

(12) Aní gw-e w-á-kúbyê?
      who 1-e   2SG.SM-PAST-hit.PERF

      'Who is it that you hit?'

      #Sí-rí-nâ.   / Te-wá-lî.
      NEG.1SG.SM-be-with NEG-16SM-be

      'Nobody.'

By comparison, an object content question in SVO order (13) can felicitously be answered by 'nobody', since there is no presupposition that anyone in fact was hit.

(13) W-á-kúbyé   ání?
      2SG.SM-PAST-hit.PERF who

      'Who did you hit?'

      Sí-rí-nâ.   / Te-wá-lî.
      NEG.1SG.SM-be-with NEG-16SM-be

      'Nobody.'

3.2 Indefinites

If the PFC contains a presupposition that some referent was involved in the event expressed by the verb, then the identification of that referent should add information and be specific. For example, it is odd to say 'who he hit was someone or other', because we already know from the presupposition contained in 'who he hit' that someone was hit. A non-specific interpretation is thus predicted to be infelicitous for the focused referent in the PFC. This is easily tested with the word *omuntu* 'person', which can receive an indefinite, non-specific interpretation ('anyone') in SVO order, as in (14a) and (15a), but

as predicted not in the PFC (14b, 15b). Instead, *omuntu* must be interpreted as a specific entity or the type named by the noun (as opposed to an entity of some other type) in (14b) and (15b).

(14) a. Y-á-kúbyé   o-mu-ntu.
          1SM-PAST-hit.PERF 1A-1PX-person

          'He hit someone.'

      b. O-mu-ntu   gw-e y-a-kubyê.
          1A-1PX-person 1-e   1SM-PAST-hit.PERF

          'It's a person he hit.'

          #'He hit someone.' / #'It is someone (non-specific) that he hit.'

(15) a. O-mu-ntu   a-gudde.
          1A-1PX-person 1SM-fall.PERF

          '(A certain) Someone fell.'

      b. O-mu-ntu   y' aa-gudde.
          1A-1PX-person 1.e 1SM-fall.PERF

          'It's a human being who fell.'

          #'Someone fell.' / #'It is someone (non-specific) that fell.'

3.3 Incomplete 'yes/no' questions

A further test involves incomplete 'yes/no' questions. These are questions that ask about a subset of the referents for which the sentence is true. Concretely, the test involves showing participants a picture of Thomas with a cat and a rabbit, and asking a 'yes/no' question on a subset of 'cat and rabbit', namely whether Thomas has a cat. Incomplete 'yes/no' questions and their answers can reveal focus effects in two ways: the choice of the answering particle ('yes' or 'no') and the choice of conjunction ('and' or 'but').[7]

With regard to the answering particle, only the answer 'yes' is predicted to appear in an answer to an SVO question, since it is true that Thomas has a cat, regardless of whether he has other animals too. By contrast, the answer to the corresponding question in the PFC can relate not just to the content, but also to the unique identification: the answer 'no' here negates not that Thomas has a cat, but that Thomas has only a cat. The answer 'no' is thus predicted to be possible for the PFC question but not for the SVO question. This is precisely what we found: an incomplete 'yes/no' question in the PFC (16a) is corrected by answering 'no' (16b), whereas for an SVO question (17a), the answer is either a simple 'yes' (17b), or a 'yes but/and' (17c), but not a straight 'no'.

---

[7] This test is modified from the Questionnaire on Information Structure (Skopeteas et al. 2006); see Onea & Beaver (2011) for an experimental application of this test for Hungarian, and Destruel et al. (2015) for further crosslinguistic testing of it-clefts.

(16) PFC question (picture is shown of Thomas having a cat and a rabbit)

    a. Kkápa Thomas   gy-e a-li-nâ?
        9.cat   1.Thomas 9-e   1sm-be-with

        'Is it a cat that Thomas has?'

    b. Nédda, sí   y-okkâ, Thomas a-li-ná   n' á-ká-myû.
        no   neg.cop 9-only   1.Thomas 1sm-be-with and 12a-12px-rabbit

        'No, not only, Thomas also has a rabbit.'

(17) SVO question (picture is shown of Thomas having a cat and a rabbit)

    a. Thomas   a-li-ná   kkápâ?
        1.Thomas 1sm-be-with 9.cat

        'Does Thomas have a cat?'

    b. Yee, Thomas   a-li-ná   kkápa.
        yes 1.Thomas 1sm-be-with 9.cat

        'Yes, Thomas has a cat.'

    c. Yee, Thomas   a-li-ná   kkápa, nayé a-li-ná   ne á-ká-myû.
        yes 1.Thomas 1sm-be-with 9.cat   but 1sm-be-with and 12a-12px-rabbit

        'Yes, Thomas has a cat, but he has a rabbit too.'

When the answer to the PFC question is not a straight negation as in (16b), the prediction is that the PFC strongly prefers 'yes but' over 'yes and'. This is because the identificational focus associated with the PFC suggests an exclusion of and contrast to alternative referents, whereas the 'yes and' answer is explicitly non-exclusive and non-contrastive. The conjunction 'but' is required here to highlight the identification associated with the PFC and to express the contrast with possible alternatives, as is confirmed by the data in (18). In (18a) the yes/no question is formed by a PFC, and the answer 'yes but' in (18b) is acceptable, but the answer 'yes and' in (18c) is not. The explicit contrast induced by 'but' is not necessary for the SVO counterpart in (19), since no exclusion or contrast is expected for SVO: both answers 'yes but' and 'yes and' are fine (19b-c).

(18) PFC question (picture is shown of Maria with a red and a yellow pepper)

    a. Kikáámuláli e-ki-myúfu Maria   ky-e a-li-ná?
        7.pepper   7a-7px-red 1.Maria 7-e   1sm-be-with

        'Is it a red pepper that Maria has?'

    b. Ye, kyokka/naye a-li-ná   ne e-ky-á   ky-énvu.
        yes only/but   1sm-be-with also 7-7-conn 7-yellow

        'Yes, only/but she also has a yellow one.'

    c. #Era/ate   a-li-ná   ne e-ky-á   ky-énvu.
        and/and.even 1sm-be-with also 7-7-conn 7-yellow

        'And she also has a yellow one.'

(19) SVO question (picture is shown of Maria with a red and a yellow pepper)

    a. Maria a-li-ná   kikáámuláli e-ki-myúfu?
       1.Maria 1SM-be-with 7.pepper   7A-7PX-red

       'Does Maria have a red pepper?'

    b. Ye, era/ate   a-li-ná   ne e-ky-á   ky-énvu.
       yes and/and.even 1SM-be-with also 7-7-CONN 7-yellow

       'Yes, and she has also/even a yellow one.'

The first part of this test ('yes/no') suggests that the PFC is associated not just with identification but with exhaustivity, that is, the exclusion and negation of all possible alternatives to the set expressed by the focused phrase ('only a cat and nothing else'). This is because a 'no' answer negates that Thomas *exhaustively* has a cat. A question is whether this is an inherent semantic part of the PFC or rather a pragmatically implied part of meaning. A suggestive answer is found in the second test ('and/but'). Specifically, if the PFC has an inherently exhaustive interpretation, we would predict 'no' to be the only appropriate answer. Instead, the fact that 'yes, but' is also compatible with the PFC suggests that the exhaustive interpretation is implied but not encoded. More specifically, it seems that there is a presupposition of contrast and an implicature of exhaustivity (cf. Byram Washburn 2013). This is why there is a need for a contrasting conjunction 'but', but there is no truth-conditional exhaustive meaning inherently present in the PFC.

3.4 Focus particles 'only' and 'even'

Focus particles also reveal the identificational nature of the PFC. The particle 'only' shows a subset of identification, namely exhaustive identification. 'Only' is compatible with the PFC for objects, as in (20), and the PFC is in fact required for subjects modified by 'only', as shown by the ungrammaticality of (21b) where the subject 'only the house' is not in the PFC. Note that 'only' requires the augmentless form of the noun, as further discussed in §4.

(20) Ka-wûnga k-okká k-e y-a-lyâ.
      12PX-ugali 12-only 12-e 1SM-PAST-eat

      'It is only ugali that he ate.'[8]

(21)   a. Nnyúmba y-okká y-e y-a-yér-ebw-a.
         9.house   9-only 9-e 9SM-PAST-sweep-PASS-FS

         'Only the house was swept.'

    b. *Nnyúmba y-okká y-a-yér-ebw-a.
         9.house   9-only 9SM-PAST-sweep-PASS-FS

         ('Only the house was swept.')

---

[8] *Ugali*, a stiff porridge made from maize flour, is the staple food in East Africa.

The particle *ne* 'even, also', however, is not compatible with the PFC, as shown in (22b) and (23b). For comparison, it *is* grammatical in a preposed topic position (22a).

(22) a. Ne   á-ká-wûngá   y-a-ká-lya.
          even 12A-12PX-ugali 1SM-PAST-12OM-eat
          'He ate even ugali.'

  b. *Ne akawûnga   k-e y-a-lyâ.
      even 12A-12PX-ugali 12-e 1SM-PAST-eat
      ('It's even ugali that he ate.')

(23) a. Né   énnyúmba y-a-yér-ebw-a.
          even 9.house   9SM-PAST-sweep-PASS-FS
          'Even the house was swept.'

  b. *Né   énnyúmba y-e y-a-yér-ebw-a.
      even 9.house   9-e 9SM-PAST-sweep-PASS-FS
      ('It is even the house that was swept.')

The opposite distribution of 'only' and 'even' in the PFC can be explained by the incompatibility between 'even' and exclusivity: 'even' includes all alternatives up to the furthest end of a scale of possibilities, thus excluding none of the alternatives. This suggests that the PFC is associated with exclusive focus, not just identification. This is also how Walusimbi (1996: 52) characterizes the PFC: as expressing the existence of a presupposition and simultaneously an exclusive set that satisfies a missing variable in the presupposition. Alternatively, however, exclusivity may not be inherent to the PFC (as suggested in §3.3 above), but instead be due to the pragmatic incompatibility between scalarity and identification of a referent. To illustrate, if one would want to identify 'who ate ugali?' the scalarity-related answer [??]'Even John ate ugali' does not seem appropriate.

In conclusion, the incompatibility of the PFC with the scalar particle 'even' is not directly expected if the PFC expresses just identificational (not inherently exclusive) focus. Further tests exploring the combination of the PFC with augmentless nouns can further elucidate our hypothesis that the PFC expresses identificational focus. We turn to these in §4 and §5.

4 The augment

Apart from the IAV position and the PFC, a third linguistic strategy related to focus is the presence or absence of the augment on nouns. Hyman & Katamba (1990) and Hyman & Katamba (1993) analyse the presence ([+A]: *o-mu-sota*) or absence ([-A]: __-*mu-sota*) of the augment as related to two licensers: negation and focus. They show that any NP following a negative verb must be [-A], as shown in (24a) and the ungrammaticality of (24b-d); this is a purely syntactic restriction and has no influence on information structure.

(24) Hyman & Katamba (1993: 224)

    a. Te-bá-á-w-a   __-b-áana   __-bi-tábó.
       ɴᴇɢ-2ѕᴍ-ᴘᴀѕᴛ-give-ꜰѕ 2ᴘх-children 8ᴘх-books

       'They didn't give the children books.'

    b. *Tebááwa abáana ebitábó.

    c. *Tebááwa abáana __-bi-tábó.

    d. *Tebááwa __-báana ebitábó.

They also identify focus as a licenser of augmentless nouns: an augmentless object in an affirmative clause receives a focus reading, as shown in the difference between (25a) and (25b).

(25) Hyman & Katamba (1993: 228)

    a. Y-a-gúla   e-bí-tábó.
       1ѕᴍ-ᴘᴀѕᴛ-buy 8ᴀ-8ᴘх-books

       'He bought books.'

    b. Y-a-gúla   __-bi-tábó.
       1ѕᴍ-ᴘᴀѕᴛ-buy 8ᴘх-books

       'He bought ʙᴏᴏᴋѕ.'

We generally accept their analysis of the augment in Luganda, and extend it by studying what the precise interpretation of this focus strategy is. Our proposal is that the focus expressed by the absence of the augment is exclusive. As explained before in §1, by *exclusive* we mean that the focused referent triggers a set of alternatives and excludes at least some of those alternatives. For example, in a sentence 'it's matooke that she ate',[9] the focused item *matooke* triggers the existence of a set of alternative things that she may have eaten, like rice, beans, porridge, and conveys that the action of eating is true for *matooke*, but not true for some (and potentially all) of the alternatives. This exclusive focus can be shown in five tests: alternative questions, correction, indefinites, focus particles, and answers to content questions.

4.1 Alternative questions

In answers to alternative questions, a selection needs to be made between alternatives. This means that one or more alternatives are excluded. The data show that in this case the [-A] form is obligatorily used in both the question (26a) and the answer (26c), which can also just be a one word fragment (26d). Example (26b) shows that nouns with the augment in object position lead to ungrammaticality of the question.

(26)   a. O-yágálá   mágí obá nnyamá?
        2ѕɢ.ѕᴍ-want 6.eggs or 9.meat

        'Do you want eggs or meat?'

---

[9] *Matooke* is a green plantain which is traditionally steamed in banana leaves and mashed for eating.

    b. *O-yágálá   a-mágí   obá e-nnyama?
       2SG.SM-want 6A-6.eggs or   9A-9.meat

       ('Do you want eggs or meat?')

    c. N-jágálá   nnyámá / #é-nnyámá.
       1SG.SM-want 9.meat   / 9A-9.meat

       'It's meat that I want', 'I want MEAT.'

    d. Nnyámá.
       9.meat

       'Meat.'

4.2 Overt contrast/correction

A second test for exclusivity involves adding an overt contrast or correction to a statement. The overt contrast or correction phrase excludes the mentioned alternative(s), and thus requires the [-A] form of the object *nnyúmba* 'house' in the correction in (27), and the object *ssúúká* 'bedsheets' in the contrast in (28).[10]

(27)   a. Y-a-yéra   o-lú-ggya.
         1SM-PAST-sweep 11A-11PX-yard

         'She swept the yard.'

      b. Nédda, y-a-yérá   nnyúmba / #e-nnyúmba.
         no   1SM-PAST-sweep 9.house

         'No, she swept the house.'

(28)   Máaka a-b-áana   y-a-bá-gúl-ír-á   ssúúká / #e-ssúúká, sí
       1.Mark 2A-2PX-children 1SM-PAST-2OM-buy-APPL-FS 10.bedsheets   NEG.COP
       bulangíti.
       blankets

       'Mark bought the children bedsheets, not blankets.'

Crucially, the [+A] form is judged as felicitous if the contrasting clause is absent, as in (29).

(29)   Máaka a-b-áaná   y-a-bá-gúl-ír-a   e-ssúúkâ.
       1.Mark 2A-2PX-children 1SM-PAST-2OM-buy-APPL-FS 10A-10.bedsheets

       'Mark bought the children bedsheets.'

---

[10] This particular example has a preverbal topical benefactive object (resumed by an object marker on the verb), which allows the theme object to be contrasted in IAV position.

4.3 Indefinites

A third test showing exclusivity involves the indefinite noun *(o)muntu* 'person'. This can be interpreted as non-specific in its [+A] form, as in (30a); but refers to a type of entity when in its [-A] form and when the scope of focus is the object, as in (30b). This is because non-specific indefinites do not allow for the exclusion of alternatives ('anyone' includes everyone), whereas types do allow for exclusion, as indicated in the following contrastive clause ('not a cat').

(30)   a.   N-á-kúbyé   o-mu-ntu.
            1SG.SM-PAST-hit.PERF 1A-1PX.person

            'I beat someone.'

       b.   N-á-kúbyé   mu-ntú,   si   kkapa.
            1SG.SM-PAST-hit.PERF 1PX-person NEG.COP 9.cat

            'I beat a person, not a cat.'

4.4 Focus particles 'only' and 'even'

The exhaustive focus particle *-okka* 'only' requires the noun it modifies to be [-A], which is expected considering that an exhaustive reading excludes all alternative referents.

(31)   Y-a-yérá   nnyúmba / *e-nnyúmba y-okká, sí   ki-yûngu.
       1SM-PAST-sweep 9.house   / 9A-9.house   9-only NEG.COP 7PX-kitchen

       'She swept only the house, not the kitchen.'

Equally expected is the finding that the focus particle 'even/also' cannot modify a [A] noun, as shown in (32).[11]

(32)   a.   Y-a-lyá   n'   á-ká-wûnga.
            1SM-PAST-eat even 12A-12PX-ugali

            'He ate even ugali.'

       b.   *Y-a-lyá   né   ká-wûnga.
            1SM-PAST-eat even 12PX-ugali

            ('He ate even ugali.')

---

[11] Interestingly, negation (which normally licenses a [-A] object) does not affect this constraint:

(i)   a.   Te-y-a-lyá   n'   ákáwûnga.
            NEG-1SM-PAST-eat even 12A-12PX-ugali

            'He didn't even eat ugali.' (though it's his favourite dish)

       b.   *Te-y-a-lyá   né   káwûnga.
            NEG-1SM-PAST-eat even 12PX-ugali

            ('He didn't even eat ugali.' (though it's his favourite dish))

If [-A] indeed encodes exclusivity, the ungrammatical status of (32b) is explained by the fact that né 'even, also, with' includes all the less likely alternatives (not excluding any alternatives), and thus is incompatible with the exclusive meaning of the absence of the augment on the noun. This ungrammaticality would not be expected if the absence of the augment merely encodes focus (cf. Hyman & Katamba 1993), since 'even ugali' is indeed the new information focus in this clause.

4.5 Restricted interpretation of answers to content questions

A fifth test involves the interpretation of answers to content questions. Either form of the object [+A/- A] is grammatical here, but whereas the [+A] answer is described as "neutral" (33b), a contrast is spontaneously indicated by our consultants for the [A] form (33c).

(33)   a. O-mu-nyá   gu-kutte   ki?
          3A-3PX-gecko 3SM-catch.PERF what

          'What has the gecko caught?'

       b. Gú-kutte   e-nsirî.
          3SM-catch.PERF 9A-9mosquito

          'It caught a mosquito.'

       c. Gú-kútté   nsírî.
          3SM-catch.PERF 9mosquito

          'It caught a MOSQUITO.'
          Situation: "not something else"

4.6 A note on the [–A] subject

Just as a subject modified by 'only' is required to appear in the PFC (see §3.4), the subject cannot be [-A] in its canonical preverbal position (34b), that is, the PFC is a precondition for [-A] subjects (34a). For comparison, the [+A] subject is grammatical in the PFC (34c) as well as in SVO order (34d).

(34)   a. Mu-sómésa y-e y-a-mú-laba.
          1PX-teacher 1-e 1SM-PAST-1OM-see

          'It's the teacher that saw him.'

       b. *Mu-sómésa y-a-mú-laba.
          1PX-teacher 1SM-PAST-1OM-see

          ('It's the teacher that saw him.')

       c. O-mu-sómésa y-e y-a-mú-laba.
          1A-1PX-teacher 1-e 1SM-PAST-1OM-see

          'The teacher saw him.'

d. O-mu-sómésa   y-a-mú-laba.
   1A-1PX-teacher 1SM-PAST-1OM-see
   'The teacher saw him.'

The logical subject can, however, be [-A] in locative inversion. Locative inversion is a construction in which the logical subject appears in a linearly postverbal position while the preverbal position is filled by a locative expression like *muno* 'here' in (35) or *mukatale* 'in the market' in (36) and (37). The function of this inversion is to detopicalise the subject, which results in a thetic interpretation when the subject is [+A] (see example 9 above), and an exclusive interpretation when it is [-A]. This exclusive interpretation can be seen in the contrast in (35), the modification by 'only' in (36), and the generic or type interpretation of the indefinite in (37).

(35) Mu-no mú-súlá-mú   mu-lalû.
     18-DEM 18SM-sleep-18LOC 1PX-crazy

     'It's a mad person who sleeps here.' (nobody else)

(36) Mu-katále mw-áá-báddé-mú   báána   b-okká.
     18-market 18SM-PAST-be.PERF-LOC18 2.children 2-only

     'In the market were only children.'

(37) Mu-katále mw-áá-báddé-mú   baantú, si   mbwa.
     18-market 18SM-PAST-be.PERF-LOC18 2.people NEG.COP 10.dogs

     'In the market were people, not dogs.'

For completeness we mention that the agent NP in a passive can be [-A] as well, and as expected it triggers the exclusive interpretation (38). The agent does not need a preposition 'by' in Luganda (see Pak 2008).

(38) E-m-mése y-a-l-ííbw-á   mú-sótâ.
     9A-9PX-rat 9SM-PAST-eat-PASS-FS 3PX-snake

     'The rat was eaten by a snake.' (implying 'not by a cat')

The syntactic licensing of augmentless nouns seems to be restricted to a vP-internal position (also depending on the analysis of the PFC as a cleft or not); cf. Halpert (2012; 2013) for this proposal for Zulu [-A] nouns, and Carstens & Mletshe (2015; 2016) on Xhosa. The syntactic structure will be left to one side here, since our main aim is to unravel the precise interpretation of focus strategies.

In conclusion, the absence of the augment encodes a focus interpretation that is more specific than Hyman & Katamba's (1993) underspecified focus, which the five tests in this section diagnose as exclusive focus. The next question is what happens when the identificational focus of the PFC and the exclusive focus of the augmentless nouns are combined.

5 Combination of PFC and [-A]

Identification and exclusion are very close to one another but are not identical, and the two focus marking strategies encoding them in Luganda can be combined. On the basis of our analysis in the previous sections, we predict the following combinations, as also represented in Table 1 below:

- a postverbal [-A] object triggers an exclusive focus reading, excluding at least some of the alternatives;
- a preverbal [+A] subject or object in the PFC is identified as the referent for which the otherwise presupposed predication is true;
- a preverbal [-A] subject or object in the PFC marks the predicate as part of the presupposition AND excludes alternatives for the focused referent.

Table 1: Predicted combinations of two focus strategies for Luganda objects

	SVO	PFC
[-A]	exclusive	exclusive identification
[+A]	neutral/new info	identification

We demonstrate the various combinations of the focus strategies and their interpretations in three example cases.

First, we show the difference in interpretation between the [+A] and [-A] form of the subject in the PFC. As an answer to the subject question in (39), the PFC will be used. Since it concerns subject focus, the PFC *must* be used,[12] and this allows us to test the influence of [-A]. Since the only difference is the presence or absence of the augment, the difference in interpretation shows that exclusivity is *only* linked to [A], not to the PFC. In both answers, there is a presupposition that there exists someone who is reading. The difference is found in the explanation that consultants provide for felicitous use of each example. When the answer has a [+A] noun in subject position (39a), this simply identifies the one reading as a/the child, whereas when the answer has a [A] noun in subject position (39b), this implies that there are others in the room and the child is identified to the exclusion of these alternatives.

(39) Sketched situation: I want to know who is reading in the room next door.

Q: 'Who is reading?'

a. O-mw-ána   y-e a-somâ.
    1A-1PX-child 1-e 1SM-read

'It's a/the child that is reading.'
Situation: There is only one person in the room.

---

[12] Locative inversion would be an alternative for subject focus, but only if there is indeed a topical locative. The point here is that the subject cannot occur in its canonical preverbal position if it is focused.

b. Mw-ána   y-e a-somâ.
1px-child 1-e 1sm-read

'It's the child that is reading.'
Situation: There are other people in the room.

Second, the difference in interpretation between the two focus strategies is also detected in a focus test that É. Kiss (2009) applies for Hungarian. Numerals are normally interpreted as a lower boundary 'at least this amount' (Horn 1972; Levinson 2000), but in exclusive focus they refer to precisely that quantity. This is because alternative (higher) values are excluded. If the PFC is identificational but not inherently exclusive or exhaustive, we would not predict the PFC to result in the 'exact amount' reading. Indeed, the lower-bound reading of 'at least one million' is found in post-verbal position with a [+A] object, as in (40a), and crucially does not change in the PFC, as shown in (41a). When the object takes its [-A] form, the meaning narrows down to the precise the value given in the focused constituent: exactly one million. This is the case regardless of whether the object is in-situ after the verb (40b), or in the PFC (41b).

(40)   a. SVO [+A]
Tw-étaaga a-ka-kadde   ka-mú o-ku-mála   é-nnyúmba.
1pl.sm-need 12a-12px-million 12-one 15a-15px-finish 9a-9.house

'We need (at least) one million to finish the house.'

b. SVO [-A]
Tw-é-taaga ka-kadde   ka-mú o-ku-mála   é-nnyúmba.
1pl.sm-need 12px-million 12-one 15a-15px-finish 9a-9.house

'We need (exactly) one million to finish the house.' (we've calculated)

(41)   a. PFC [+A]
A-ka-kadde   ka-mú k-e tw-etaagá o-ku-mála   é-nnyúmba.
12a-12px-million 12-one 12-e 1pl.sm-need 15a-15px-finish 9a-9.house

'(About) one million we need to finish the house.'

b. PFC [-A]
Ka-kadde   ka-mú k-e tw-etaagá o-ku-mála   é-nnyúmba.
12px-million 12-one 12-e 1pl.sm-need 15a-15px-finish 9a-9.house

'(Exactly) one million we need to finish the house.'

These interpretations can be accounted for if the PFC merely presupposes that a certain amount of money is needed to finish the house, which is then identified to be one million (at least), while the absence of the augment encodes exclusivity, meaning that alternatives to the amount given are excluded, resulting in an exact reading.

A final example shows the same interpretations in the felicitous situations sketched for the different [+A/-A] forms provided by our consultants. The [-A] variant of the preverbal focused subject is interpreted as corrective or contrastive, i.e. excluding alter-

natives, as indicated for (42a). This is not the case for the [+A] variant that serves as an identificational answer in (42b).[13]

(42)   a. Aní   á-sula   mu nnyúmba eyo?
            1.who REL1SM-sleep 18   9.house   9.DEM

       'Who sleeps in this house?'

       b. Mu-wála ye a-súlá-mû.
            1PX-girl   1.e 1SM-sleep-18LOC

       'A girl sleeps there.'
       Situation 1: There are a boy and a girl. Which one sleeps here? (contrastive)
       Situation 2: You expect a man to be sleeping there. (corrective)

       c. O-mu-wála ye a-súlá-mû.
            1A-1PX-girl   1.e 1SM-sleep-18LOC

       'The girl sleeps there.'
       Situation: You know the girl.

Testing the combination of the PFC and augmentless nouns reinforces our analysis of the precise focus interpretation of these strategies: the PFC triggers a presupposition and identificational focus, whereas [-A] has an exclusive focus effect.

6 Conclusion and further research

Our research questions for this paper were (i) What morphosyntactic strategies does Luganda use to express focus? (ii) What type of focus do these express? and (iii) How can different types of focus be identified in elicitation? We have shown that Luganda uses the Immediate After Verb position, the Preverbal Focus Construction and augmentless nouns to express focus, and that, while the IAV position expresses an underspecified type of focus (in the sense of merely triggering a set of alternatives), the PFC expresses identificational focus and the augmentless nouns encode exclusive focus (in affirmative clauses). As regards methodology, while we acknowledge the value and necessity of spontaneous speech and corpora (especially in the area of information structure) we have shown how specific diagnostics for focus are useful, and in fact indispensable, in elicitation if one aims to establish the precise semantic and pragmatic interpretation of a certain strategy that is associated with focus. The spontaneous and elicited data should ideally go hand in hand.

   This narrow study of the interpretation of focus in Luganda triggers further questions, both for Luganda and for the study of information structure in general. We mention three of the issues that arise.

   First, we need to clarify the syntactic status of the PFC and [-A] nouns, as indicated above. The PFC shows properties of both cleft and non-cleft constructions, suggesting

---

[13] Note again that the subject cannot be focused in its canonical preverbal position.

it is halfway in the grammaticalisation path from a cleft to a monoclausal focus construction (cf. Heine & Reh 1983; Harris & Campbell 1995; van der Wal & Maniacky 2015). What should the synchronic syntactic analysis be? With respect to the augmentless nominals, their appearance seems to be restricted by position (vP-internal) and licensing by focus or negation. Could this be related to abstract Case licensing (cf. Halpert 2012; 2013; Carstens & Mletshe 2015; 2016)? Does the "size" of the nominal phrase play a role, with demonstratives also licensing [-A]?

Second, which features do we assume in the structural analysis of focus? Considering the relevance of exclusivity, more specific features than just an underspecified [focus] are clearly needed (cf. Neeleman et al. 2009). A further question concerns whether the morphosyntactic and interpretational effects are part of the syntax, or belong in (the interface of morphosyntax with) semantics and pragmatics.

This is related to a third point, which is the broader question of how we define focus and the pragmatic/semantic divide. The triggering of an alternative set, as proposed in Rooth's (1985; 1992; 1996) Alternative Semantics, can be seen as a unified core function of focus, and the various types of focus are then the outcomes of additional pragmatic and semantic factors (Zimmermann & Onea 2011). Rather than having a dichotomy between identificational vs. new information (É. Kiss 1998), or Kontrast vs. Rheme (Vallduví & Vilkuna 1998), Bazalgette (2015) suggests a four-way distinction between types of focus, according to their semantic complexity: simple focus (triggering alternatives under Roothian semantics of focus and nothing else), focus with an implicature (triggering alternatives, and pragmatically implying exclusion or scalarity), focus with a presupposition, and focus with truth-conditional effects (triggering alternatives and operating upon them). The usefulness of this distinction is illustrated in this paper by the different linguistic strategies for expressing focus in Luganda and their specific interpretation as identificational focus and exclusive focus.

Acknowledgements

Abbreviations and symbols

High tones are marked by an acute accent, falling by a circumflex. Numbers in glosses refer to noun classes, but to persons when followed by SG or PL.

*	ungrammatical	NEG	negation
#	infelicitous	OM	object marker
A	augment	PASS	passive
APPL	applicative	PAST	past tense
CAUS	causative	PERF	perfective
CONN	connective	PFC	preverbal focus construction
COP	copula	PX	noun class prefix
DEM	demonstrative	RED	reduplication
FS	final suffix	REL	relative
IAV	immediate after verb	SM	subject marker
LOC	locative		

References

Bazalgette, Timothy. 2015. *Algorithmic acquisition of focus parameters*. Cambridge: Cambridge University (Doctoral dissertation.)

Byram Washburn, Mary. 2013. *Narrowing the focus: Experimental studies in exhaustivity and contrast*. Los Angeles: University of Southern California. (Doctoral dissertation.)

Carstens, Vicki & Loyiso Mletshe. 2015. Implications of Xhosa expletive constructions. *Linguistic Inquiry* 46(2). 187–242.

Carstens, Vicki & Loyiso Mletshe. 2016. Negative concord and nominal licensing in Xhosa and Zulu. *Natural Language and Linguistic Theory* 34. 761–804.

Destruel, Emilie, Daniel Velleman, Edgar Onea, Dylan Bumford, Jingyang Xue & David Beaver. 2015. A cross-linguistic study of the non-at-issueness of exhaustive inferences. In Florian Schwarz (ed.), *Experimental perspectives on presuppositions*, 135–156. Berlin: Springer.

Dik, Simon C. 1997. *The theory of Functional Grammar*. Berlin/New York: Mouton de Gruyter.

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74(2). 245–273.

É. Kiss, Katalin. 2009. Deriving the properties of structural focus. In Arndt Riester & Edgar Onea (eds.), *Focus at the Syntax-Semantics interface* (Working Papers of the SFB 732.3), 19–33. Stuttgart: University of Stuttgart.

Halpert, Claire. 2012. *Argument licensing and agreement in Zulu*. Cambridge, MA.

Halpert, Claire. 2013. Structural case and the nature of vP in Zulu. In Shayne Sloggett & Stefan Keine (eds.), *Proceedings of the 42nd Meeting of the North East Linguistic Society* (NELS 42), 209–222. Amherst, MA: Graduate Linguistics Student Association (GLSA), University of Massachusetts.

Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross linguistic perspective*. Cambridge: Cambridge University Press.

Heine, Bernd & Mechtild Reh. 1983. Diachronic observations on completive focus marking in some African languages. *Sprache und Geschichte in Afrika* 5. 7–44.

Horn, Laurence R. 1972. *On the semantic properties of logical operators in English*. Los Angeles: University of California.

Hyman, Larry M. & Francis X. Katamba. 1990. The augment in Luganda tonology. *Journal of African Languages and Linguistics* 12. 1–45.

Hyman, Larry M. & Francis X. Katamba. 1993. The augment in Luganda: Syntax or pragmatics? In Sam Mchombo (ed.), *Theoretical aspects of Bantu grammar*, 209–256. Stanford: CLSI.

Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Working papers of the SFB 632, interdisciplinary studies on information structure,* 13–55. Potsdam: Universitätsverlag Potsdam.

Levinson, Stephen C. 2000. *Presumptive meanings*. Cambridge, MA: MIT Press.

Neeleman, Ad, Elena Titov, Hans van de Koot & Reiko Vermeulen. 2009. A syntactic typology of topic, focus and contrast. In Jeroen Van Craenenbroeck (ed.), *Alternatives to Cartography,* 15–52. Berlin: Mouton de Gruyter.

Nurse, Derek & Gérard Philippson. 2003. Introduction. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 1–12. London: Routledge.

Onea, Edgar & David Beaver. 2011. Hungarian focus is not exhausted. In Ed Cormany, Satoshi Ito & David Lutz (eds.), *Proceedings of semantics and linguistic theory (SALT)*, 342–359. eLanguage. Retrieved from: http://journals.linguisticsociety.org/proceedings/index.php/SALT/issue/view/90.

Pak, Marjorie. 2008. *The postsyntactic derivation and its phonological reflexes*. Philadelphia: University of Pennsylvania. (Doctoral dissertation.)

Rooth, Mats. 1985. *Association with focus*. Amherst, MA: University of Massachusetts. (Doctoral dissertation.)

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Rooth, Mats. 1996. Focus. In Shalom Leppin (ed.), *The handbook of contemporary semantic theory*, 271–297. Oxford: Blackwell.

Skopeteas, Stavros, Ines Fiedler, Samantha Hellmuth, Anne Schwarz, Ruben Stoel, Gisbert Fanselow, Caroline Féry & Manfred Krifka. 2006. *Questionnaire on information structure (QUIS)* (Interdisciplinary Studies on Information Structure 4). Potsdam: University of Potsdam.

Vallduví, Enric & Maria Vilkuna. 1998. On rheme and kontrast. In Peter W. Culicover & Louise McNally (eds.), *The limits of syntax*, 79–106. New York: Academic Press.

van der Wal, Jenneke. 2016. Diagnosing focus. *Studies in Language* 40(2). 259–301.

van der Wal, Jenneke & Jacky Maniacky. 2015. How 'person' got into focus: Grammaticalisation of clefts in Lingala and Kikongo areas. *Linguistics* 53(1). 1–52.

Walusimbi, Livingstone. 1996. *Relative clauses in Luganda*. Cologne: Rüdiger Köppe Verlag.

Zimmermann, Malte & Edgar Onea. 2011. Focus marking and focus interpretation. *Lingua* 121(11). 1651–1670.

Chapter 20

Variation in the expression of information structure in eastern Bantu languages

Steve Nicolle

Canada Institute of Linguistics / SIL International

Despite significant typological similarities, eastern Bantu languages differ in how information structure is expressed, but much of this variation only becomes apparent when discourse considerations are taken into account. Using data from narrative texts in eleven eastern Bantu languages I highlight three parameters of variation. First, in some languages surveyed all topics must be left dislocated, while in others certain kinds of topics in specific discourse contexts may be right dislocated. Second, all languages surveyed express argument focus on non-subjects through cleft constructions and right dislocation. However, while argument focus can be expressed through right dislocation of the subject in most languages surveyed, it can be expressed only through the use of cleft constructions in three languages, one of which allows right dislocation of subjects in response to content questions. Third, whilst all of the languages have thetic (topicless) sentences with VS constituent order, most also allow SV constituent order in the orientation sections of narratives, and one language allows SV thetic sentences elsewhere as well.

1 Introduction

The eastern Bantu languages[1] are very similar as far as general morphosyntactic parameters are concerned: they all exhibit typical Bantu noun class systems with obligatory subject marking on verbs, optional object marking, and S>V>IO>DO>X constituent order in pragmatically neutral clauses. However, they also exhibit various kinds of morphosyntactic variation (Marten, Kula & Thwala 2007; van der Wal & Biberauer 2014), including variations in constituent order due to different instantiations of information structure (Zerbian 2006; Buell, Riedel & van der Wal 2011; Yoneda 2011; Downing & Hyman 2016). Most studies of information structural variation in Bantu languages to date

---

[1] The term *eastern* here indicates the geographical distribution of these languages within Africa rather than a genetic affiliation.

have relied predominantly on data collected through elicitation. Such data is well suited to isolating salient interpretational effects in individual sentences, but is less appropriate when investigating the effect of larger stretches of discourse on constituent order variation. This methodological bias is found in studies of African languages more generally, so that, for example, Güldemann, Zerbian & Zimmermann (2015: 157) in their review of research into information structure (IS) in African languages, stated, "we largely restrict the discussion to IS phenomena that hold within a sentence and exclude those in discourse units above this level." In part, this situation reflects a lack of available text corpora. Mwamzandi's (2014) study of variation in adnomial demonstratives and reciprocal constructions in Swahili is one of the few investigations into information structure in a Bantu language that is based on an extensive text corpus (the Helsinki Corpus of Swahili, consisting of 12.5 million words from written Swahili texts).

In this paper, I attempt to complement existing research into information structure in Bantu languages in two ways. First, this study is based on the expression of information structure in narrative texts rather than in elicited data.[2] There are a number of limitations but also benefits to be derived from using narrative texts as data sources. The most important limitations are that a) not every type of construction (contrastive topic, given topic, contrastive focus, corrective focus, verum focus, etc.) is attested in each text corpus, and texts representing other genres could well reveal additional constructions and parameters of variation; b) corpus data do not yield negative evidence – that is, textual data only reveals what is possible, not what is impossible (that is, ungrammatical or pragmatically implausible); and c) prosody is only available in cases where a narrative was transcribed from an oral source and that source is still available (punctuation only represents prosody indirectly and imperfectly, and there is no guarantee that prosody will be appropriately reproduced when written narratives are read aloud). The main benefits of a textual approach are that it reveals patterns beyond the sentence, and it is not dependent on speakers' judgements, invented contexts, or translation. The second way in which this study is designed to complement existing research is that it describes variation in information structure in eleven Bantu languages, thereby complementing more detailed studies of individual languages.[3] These languages and the associated data sources are listed in Table 1.

2 Information structural generalizations

Information structure concerns the way in which an utterance or text is structured to accommodate the (assumed) knowledge state of the addressees, thereby helping addressees to arrive at a coherent interpretation of the utterance or text. Information structure in eastern Bantu narrative texts is primarily expressed through variations in the relative order of subject, verb, object and oblique constituents in a sentence, although intonation,

---

[2] Unless otherwise stated, all examples are from texts that were either published in the sources listed in Table 1 or used in their preparation. All examples are in the orthographies used in the published sources.

[3] The current study is part of a larger investigation into narrative discourse in eastern Bantu languages (Nicolle 2015a).

Table 1: Languages included in this study

Language	Clf.[a]	Location	Data	Source
Fuliiru	[flr] DJ63	South Kivu, Democratic Republic of Congo	153 texts studied; data drawn from 13 (1,000 clauses approximately)	Van Otterloo (2011; 2015)
Digo	[dig] E73	Kwale District, Kenya	5 lightly edited oral texts and 2 written texts (864 clauses)	Nicolle (2015b)
Jita	[jit] EJ25	Mara Rural District, Tanzania	10 lightly edited oral texts (1,096 clauses)	Pyle & Robinson (2015)
Kwaya	[kya] EJ251	Mara Rural District, Tanzania	10 lightly edited oral texts (1,015 clauses)	Odom (2015)
Suba-Simbiti	[ssc] EJ403	Mara Rural District, Tanzania	8 lightly edited oral texts (513 clauses)	Masatu (2015)
Kabwa	[cwa] EJ405	Mara Rural District, Tanzania	9 lightly edited oral texts and 2 written texts (530 clauses)	Walker (2011)
Rangi (Langi)	[lag] F33	Kondoa District, Tanzania	66 texts (3,200 clauses)	Stegen (2011)
Mwani	[wmw] G403	Cabo Delgado Province and Quirimba archipelago, Mozambique	7 texts (number of clauses not known)	Floor (2005)
Bena	[bez] G63	Wanging'ombe District and Njombe District, Tanzania	3 written and 3 lightly edited oral texts (674 clauses)	Broomhall (2011); Eaton (2015a)
Malila	[mgq] M24	Mbeya Rural District, Tanzania	10 lightly edited oral texts and 4 written texts (755 clauses)	Eaton (2015b)
Makonde	[kde] P23	Makonde Plateau, Mozambique	8 lightly edited oral texts (585 clauses)	Leach (2015)

[a] Languages are classified using the ISO 639-3 language code as cited in *Ethnologue* (Lewis, Simons & Fennig 2015), indicated in square brackets in this table; and Maho's (2003) updated version of Guthrie's (1967-71) referential classification of the Bantu languages, indicated by capital letter(s) plus numeral.

pauses and – in certain languages – focus markers also play a role. Three parameters of variation will be discussed in this paper:

1. the position of the topic, specifically whether any topics can be right dislocated;
2. how argument focus on the subject is expressed;
3. how thetic sentences are expressed, specifically whether SV constituent order is possible.

Before looking in detail at how each of these parameters is instantiated in the data, I shall define the terms topic, focus and thetic sentence, and related concepts.

Although topic and focus are complementary notions, they belong to separate dichotomies (De Cat 2007: 66; Erteschik-Shir 2007: 42). The first dichotomy is topic-comment (corresponding to what Halliday (1967) and Dik (1981; 1989) call *theme-rheme*): the (conceptual) topic of a sentence is what a sentence provides information about, and the comment is whatever is predicated about the topic. Topic expressions[4] may be syntactically integrated with the comment (internal topics) or syntactically non-integrated (external topics, also called *topic frames* or *themes*[5]). Both possibilities are illustrated in example (1) below. The context is that the animals have dug a well, but Hare did not help; Hare has repeatedly stolen water, so Tortoise plans to catch Hare and cut off his tail as a punishment. The topic of each clause is *oyo* 'that', referring to Hare. In the first clause this is a non-integrated (external) topic, and in the second clause it is an integrated (internal) topic, as it is coreferential with the object prefix *mu-*.

(1) Jita ('Well' text, line 33)
    [Oyo]_{TOPIC} [munane   era ripanga,]_{COMMENT}
     that_one   give(PL).me just machete
    [oyo]_{TOPIC} [enimugwata   ara.]_{COMMENT}
     that_one   I.will.him.catch just

    'That one, just give me a machete, that one, I'll just catch him.'

The second dichotomy is focus-presupposition: The presupposition of a sentence is the grammatically or lexically expressed information "which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered" (Lambrecht 1994: 52). The focus is defined in contrast to the presupposition as the grammatically or lexically expressed information "which cannot be taken for granted at the time of speech" (Lambrecht 1994: 207) and which, moreover, is considered by the speaker to be the most important information in the sentence (Dik 1989: 277). Within this broad definition, focus is often correlated with new information (Good 2010: 39), where information can be new either in relation to the discourse as a whole or in relation to the

---

[4] Henceforth *topic* will be used to designate a topic expression (i.e. a linguistic form) and *conceptual topic* will be used where the referent rather than the referring expression is intended.

[5] Dik (1981: 129–144) makes a distinction between *theme*, which is defined as an expression designating "a domain or universe of discourse with respect to which it is relevant to pronounce the following predication" (ibid. 130), and *topic*, which is part of that predication.

predicate in question (referred to in Gundel & Fretheim 2004 as referential givenness-newness and relational givenness-newness respectively), but also includes contrastive and identificational focus (see Gibson et al. To appear).[6]

Finally, thetic sentences are sentences in which neither the topic-comment dichotomy nor the focus-presupposition dichotomy is applicable. Thetic sentences are typically used to introduce new participants into a narrative, and to report events or situations in which neither the action nor the participants can be taken for granted, for example, in answer to a general question such as "What happened?" As such, they do not contain a topic. It has be argued that thetic sentences contain neither topic nor focus expressions (Yoneda 2011: 755), however following Lambrecht (1994; 2000) and Nicolle (2015a: 55), we will assume that, since the whole sentence is informationally prominent (that is, it is asserted), a thetic sentence exhibits sentence focus.

3 Obligatory and optional left dislocation of topics

The pre-verbal domain is restricted to lexical topics and non-focus subjects in all of the eastern Bantu languages surveyed; that is, focus elements may not occur pre-verbally. This may be true for most eastern Bantu languages (see Zerbian 2006; van der Wal 2009; Yoneda 2011), except for languages such as Kikuyu [E51], which has a pre-verbal focus marker that is derived historically from a copula construction (Schwarz 2003; 2007; van der Wal 2014). Pre-verbal topics are said to be *left dislocated*, meaning that they occur outside of the clause nucleus. However, there is variation concerning whether all topic expressions must be left dislocated or whether certain topics may be right dislocated in specific discourse contexts.

Left dislocation is used in a broad sense (as in Shaer et al. 2009) which subsumes both left dislocation proper and topicalization.[7] Left dislocation proper involves a resumptive element, such as the class 9 object marker *i-* in example (2) below, which is coreferential with the topic *barabara* 'road'. Topicalization involves a non-resumptive ('gap') construction, as in (3) where there is no object prefix corresponding to the topic *pesa* 'money'.

(2) Digo (Nicolle 2013: 237)
     [Barabara]_{TOPIC} ndipho [a-ka-**i**-rich-a.]_{COMMENT}
      9.road   then   3SG-SEQ-9.OM-leave-FV

     'The road, then, he left it.'

---

[6] For a critical evaluation of the notion of focus as a universal linguistic category, see Matić & Wedgwood (2013).

[7] In this paper, *topicalization* refers to a structural property and does not entail that the constituent exhibiting this property also functions as a topic in discourse. In the literature the term *fronting* is also used to refer to both topicalization and left dislocation proper (cf. Cohen 2009: 313).

(3) Digo (Nicolle 2013: 228)
    [Pesa]_{TOPIC} sino [hu-na-Ø-hew-a.]_{COMMENT}
     10.money us   1PL-PRES-Ø-be.given-FV
    'Money, us we are given.'

The topics in the examples above correspond to the grammatical object of each clause. In Makonde, Fuliiru, Mwani, and possibly in other languages (but not in Digo), left dislocation of the object can also be used to give prominence to the final constituent of the sentence. In (4) below, moving the object *kirya kijumba* 'that box' out of the comment emphasizes the fact that it was laid on the bed (as opposed to being opened, laid elsewhere, etc.). The object is left dislocated but follows the subject, which is also left dislocated, giving two topics and giving focal prominence to the comment.[8]

(4) Fuliiru (Van Otterloo 2011: 350)
    [Ulya munyere]_{TOPIC1} [kirya kijumba]_{TOPIC2} [a-na-ki-gwejez-a ku
     1.that 1.girl   7.that 7.box   3SG-SEQ-7-lay-FV on
    ngingo.]_{COMMENT}
    bed
    'That girl, that box, she laid it on the bed.'

Most topics, however, are subjects, and such topic-comment constructions conform to the canonical SVO constituent order. Nevertheless, subjects in topic position can be separated from the verb by a non-core element such as an exclamative or adverbial phrase, such as *bhuri rusiku* 'every day' in the following example:[9]

(5) Jita (Pyle & Robinson 2015: 32)
    Eyo   mw=ibhara [wamembe]_{TOPIC} [bhuri rusiku :aa-jaga[10]   mu=mugunda
    there in=forest   hyena   every day he.PST-went in=field
    gwaye.]_{COMMENT}
    his
    'There in the forest, Hyena, every day he went to his field.'

Topics may also be extracted from their host clause(s), which then intervene between the topic and the comment. In (6), the topic *rhibuyi eryo* 'that rock' is the object of a purpose clause, which in turn is the complement of a possessive clause, which is the complement of 'see' in a temporal clause. (The embedded clauses are indicated in square brackets. Extraction from multiple embedded clauses such as this may not be possible in all languages surveyed.)

---

[8] For discussion of a similar construction in Matengo, see Yoneda (2011: 756–758).

[9] In this and subsequent examples glossing will be by word rather than by morpheme as a morpheme-based gloss is not necessary for the analysis of constituent order. In long examples, for reasons of space, only a free translation will be provided.

[10] In examples from Jita and Kwaya, the symbol <:> at the beginning of a verb indicates far past tense and the symbol <ˆ> at the beginning of a verb indicates narrative tense with 3SG subjects. The past anterior and the far past, and the 3SG form of the narrative and the 1SG form of the anterior are only distinguished through tone, which is not marked in the orthography. Thus, these symbols differentiate the forms.

(6)   Jita (Pyle & Robinson 2015: 32)
      Ribhuyi eryo$_i$ [ejire arora [atari   na   [ja kwasisya Ø$_i$]]] eeganirisya muno.
      rock   that   when he.saw he.is.not with of to.break Ø   he.thought much
      'That rock$_i$, [when he saw [he has nothing [to break (it$_i$)]]], he thought a lot.'

In all of the languages surveyed, topics are left dislocated at points of textual disconti-
nuity. Textual discontinuity occurs at episode and paragraph breaks, when events are
presented in a non-iconic order (including elaborations and parenthetical material), and
when the topic is a *switch topic*. A switch topic (also called *shifted topic* and *link* in
Vallduví 1992: 109–110) occurs when the conceptual topic of the sentence under consid-
eration is different from the conceptual topic of the immediately preceding sentence. In
the following example, the topic changes from *yuno* 'that one', referring to the speaker's
brother, to *go mafuha ga taa* 'that lamp oil'.

(7)   Digo ('Sababu ya Bahari Kuhenda Munyu' text, line 11)
      Pho [yuno]$_{\text{TOPIC}}$ [kanipha   vitu]$_{\text{COMMENT}}$ ananipha   mafuha ga taa
      so   that.one   he.does.not.give.me things   he.gives.me oil   of lamp
      bahi, [go   mafuha ga taa]$_{\text{TOPIC}}$ ndo [n'yarya?]$_{\text{COMMENT}}$
      only   that oil   of lamp   EXCL I.go.eat
      'So that one doesn't give me anything, he only gives me lamp oil, that lamp oil –
      can I go eat it?'

A switch topic is distinct from a continued topic, which is a topic that was also the
topic of the previous sentence. Both switch topics and continued topics that occur at
points of discontinuity are typically expressed lexically, and always in a pre-verbal posi-
tion (with the exception of renewed topics in Mwani and temporary topics in Rangi, to
be discussed below). Continued topics at points of discontinuity include conceptual top-
ics that continue across paragraph and episode boundaries, and topics that are repeated
when describing events narrated out of sequential order, such as elaborations. In the
following example, the conceptual topic remains the man who is mentioned in the first
sentence (indicated in bold). The sentence after the direct quotation starts a new para-
graph, indicated by the use of the past tense *a-* in *warima* 'he farmed' as opposed to the
consecutive tense *chi-* in *achiamba* 'he said' in the previous sentence. The sentence ini-
tiating the new paragraph starts with the continued topic *yuya bwana* 'that gentleman'
in pre-verbal position:

(8)   Digo (Nicolle 2015b: 65)
      Yuya mlume achiamba, "Mino rivyo nchirima tsula n'naphaha, phahi ndarima
      dza phapha na ko Mwamtsola, ili niphahe vitu vinji vyanjina niguze nigule
      ng'ombe."
      Yuya bwana warima munda uchifika dza Mazera, na hiku uchifika dza Malindi
      ela kaguwire hata tsere mwenga.
      '**That man** said [CONS], "Me when I farmed a termite mound I was getting (a lot
      of food), so I shall farm from here to Mwamtsola, so that I will get lots of things
      to sell so I can buy a cow."

That gentleman farmed [PST] a field as large as from from Mazeras to Malindi, but he didn't get even one grain of maize.'

The majority of continued topics do not occur at points of discontinuity, and as most topics are subjects, continued topics are typically expressed using just a subject prefix on the verb. Since the subject prefix is obligatory in all of the languages surveyed, strictly speaking there is no overt *lexical* topic expression in such clauses, though there is an understood conceptual topic. Occasionally, however, a continued topic is expressed lexically even when there is no discontinuity in the text. When this occurs, eastern Bantu languages differ in how such topics are expressed. In most eastern Bantu languages, topics must occur pre-verbally, whereas in others, topics in certain discourse contexts may occur post-verbally.

In most languages surveyed, all topics are left dislocated, and any post-verbal element is interpreted as the focus. The following examples from Jita and Fuliiru texts illustrate left dislocated continued topics in which the topic (*wamembe* 'hyena' in Jita and *wandare* 'lion' in Fuliiru) is a subject (preserving the canonical SV constituent order); the following verbs contain subject agreement and are inflected for narrative tense (Jita) and sequential tense (Fuliiru):

(9) Jita ('Hare and Hyena' text, lines 8-10, Allison Pyle & Holly Robinson, p.c.)
Wamembe nayomba, wamembe nasurumbara, wamembe najira obhuramusi,
hyena   he.spoke hyena   he.lamented hyena   he.got decision
bhwokuja-otema obhurembo, natura mumugundu gwaye.
he.went-cut   sap/birdlime he.put in.field   his
'Hyena spoke, hyena lamented, hyena made up his mind, he went and cut birdlime (tree sap), (and) he put (it) in his field." (Free translation by Steve Nicolle)

(10) Fuliiru (Van Otterloo 2011: 541)
Wandare anayuvwa   kwâkola mulirira   umwana. Wandare anabwîra
lion   he.heard(SEQ) that   it.is.crying.for child   lion   he.told(SEQ)
uyo mushaaja...
that old_man
'Lion heard that it [the cow] is crying for (its) child. Lion told that old man...'

In Digo, Mwani and Rangi, topics may be right dislocated under certain conditions. It should be noted that right dislocated topics are distinct from post-verbal subjects which may occur in thetic sentences (see §5 below). Right dislocated topics refer to specific, identifiable participants, are often marked as such (for example, they are modified by demonstratives or are proper names), and are separated from the verb by a pause and sometimes by non-core elements. Like left dislocated constituents, right dislocated constituents are outside of the clause nucleus. In contrast, post-verbal subjects are grammatical subjects in their own right and are never separated from the verb by a pause or by non-core elements.

In Digo and Mwani, continued topics are right dislocated when there is textual continuity, that is, when events are presented in sequence within a single thematic unit (i.e.

a paragraph). In the following example, the continued topic *mutu yuyu* 'this person' is right dislocated. This is possible because there is no change of topic and no discontinuity; the use of the consecutive tense marker *chi-* in *achinyamala* 'he stayed silent' indicates a sequential action within a single paragraph:

(11) Digo (Nicolle 2015b: 27–28)
Achidziuza mwakpwe rohoni, "Pho munda, nkauhenda mkpwulu na sikaphaha hata tsere mwenga, kpwani nini?" Lakini [achinyamala]_{COMMENT} [**mutu yuyu**]_{TOPIC} wala kayagomba na mutu.
'He asked himself [CONS] in his heart, "That field, I have made it big and I have not got even a single maize cob, but why?" But he stayed silent [CONS] **this man**, neither did he speak with anyone.'

In (12), from Mwani, the conceptual topic does not change but is referred to using a noun phrase and proper names in the last clause. There is no discontinuity in the text as the order of events is iconically represented: they were called and then they came, and so the topic is right dislocated.

(12) Mwani (Floor 2005: 10)

Wakati	waifikire	sumana	yawasikizane,	wakitíwa,
time	when.it.arrived	week	that.they.agree	they.were.called

[wakíja]_{COMMENT}	[wó-wawiri,]_{TOPIC}	Anli	na	Ntendaji.
they.came	those-two	NAME	and	NAME

'When the week that they agreed upon arrived, they were called, they came those two, Anli and Ntendaji.'

In addition, in Mwani a topic may also be right dislocated if it is a renewed topic, that is, an element that has previously functioned as a topic but not in the immediately preceding clause (Floor 2005: 10-11). (It is not clear whether this is always the case or is an option.) Example (13) shows the difference between a right dislocated renewed topic, *muka ire* 'that woman', which was last mentioned two clauses earlier, and a left dislocated switch topic, *vinu vire* 'those things', which has not previously functioned as a topic.

(13) Mwani (Floor 2005: 10–11)

Sambi	[akikála]_{COMMENT}	[muka	ire]_{TOPIC}	na	[vinu	vire]_{TOPIC}	[akipíka
now	she.sat	woman	that	and	things	those	she.cooked

akiwápa	wanu.]_{COMMENT}
she.gave.to.them	people

'Now she sat down that woman and those things she cooked (them) and gave to the people.'

In Rangi, a switch topic may be right dislocated if it is only temporary; that is, if it functions as the topic of a single clause but the previous conceptual topic is resumed

immediately after. This is optional, however, as not all temporary topics are right dislocated. In (14), the subject twice changes from the elder to the boys and immediately back to the elder. The first clause involving the boys has SV order but the second has VS order. The boys are a temporary topic whereas the elder is more permanent: he is introduced formally in the first clause and it is he who speaks at the end of this passage; these are typical features of major participants (Dooley & Levinsohn 2001: 119).

(14) Rangi (Stegen 2011: 532–533)
Aho kalɨ kwijáa na **moosi ʉmwɨ**, afuma iʉndii, maa **vatavana** navo viyokʉʉja, maa akavasea, "Mpokeri isɨrɨ raanɨ." [Maa vakasiita]COMMENT [**vara vatavana**.]TOPIC Maa [**ʉra moosi**]TOPIC [akasea…]COMMENT
'In times of old there was **one elder**, he came from the field, and **the boys** and they, they are coming, and he told them, "Carry my hoe." And they refused, **those boys**. And **that elder** said…'

In other languages, temporary topics are left dislocated. In the Bena example in (15), *inyama* 'meat' is a left dislocated temporary topic; *u-Mbwa* 'Dog' occurs after the verb because it receives argument focus (see §4).

(15) Bena (Eaton 2015a: 34)
UDuuma aaheliye kwa mwipwave kuhungila kivembo.   [Inyama]TOPIC
Leopard he.went to uncle   to.greet   misfortune meat
[aalekile,]COMMENT [iloleela]PRESUPPOSITION [uMbwa.]FOCUS
 he.left   he.looks.at   Dog
'Leopard went to console his uncle for his bereavement. He left the meat, Dog was looking after it.'

Given the available data, it seems that the majority of the languages surveyed pattern like Jita and Fuliiru in that all topics are obligatorily left dislocated. However, it is possible that evidence may emerge of specific discourse contexts that trigger right dislocation of topics in other languages.

4 The expression of subject argument focus

Argument (or *term*) focus arises when non-predictable information is expressed by a noun phrase. It is found in declarative sentences when a certain event or situation is presupposed, but the speaker assumes that the addressee does not know the identity of one of the participants in that event or situation. In (15) above, it is presupposed that Leopard will not leave his meat unattended and the post-verbal subject *u-Mbwa* 'Dog' identifies who has been left to guard it.

Argument focus in Bantu languages may be expressed in-situ, immediately after the verb (IAV), or in clause-final position (Yoneda 2011: 761–762; Gibson et al. To appear). In most eastern Bantu languages surveyed it appears that argument focus is associated with

clause-final constituents, an exception being Makonde which appears to use the IAV position, in line with most other Bantu languages exhibiting a conjoint/disjoint distinction (Gibson et al. To appear). There is evidence that elements in this position are right dislocated in at least some cases, as the occurrence of the adverbial *woori* 'now' between the presupposition and focus elements in (16) suggests. The context for this example is that the animals have dug a well, but Hare is coming at night and stealing the water. The animals plan to post a guard by the well, and this constitutes the presupposition; the focus identifies who was chosen.

(16) Jita (Pyle & Robinson 2015: 18)
      [Mbamuta-ko]_{PRESUPPOSITION} woori [nyawatare.]_{FOCUS}
       they.put.him-there   now   lion
      'Now they put lion there.'

Makonde does not allow both an object and subject to occur after the verb, and so when there is a post-verbal subject the object is left dislocated (Leach 2015: 91). In Example (17) below, *shakulya* 'staple food' is always served with *imbogwa* 'sauce' and so it is assumed that someone will provide this; the post-verbal subject identifies this person as the speaker.

(17) Makonde (Leach 2015: 92)
      Paukile ndawika kukaja   kumwaulila ndyagwe do: "Ndyangu, taleka
      when   he.went and.arrived he.told.her his.wife QUOT my.wife   cook
      shakulya,   [imbogwa namanya]_{PRESUPPOSITION} [nimwene.]_{FOCUS}"
      staple.food sauce   I.will.know   myself
      'When he got home he told his wife, "Get some food ready for me, wife – but as
      for the meat sauce, I'll deal with that."'

Argument focus can also be expressed through cleft constructions in all languages surveyed. In such constructions, the focused element and the presupposition – often in the form of a relative clause or verbless predicate – are connected using a copula or focus marker. The orders focus>presupposition and presupposition>focus are both found:

(18) Kabwa (Walker 2011: 25)
      Kumbe [omukari wunu]_{FOCUS} ng'we [yankorera   eng'ana yinu.]_{PRESUPPOSITION}
      EXCL   woman   this   COP   she.did.to.me thing   this
      'Gosh, it was this woman who did this thing to me.'

(19) Suba-Simbiti (Masatu 2015: 28)
      [Omoremo ghono yaamanyirë]_{PRESUPPOSITION} [no-bhötëghi ubhwa sinswe.]_{FOCUS}
       work   REL   he.knew   FOC-trapping of   fish
      'The work which he knew is fishing.'

All languages surveyed use both clause-final (or IAV) position and cleft constructions for non-subject argument focus, but there is cross-linguistic variation with subject focus

marking. The fact that subject focus behaves differently from other kinds of argument focus is not surprising. What Güldemann, Zerbian & Zimmermann (2015: 159) refer to as the "default topic-hood of subjects/agents" probably underlies the fact that focus marking on subjects sometimes differs from focus marking on other constituents, for example by requiring explicit focus marking even when a non-subject focus element is unmarked (ibid. 170; see also Fiedler et al. 2010 for a discussion of this phenomenon in Gur, Kwa and West Chadic languages).

In all the languages surveyed, argument focus on the subject can be indicated by cleft constructions. Argument focus on the subject can also be indicated by clause-final position in Jita, Kabwa, Kwaya, Suba-Simbiti, Bena, Malila and Rangi, and by IAV position in Makonde. The Jita example in (20) illustrates both strategies.[11] The post-verbal subject in the first clause identifies 'women only' as those who were living in the land (that the land was inhabited by someone is a presupposition); the cleft construction (with a copula clitic) in the second clause identifies these women as the ones who had stolen Mariro's cows, an event of which the audience is already aware.

(20) Jita (Pyle & Robinson 2015: 34)
Echaaro echo [:bhaariga bheekaye-mo]$_{\text{PRESUPPOSITION}}$ [abhagasi era,]$_{\text{FOCUS}}$
land   that   they.were living-there   women   only
[ni=bho]$_{\text{FOCUS}}$ [:bhaariga bheebire   jing'a ja Mariro.]$_{\text{PRESUPPOSITION}}$
 COP=3PL   they.were they.had.stolen cows of Mariro

'In that land were living women only, it was they who had stolen Mariro's cows.'

Digo and Fuliiru do not allow post-verbal subjects to receive argument focus. In both languages, subject argument can only be expressed using cleft constructions. In Fuliiru, the cleft construction is formed with a 'focus copula' which is cliticized to the following verb:

(21) Fuliiru (Van Otterloo 2015: 345)
[Yàbá bágénì]$_{\text{FOCUS}}$ [bó=bàgírá   yìbì.]$_{\text{PRESUPPOSITION}}$
 these guests   FOC=they.do these

'These guests, they (are the ones who) did these things.'

In Digo, the cleft construction consists of the copula prefix *ndi* (*si* in the negative) plus a referential marker, and is typically, although not always, followed by a relative clause:

(22) Digo (Nicolle 2015b: 55)
Ndipho atu   achimanya kukala [iye]$_{\text{FOCUS}}$ [ndi=ye   ariyehenda mambo
then   people they.knew that   she   COP=1.REF who.did   things
higo.]$_{\text{PRESUPPOSITION}}$
those

'Then people knew that it was her who did those things.'

---

[11] The precise differences in interpretation between different focus positions could not be determined on the basis of narrative corpus data alone.

Finally, in Mwani, argument focus is post-verbal in response to a content question, but when there is no prior question a cleft construction is used (Floor p.c. 8 April 2014; see also Floor 2005: 9):

(23)   Mwani (Floor p.c.)

   a.   Kitabu atwarire   nani?   [Katwala]PRESUPPOSITION [Saidi.]FOCUS
        book   3SG.REL.PST.take who(SUBJECT)   3SG.PST.take   Saidi

        'Who took the book? Saidi took it.'

   a.   [Atwarire]PRESUPPOSITION ndi [Saidi.]FOCUS
        3SG.REL.PST.take   COP Saidi

        'It was Saidi who took it.'

5 VS and SV thetic sentences

As we have seen, the pre-verbal domain in eastern Bantu languages is restricted to topics and non-focus subjects, and so a sentence with canonical SV constituent order will normally be interpreted as expressing topic-comment sentence articulation. It is therefore not surprising that thetic sentences, in which there is no topic, exhibit VS constituent order. These post-verbal subjects are not right dislocated; post-verbal subjects in thetic sentences are grammatical subjects in their own right, they are not topics (and so are generally non-specific or non-established), and they are never separated from the verb by a pause or by non-core elements. The following examples illustrate this.

(24)   Kwaya (Odom 2015: 29)
       Woori bhunu :aariga acheeganiirisha   mmbe ˆn-aa-j-a waarukerwe.
       now   while   he.was he.still.be_thinking so   he.came frog
       'Now while he was still continuing to think, came a frog.'

(25)   Digo ('Mhegi wa Mihambo' text, line 25b)
       ratuluka   fisi,   rina   chitswa dza cha mutu...
       it.emerged hyena it.has head   of   like person
       '...a hyena emerged, it had a head like a person's...'

A common function of thetic sentences is to introduce participants into a narrative. When participants are introduced using a verb of arrival ('come', 'emerge', 'appear', etc.), agreement is with the agent noun phrase in all languages (see the examples above) except Fuliiru. However, when participants are introduced using existential verbs, the verb agrees with the noun phrase in Makonde,[12] Bena, Malila, Jita, Kabwa, Kwaya, Suba-Simbiti and Ekoti, but with a locative noun class in Fuliiru, Digo and Rangi (see Nicolle 2015a: 17–20).

---

[12] van der Wal (2008) reports that languages which distinguish conjoint and disjoint verb forms – which includes Makonde – differ concerning which form is used in thetic sentences; for example, Sesotho [S32/33] uses a conjoint verb form, whereas Makhuwa [P31] uses a disjoint form. Makonde patterns like Makhuwa.

In thetic sentences, only VS constituent order is found in the Digo, Jita and Kabwa text corpora. However, a few texts in the other languages surveyed begin with SV clauses where the subject is not a topic. Most of the subjects in these SV clauses are either well-known folk-tale or animal characters – as in (26) below from Bena – or they refer to non-specific participants, such as 'children' in the Kwaya example (27) and 'one man' in the Fuliiru example (28). This suggests that at the start of a narrative, where no established referents are available to be topics, SV constituent order may be used in thetic sentences when the subject is known to the audience or is (currently) non-specific.

(26) Bena (Eaton 2015a: 54)
      Pa vutalilo [u-Mbwa nu Duuma]$_S$ [vaali]$_V$ nu   wunyalumwinga.
      at start   dog   and leopard   they.were with unity
      'In the beginning Dog and Leopard were together.'

(27) Kwaya (Odom 2015: 60)
      Rusuku rumwi [abhaana]$_S$ [mbaja]$_V$ okureebha.
      day   one   children   they.came to.herd
      'One day children went to herd.'

(28) Fuliiru (Van Otterloo 2015: 104)
      [Mushoshi muguma]$_S$ [akagira]$_V$ lusiku likulu ha'mwage.
       man   one   he.held   day   great at.home
      'One man had a feast at his house.'

Although VS thetic sentences are the norm in most of the languages surveyed, with SV thetic sentences occasionally at the start of narratives, this is not the case in the Suba-Simbiti text corpus. Only one Suba-Simbiti text starts with a VS thetic sentence; of the other texts, three start with SVO thetic sentences, one with a VO thetic sentence, and one with a subject in a copula construction followed by a past form of the verb *rë* 'be':

(29) Suba-Simbiti (Masatu 2015: 16)
      [Musimbëtë na Mohaasha]$_S$ m=bhaana   abha enda   ëmwë [bha-a-rë.]$_V$
       Musimbiti and Mohaasha   COP=2.children 2.ASS 9.stomach 9.one 3PL-PST-be
      'Msimbiti and Mohaasha were siblings, they were.'

Unusually for an eastern Bantu language, Suba-Simbiti also allows SV thetic sentences after the start of a narrative, as in (30). No buffalo has been mentioned previously in the text, and so a topic-comment reading is ruled out. (Two topic-comment sentences follow, in which the buffalo and then the youth function as topics.)

(30) Suba-Simbiti (Masatu 2015: 44)
      Bhoono hano yaarëësyanga   urusikö urwöndë, [eng'era]$_S$ [ekaasha]$_V$
      now   when he.was.herding day   another buffalo   it.came
      mu-rihisho irya waabho   riyo.
      in-group   of   their.place that
      'Another day when he was herding, a buffalo came among their herd.'

6   Conclusions

Based on the available narrative texts, a number of generalizations can be made concerning information structure in the eastern Bantu languages surveyed.

- Obligatory and optional left dislocation of topics

All languages surveyed have left dislocated (pre-verbal) topics, and topic-comment sentence articulation is very common. When there is textual discontinuity (switch topics, episode or paragraph breaks, non-iconic order of events, etc.), topics in all languages are left dislocated. Moreover, all topics are left dislocated in Jita and Fuliiru, and probably in most other eastern Bantu languages. In Jita and Fuliiru, left dislocation of topics is never overridden by other discourse factors. However, certain types of topic are right dislocated in a few languages. Continued topics are right dislocated in Digo and Mwani when there is textual continuity (i.e. within a paragraph); renewed topics may be right dislocated in Mwani; and temporary topics are optionally right dislocated in Rangi. In Digo, Mwani and Rangi, therefore, the default left dislocation of topics can be overridden by discourse factors.

- The expression of subject argument focus

Clause-final or IAV constituents other than the subject may express argument focus in all languages. In Jita, Kabwa, Kwaya, Suba-Simbiti, Bena, Malila, Makonde and Rangi, argument focus on the subject can be expressed both through clause-final or IAV position and through cleft constructions. However, in Fuliiru and Digo, argument focus on the subject can only be expressed using a cleft construction; this appears to be a grammatical constraint and is not tied to specific discourse contexts. In Mwani, argument focus involves right dislocation in response to a content question but a cleft construction is used where there is no prior question.

- VS and SV thetic sentences

Post-verbal subjects are the norm in thetic sentences and are the only possibility in Digo, Jita and Kabwa. However, SV thetic sentences occur in Fuliiru, Kwaya, Suba-Simbiti, Rangi, Bena, Malila, and Makonde for the presentation of new participants at the start of narratives. The Suba-Simbiti corpus also includes SV thetic sentences elsewhere. This suggests that in most eastern Bantu languages surveyed, the default VS constituent in thetic sentences can be overridden by discourse factors.

Abbreviations

1PL	1st person plural	OM	object marker
3PL	3rd person plural	PRES	present tense
3SG	3rd person singular	QUOT	quotative marker
COP	copula	REF	referential marker
EXCL	exclamation	REL	relative marker
FOC	focus marker	SEQ	sequential tense
FV	final vowel		

References

Broomhall, Elizabeth. 2011. Bena text linguistics. Africa International University. Unpublished ms.

Buell, Leston C., Kristina Riedel & Jenneke van der Wal. 2011. What the Bantu languages can tell us about word order and movement. *Lingua* 121. 689–701.

Cohen, Ariel. 2009. Fronted quantificational adverbs. In Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds.), *Dislocated elements in discourse: Syntactic, semantic, and pragmatic perspectives*, 312–327. New York & London: Routledge.

De Cat, Cécile. 2007. *French dislocation: Interpretation, syntax, acquisition*. Oxford: Oxford University Press.

Dik, Simon C. 1981. *Functional grammar*. 3rd edn. Dordrecht: Foris.

Dik, Simon C. 1989. *The theory of functional grammar, part 1: The structure of the clause*. Dordrecht: Foris.

Dooley, Robert A. & Stephen H. Levinsohn. 2001. *Analyzing discourse: A manual of basic concepts*. Dallas, TX: SIL International.

Downing, Laura J. & Larry M. Hyman. 2016. Information structure in Bantu. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 790–813. Oxford: Oxford University Press.

Eaton, Helen. 2015a. *Bena narrative discourse* (SIL Language and Culture Documentation and Description 25). Dallas, TX: SIL International.

Eaton, Helen. 2015b. *Malila narrative discourse* (SIL Language and Culture Documentation and Description 27). Dallas, TX: SIL International.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.

Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz & Malte Zimmermann. 2010. Subject focus in West African languages. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives*, 234–257. Oxford: Oxford University Press.

Floor, Sebastian J. 2005. Marked word-order and focus in KiMwani. Unpublished ms.

Gibson, Hannah, Andriana Koumbarou, Lutz Marten & Jenneke van der Wal. To appear. Locating the Bantu conjoint/disjoint alternation in a typology of focus marking. In Jen-

neke van der Wal & Larry M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu.* Berlin: De Gruyter.

Good, Jeff. 2010. Topic and focus fields in Naki. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure: A documentation of its diversity across Africa*, 35–67. Amsterdam/Philadelphia: John Benjamins.

Güldemann, Tom, Sabine Zerbian & Malte Zimmermann. 2015. Variation in information structure with special reference to Africa. *Annual Review of Linguistics* 1. 155–178.

Gundel, Jeanette K. & Thorstein Fretheim. 2004. Topic and focus. In Laurence R. Horn & Gregory Ward (eds.), *The handbook of pragmatics*, 175–196. Oxford: Blackwell.

Guthrie, Malcolm. 1967-71. *Comparative Bantu. An introduction to the comparative linguistics and prehistory of the Bantu languages. (4 vols.)* Farnborough: Gregg International.

Halliday, M. A. K. 1967. Notes on transitivity and theme in English II. *Journal of Linguistics* 3. 199–244.

Lambrecht, Knud. 1994. *Information structure and sentence form.* Cambridge: Cambridge University Press.

Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language* 24(3). 611–682.

Leach, Rhoda Martyn. 2015. *Makonde narrative discourse* (SIL Language and Culture Documentation and Description 29). Dallas, TX: SIL International.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the world.* 18th edn. Dallas, TX: SIL International. http://www.ethnologue.com.

Maho, Jouni. 2003. A classification of the Bantu languages: An update of Guthrie's referential system. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 639–651. London & New York: Routledge.

Marten, Lutz, Nancy C. Kula & Nhlanhla Thwala. 2007. Parameters of morphosyntactic variation in Bantu. *Transactions of the Philological Society* 105. 253–338.

Masatu, Amanda. 2015. *Suba-Simbiti narrative discourse* (SIL Language and Culture Documentation and Description 31). Dallas, TX: SIL International.

Matić, Dejan & Daniel Wedgwood. 2013. The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics* 49. 127–163.

Mwamzandi, Mohamed Y. 2014. *Swahili word order choices: Insights from information structure.* Doctoral dissertation. Arlington: University of Texas at Arlington.

Nicolle, Steve. 2013. *A grammar of Digo (a Bantu language of Kenya and Tanzania).* Dallas, TX: SIL International.

Nicolle, Steve. 2015a. *A comparative study of eastern Bantu narrative texts* (SIL Electronic Working Papers 2015–003). Dallas, TX: SIL International.

Nicolle, Steve. 2015b. *Digo narrative discourse* (SIL Language and Culture Documentation and Description 26). Dallas, TX: SIL International.

Odom, Shannon Ronit. 2015. *Kwaya narrative discourse* (SIL Language and Culture Documentation and Description 30). Dallas, TX: SIL International.

Pyle, Allison & Holly Robinson. 2015. *Jita narrative discourse* (SIL Language and Culture Documentation and Description 28). Dallas, TX: SIL International.

Schwarz, Florian. 2003. Focus marking in Kikuyu. *ZAS Papers in Linguistics* 30. 41–118.

Schwarz, Florian. 2007. Ex-situ focus in Kikuyu. In Enoch O. Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, 139–160. Berlin: Mouton de Gruyter.

Shaer, Benjamin, Philippa Cook, Werner Frey & Claudia Maienborn. 2009. Introduction. In Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds.), *Dislocated elements in discourse: Syntactic, semantic, and pragmatic perspectives*, 1–27. New York & London: Routledge.

Stegen, Oliver. 2011. *In quest of a vernacular writing style for the Rangi of Tanzania: Assumptions, processes, challenges.* Edinburgh: University of Edinburgh. (Doctoral dissertation.) http://www.sil.org/resources/archives/42655.

Vallduví, Enric. 1992. *The informational component.* New York: Garland.

van der Wal, Jenneke. 2008. Agreement in thetic VS sentences in Bantu and Romance. In Cécile De Cat & Katherine Demuth (eds.), *The Bantu-Romance connection. A comparative investigation of verbal agreement, DPs and information structure*, 323–350. Amsterdam/Philadelphia: John Benjamins.

van der Wal, Jenneke. 2009. *Word order and information structure in Makhuwa-Enahara.* Utrecht: LOT (Landelijke Onderzoekschool Taalwetenschap) Publications.

van der Wal, Jenneke. 2014. Morphological marking of term focus. (Paper presented at Information Structure in Head-Marking Languages workshop, Max Planck Institute Nijmegen, 28-29 March 2014.)

van der Wal, Jenneke & Theresa Biberauer. 2014. From macroparameters to nanoparameters: A comparative Bantu case study. (Paper presented at ComSyn meetings, Leiden, 5 February 2014.)

Van Otterloo, Roger. 2011. *The Kifuliiru language, vol. 2: A descriptive grammar.* Dallas, TX: SIL International.

Van Otterloo, Roger. 2015. *Fuliiru narrative discourse* (SIL Language and Culture Documentation and Description 32). Dallas, TX: SIL International.

Walker, John R. 2011. Discourse features of Kabwa and accompanying texts. Unpublished ms.

Yoneda, Nobuko. 2011. Word order in Matengo (N13): Topicality and informational roles. *Lingua* 121. 754–771.

Zerbian, Sabine. 2006. *Expression of information structure in Northern Sotho.* Doctoral dissertation. Berlin: Humboldt University.

Part II

(Possible) Nilo-Saharan

Chapter 21

Number marking in Lopit, an Eastern Nilotic language

Jonathan Moodie

University of Melbourne

Nilotic and other Nilo-Saharan languages have rich number marking systems and it has been difficult to establish what rules might govern these systems (Dimmendaal 2000). This paper addresses the question of how number is marked in Lopit and describes some possible rules for number marking. Lopit has a three-way system for number marking involving plurative, singulative and replacement marking. Lopit has a greater plural in addition to the normal plural. It also has a special form of number marking whereby the marked singular can be used to denote very large numbers. This form of number marking has not been observed in the literature and I call it the greater singular.

1 Introduction

Lopit is an Eastern Nilotic language spoken by around 50,000 people living in the Eastern Equatoria province of South Sudan as well as by diaspora groups elsewhere in Africa and the world. It is part of the Lotuxo sub-group of the Lotuxo-Maa languages. Until recently, the Lopit language has received little descriptive attention. Some observations on Lopit were made by authors working on the related Otuho (Lotuko) language (Muratori 1938) and in comparative wordlist data collected by Driberg (1932) and Vossen (1982). Lopit has only been the focus of linguistic description and documentation in more recent years. The language has six different dialects (Ngaboli, Dorik, Ngutira, Lomiaha, Lohutok and Lolongo), and data collected with speakers of a number of these has led to observations on aspects of Lopit phonetics and phonology (Turner 2001; Stirtz 2014; Billington 2014) and morphology and syntax (e.g. Ladu et al. 2015). Data presented in this paper are based on a corpus of elicited, storytelling and conversational materials collected with six Dorik speaking members of the Lopit community in Melbourne, Australia, transcribed and compiled using ELAN and Fieldworks. These speakers are aged between 30 and 55 and have migrated to Australia in the last 10 years.

The aim of this paper is to describe how number is marked in Lopit. There appears to be nothing in the literature about this subject apart from some singular and plural nouns in word lists given by Driberg (1932) and Vossen (1982). I describe the different

number marking patterns used, how these patterns are distinguished and what general-isations can be made about the selection of the number marking morpheme. In addition, I describe how Lopit uses two methods for expressing larger amounts than are normally described with plurals. I use the terms *greater plural* and *greater singular* to describe these methods.

2 An overview of number marking

In many languages, a distinction can be made between countable and non-countable nouns. Non-countable nouns cannot be differentiated on the basis of number. They are called mass nouns. In Lopit, mass nouns are inherently either singular or plural. The words for 'milk' and 'water' take plural agreement, whereas those for 'air', 'grass' and 'flour' take singular agreement. Examples are shown in (1) and (2). In Lopit, relative clause pronouns inflect for the gender and number of the noun they modify.[1]

(1) rè   hùná   l-á-rá   h-ìtúrá
     milk REL.F.PL SBO-3SG-be INF-pour

     'sour milk' (lit. 'milk which has been poured')

(2) a. lòyyámì nà   l-ó-nók
         air   REL.F.SG SBO-3SG-be.hot

         'hot air' (lit. 'air which is hot')

     b. *lòyyámì hùná   l-ó-nók
         air   REL.F.PL SBO-3SG-be.hot

Variation in the treatment of mass nouns also occurs in Turkana (Dimmendaal 1983: 224) and some Bantu languages (Corbett 2000: 173). Dimmendaal ascribes these differences to the etymological origin of each particular term (Dimmendaal 2000: 230) rather than to any semantic conceptualisation.

For countable nouns, Dimmendaal (1983: 224; 2000: 214) has examined the patterns of number marking found in Nilo-Saharan languages and describes a tripartite system which is found in many of the languages in this family. This system is shown in Table 1,

Table 1: The tripartite system of number marking

System		Distinction		
plurative marking		base	⟷	plural
singulative marking	singular	⟷	base	
replacement marking	singular		⟷	plural

---

[1] Lopit is a tone language with high, low and falling tones. It has ATR distinctions which are not substantive in relation to number marking and are ignored in this paper. In general the Latin script is used as in English except that the sounds /x/, /ɲ/ and /ŋ/ are represented by <h>, <ny> and <ng> respectively.

Table 2: Tripartite system of number marking in Lopit

Marking System	Singular	Plural	English
Singulative	*hófír-í*	*hófir*	'feather', 'hair'
	hàlá-tí	*hàlà*	'tooth'
	áhér-í	*àhèr*	'star'
Plurative	*bérèt*	*bèrét-ì*	'flag'
	hírrí	*hìrrì-jà*	'waterhole'
Replacement	*húng-ú*	*hùng-à*	'knee'
	fàìt-î	*fàìt-ô*	'ebony tree'

Table 3: Irregular singular/plural forms

Singular	Plural	English
háná	*hàss*	'hand'
máná	*máttá*	'farm'
híténg	*hísúng*	'cow'
sòhínè	*sáng*	'thing'
wór	*wónnì*	'valley'

which has been adapted from Dimmendaal (1983: 224) and Corbett (2000: 156). In Dimmendaal's terminology, plurative marking is where the plural form has a morphological marker and the singular form is the unmarked base or root. Singulative marking is where the singular form has a morphological marker and the plural form is the unmarked base. Replacement marking is where both the singular and the plural forms have a morphological marker and the base is not specified for number and is not found as a word. The morphological markers are usually suffixes.[2]

Lopit follows the tripartite system of singulative, plurative and replacement marking. Some examples are shown in Table 2. Note that there is a large range of segmental morphemes which can be used to mark number. This is discussed further in §4.

Some singular/plural relationships appear to involve irregular or suppletive forms. Some examples of these are shown in Table 3. There is some morphophonemic similarity between the singular and plural in all these examples.

Another method of number inflection is tonal modification, although this has only been observed for a small number of nouns. No pattern has yet been discerned for the tonal alternations between singular and plural forms. Some examples are shown in Table 4.

---

[2] Corbett (2000: 156) has also examined tripartite systems, largely using Nilo-Saharan examples. He uses a similar classification to Dimmendaal except that he uses the terms Type A, B and C for the plurative, singulative and replacement marking systems respectively. In this paper, I use Dimmendaal's terms.

Table 4: Examples of tonal number inflection

Singular	Plural	English
hínè	*hìnè*	'goat'
yànì	*yání*	'fruit, tree'
mònyè	*mònyê*	'father'

I examined 446 Lopit nouns to determine the distribution of the various systems of number marking and to investigate potential rules. The frequency distribution of the different systems is shown in Table 5. The plurative marking system is the most common with 58%, but the proportions of singulative and replacement marking patterns are both significant.

Table 5: Distribution of number marking systems in Lopit

System	Number	Percent
Plurative	257	58
Singulative	85	19
Replacement	79	18
Irregular & tonal	25	5
Total	446	100

So far we have seen that Lopit has a variety of number marking morphemes and that most follow a tripartite marking system. We will now look at how the singulative and plurative marking systems in Lopit can be differentiated.

3 The singulative and plurative distinction

There is a semantic basis for the assignment of a lexeme to singulative versus plurative number marking pattern. The singulative pattern generally applies to those nominal lexemes which name referents which normally occur in groups or large numbers. In addition to the examples in Table 2, other nouns which are unmarked in the plural and take singulative marking include *morro* 'beans', *sana* 'branches' and *sohot* 'coconuts'. The singulative pattern is also used for nouns which are in pairs or finite sets (*hafiela* 'fingers', *iwwa* 'wings'). This is common amongst Nilo-Saharan languages (Dimmendaal 2000: 216; Creissels et al. 2008: 119).

The distinction between singulative and plurative patterns can be related to the concept of individuation in number marking. Corbett (2000: 217) points out that "the groups which we quantify with large numbers are the groups which are less individuated and conversely are more likely to be viewed as a unit". Thus referents of the Lopit words

aher, 'star', *balang*, 'salt' and *hofir*, 'hair' are found in large numbers and are not easily differentiated into single items.

When they are individuated, singulative singulars can often have a specific meaning which refers to a separated item of the referent. Some examples are shown in Table 6.

Table 6: Nouns that take the singulative patterns

Singular	English	Plural	English
ngámá-rì	'grain of sorghum'	*ngàmà*	'sorghum'
báláng-á	'grain of salt'	*báláng*	'salt'
hófìr-í	'strand of hair'	*hófìr*	'hair'

The distinction between singulative and plurative has been examined in some detail by Grimm (2012) for the Dagaare language (Gur: Niger-Congo).[3] He tested 1500 words in Dagaare and found the following (Grimm 2012: 50):

i. Nouns for higher-level (more salient) animals are more likely to be unmarked in the singular than nouns for insects.
ii. Nouns for trees are typically unmarked in the singular in comparison to nouns for vegetation which are typically unmarked in the plural.
iii. Nouns for tools are more likely to be unmarked in the singular than the converse.
iv. Nouns for body parts which inherently come in pairs or groups are more likely to be unmarked in the plural than not; while nouns for body parts which inherently come in single units are more likely to be unmarked in the singular.[4]

I carried out a similar analysis on the set of 446 Lopit nouns and tested the categories of mammal, bird, reptile, insect, tree, vegetation and tool. The results are shown in Figure 1. The resulting trends were somewhat similar to those found by Grimm for the same categories.

For most semantic categories tested, the preferred number marking pattern was clear. However, it was not so clear for the insects and the data was examined in more detail. The examples of insects are listed in Table 7. Those insects which tend to be larger and more likely to be seen individually, such as butterflies, caterpillars and large wasps, take the plurative system. Conversely, those insects which are smaller and/or seen in large numbers such as mosquitoes, lice and flies have the singulative system. Whilst not conclusive, this data tends to support the role of individuation in determining the choice between singulative and plurative patterns for number markings.

The number pattern for the terms for body parts was also investigated. The terms were grouped into those parts which are found singly ('face', 'tongue', 'head') and those which are found in pairs or sets ('eyes', 'fingers', 'hands'). The results are shown in Figure 2

---

[3] Grimm (2012) uses the term *marked singular* rather than *singulative*.
[4] There were a number of exceptions which related to some specific semantic aspects, borrowings and some derived forms.
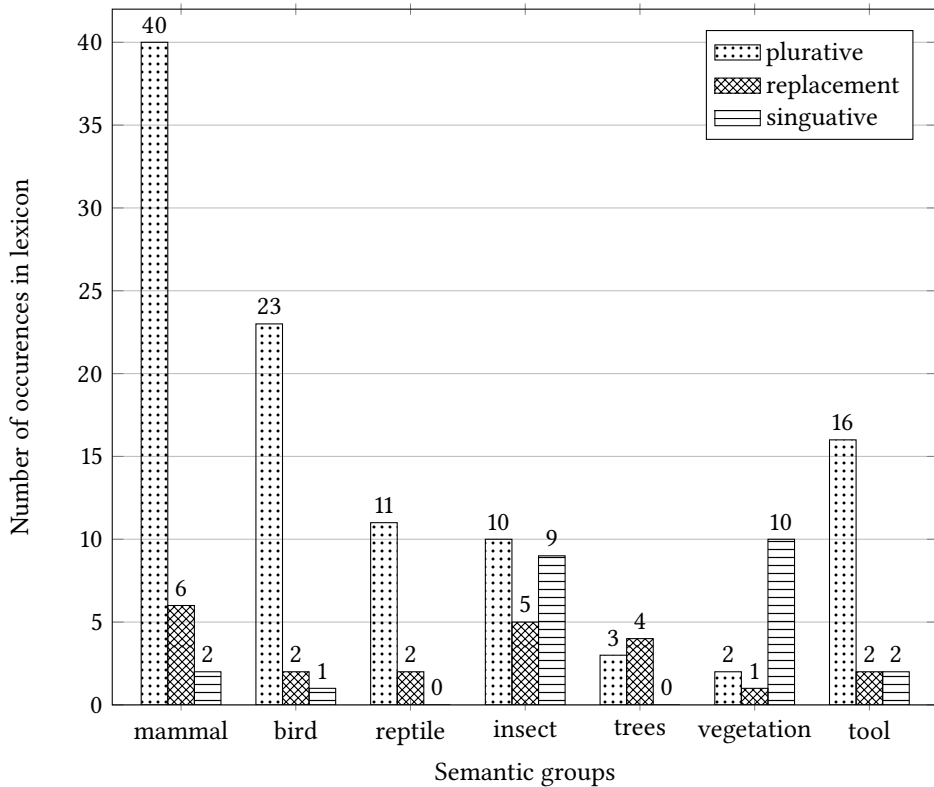
Figure 1: Number marking across semantic domains

and these also resemble the findings of Grimm. That is, body parts which occur singly are more likely to have the plurative pattern and those parts found in pairs or sets are more likely to have the singulative pattern.

It is worth noting that there are a considerable number of examples of replacement marking in the body part data. Replacement marking is not found in Dagaare so this was not examined in Grimm's (2012) work. The proportion of the replacement system in some categories in Lopit is higher than for the singulative system. In some groups, like 'insects' in Figure 1 and 'pairs or groups' in Figure 2, the proportion of replacement marking is quite high. From this data, it might be possible to infer that the proportion of replacement marking is higher for a particular semantic group when there are also high proportions of both singulative and plurative marking in the group. It could be that replacement marking is some kind of intermediate or derived pattern. This is discussed further in §4.3.

Table 7: Plurative and singulative insect nouns

Plurative			Singulative		
Singular	Plural	English	Singular	Plural	English
ífòrí	*ìfòrí-há*	'butterfly'	*hìlófìr-î*	*hìlòfìr*	'bee sp.'
lòmólòrúk	*lòmólòrùh-í*	'ant sp.'	*hìmúrùt-í*	*hìmùrùt*	'mosquito'
hùtèlèk	*hùtèléh-í*	'caterpillar'	*lòngóròm-í*	*lòngòròm*	'termite'
ídóló	*ìdòló-hó*	'locust'	*lòfér-ìtí*	*lóférr*	'tick'
lótáhùlòng	*lòtàhùlóng-í*	'wasp, large mud dauber'	*múhúny-ì*	*múhúny*	'ant, small black'
			ítíngílìyè-tí	*ítíngílìyè*	'ant, small brown'
			làkiè-tí	*làkié*	'louse'
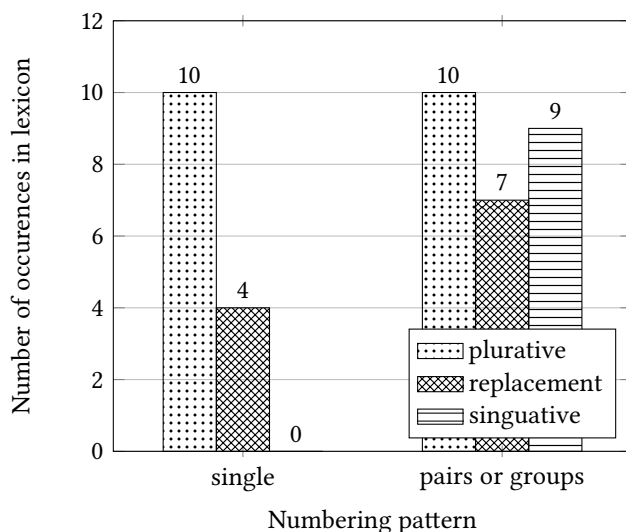
Figure 2: The numbering patterns for body parts

One finding from this study does not appear to be reported by Dimmendaal or Grimm. This is the finding that some words are regarded as singulative by some speakers and plurative by others. Some examples are given in Table 8. For example, some speakers (AL, VH) consider the base of the concept of 'rib' or 'ribs' in Lopit (*mari*) to be singulative in pattern with the singular marked by the suffix *-ti*. A third speaker (DA) considers the base to be of the plurative pattern with the plural marked by the suffix *-jin*.

Table 8: Examples of different speaker's choice of singulative and plurative

AL		VH		DA		JL		English
SG	PL	SG	PL	SG	PL	SG	PL	
màrí-tí	*mârì*	*màrí-tí*	*mârì*	*màrí*	*màrí-jín*			'rib'
kàlà-ì	*kàlá*	*kàlà-ì*	*kàlá*	*kál*	*kál-ì*			'side'
ì-túrèt	*tùrèt*		*tùrèt / ì-túrèt*	*tùrèt*	*turet-i / ì-túrèt-í*	*ì-túrèt*	*tùrèt*	'twin'

This suggests that there is some kind of semantic boundary zone between those concepts which are more likely to take the singulative pattern and those that are more likely to take plurative pattern and that there might be variation between individuals on where this boundary sits. Ribs, sides and twins are found in sets with small numbers of members and are semantically more like individuated things. Thus they would be expected to be closer to the plurative marking. This suggests that, from a semantic perspective, these concepts might be viewed differently to concepts like hair and teeth.

4 Regularity of number marking in Lopit

4.1 Introduction

Many studies have commented on the complexity of number marking patterns in Nilo-Saharan languages and have pointed out that it is very difficult to predict the plural of a plurative pattern lexeme given the singular form of the lexeme (e.g. Tucker & Mpaayei 1955: 4; Hilders & Lawrance 1957: 3). On the other hand, Dimmendaal (2000: 255) states that Nilo-Saharan languages "have a finite system governed by rules" although he acknowledges that further research is required to understand these rules.

This study has found that the forms of plurative, singulative and replacement marking in Lopit are very diverse. A large range of number suffixes was identified and this is common amongst Nilo-Saharan languages (Dimmendaal 2000: 219). An attempt has been made to determine if there is a "finite system governed by rules". The list of 446 nouns was examined and several patterns have been identified. These are shown in Table 9.

Table 9: Patterns for singular and plural formation

Plurative		Singulative		Replacement		
Plural	Number	Singular	Number	Singular	Plural	Number
-i	73	*-i*	30	*-ni*	*-k*	18
-hi	5	*-ti*	34	*hi-*	*-i*	13
-a	23	*-hi*	3	*-V₁*	*-V₂*	12
-ha	32	other suffixes	18	other affixes		36
-o	23					
-ho	16					
-jin	21					
-(h,s,sh,c)in	16					
-Vn	11			Tonal		9
other suffixes	37			Irregular		16

Table 10: Examples of number marking using *-i* and *-Ci*

Plurative			Singulative		
Singular	Plural	English	Singular	Plural	English
tìàng	*tìáng-ì*	'animal'	*cèng-î*	*cèng*	'bird'
sùkùl	*sùkúl-í*	'school'	*fófóng-ì*	*fófóng*	'cactus'
gálám	*gálàm-ì*	'pen'	*sùkár-í*	*sùkár*	'sugar'
bùk	*bùh-í*	'book'	*muáráh-ì*	*muàràk*	'horn'
sà	*sà-tí*	'hour'	*mòró-tí*	*mòró*	'bean'
hò	*hò-sî*	'head'			

The most common form (around 32% of this sample) involves the suffix *-i* or *-Ci*. Note that this is found in both plural and singulative systems. Some examples are shown in Table 10.

A number of these examples are loan words (e.g. the words for 'sugar', 'school', 'pen' (Arabic) and 'book') and this suggests that these forms are productive.[5]

The suffix forms *-a* and *-o* also occur with both singular and plural meaning, as shown in Table 11.

For singulative marking, the suffix form appears to be more predictable. In 75% of the sample, the singular is marked with the suffix *-i* after a stem-final consonant and with the suffix *-ti* after a stem-final vowel. This is similar to the neighbouring Eastern Nilotic language, Lotuko, where Arber (1936: 7) describes *-i* as the "common form" for the singular suffix for singulative nouns

---

[5] It is worth noting that, when the *-i* suffix is added to a word ending in *-k,* the consonant is modified to the velar fricative /x/ (represented here as <h>).

Table 11: Examples of number marking with *-a* and *-o*

	Plural			Singulative		
Singular	Plural	English		Singular	Plural	English
hùnòm	*húnòm-ò*	'cave'		*mòrù-ó*	*mórú*	'stone'
hùró	*hùró-hó*	'goat kid'		*báláng-á*	*bàlàng*	'salt'
hómuóng	*hòmuòng-â*	'face'				
hárí	*hárí-yà*	'river'				

4.2 Suffixes for plurative number marking

One of the main things seen in the data in Table 10 and Table 11 is that it is not obvious how the choice between the *-i*, *-o*, and *-a* suffixes is determined. Table 12 shows the range of stem endings (i.e. the last two phonemes) that are associated with the various plural markers.

A large range of final vowel, consonant, and vowel/consonant combinations can be found with each of the three suffix types. This shows that suffix choice cannot be predicted based on how the noun stem ends. However, there were occasional tendencies for words with similar structures to pattern together, so this was examined further.

To explore what suffix choices might be regularised, three consultants were tested on their intuitions about the plural marking of loan words from English and Arabic and of nonce words.[6] A smaller sample of 66 words was used and several patterns were identified. The syllable structure of the stem appears to play a role. The results are shown in Table 13 together with some examples.

The tests with borrowed and nonce words show patterns which are also present to some extent in the previous lexical data, but with much more irregularity in that lexical data. This suggests that there are some preferences governing choice of number-marking morpheme, allowing the system to be productive, but that historical or other processes may have obscured parts of this system in the Lopit lexicon. On the basis of the tests, the following generalisations for plural formation are provisionally proposed.[7]

(3)   If the noun ends in a consonant, the plural suffix is *-i*.
       If the noun has the form CV, the plural suffix is *-si*.
       If the noun has the form CVCV, the plural suffix is *-sin*.
       If the noun has the form CVCVCV and at least one V is *o*, the plural suffix is *-ho*.
       If the noun has the form CVCVCV and at least one V is *a*, the plural suffix is *-ha*.

---

[6] A nonce word is one which has been made up for testing purposes. Since consonant clusters are not normally found in Lopit words, the testing was based on CV(CV)(C) words. Sometimes, consultants also gave Arabic plurals e.g. *kabaya*, *kabaya-t* and these have been ignored.

[7] I assume that there is a single morpheme (provisionally called *-sin*) which has variable pronunciations which may in part be lexically specified and may also differ depending on the individual. These include *-shin,-sin*, *-cin* and *-jin*.

Table 12: Plural suffixes with the stems to which they attach

General suffix form	Suffix	Last two phonemes on the stem
-(C)i	*-i*	*-ak, -am, -an,-ang, -at, -eng, -ek, -el, -er, -et, -ing, -ir, -it, -of, -ol, -om, -ong, -or, -os, -ra, -re, -ru, -uk, -ul, -um, -uny, -ur, -ut, -wa*
	-hi	*-du, -fi, -fʷo, -fa*
	-ni	*-ra*
	-si	*-bu, ho,*
	-ti	*sa, -ju, -iu, -me, -re, -ri, -ro, -ye,*
-(C)a	*-a*	*-ak, -al, -ar, -ang, -bu, -ef, -en, -er, -ir,-ok, -ol, -ong, -ni*
	-ha	*-ga, -gu, -ia, -ja, -la, -le, -ne, -ra, -re, -ri, -ti, -tu*
	-ja	*-ri, -tu*
	-na	*-ge, -ha*
	-ta	*-he, -ri*
	-ya	*-me, -ni, -ri, -te*
-(C)o	*-o*	*-ing, -ol, -om, -ong, -ony, -ri, -ru, -ti*
	-ho	*-lo, -me, -mu, -ri, -ro, -ti, -wo*
	-jo	*-ti*
	-so	*-he*
	-yo	*-mi, -ni, -ri*

Variation in plural formation in Turkana is similarly related to word structure. Dimmendaal (2000: 235) gives examples of the plurative suffix *-a*, which is found after CVCVC roots, and the suffix *-in*, which is found after CVC roots.

It is also worth noting that there is some variation between speakers. Some of this is may be related to different locations within the same dialect area. However, there is also some variation with the same speaker such as alternating between common suffixes like *-ho* and *-sin*. This was also observed by Unseth (1988: 76) in his study of the Nilo-Saharan Surmic language, Majang. He commented that "A certain amount of variation for marking number on some nouns is noticeable, even by one speaker", and that "Generally, the variation consisted of alternate suffixes". This variation is not surprising given the complexity of the number-marking systems. In addition, alternating some of the common suffixes for plurative nouns is unlikely to cause much ambiguity.

Table 13: Results from the study with loan and nonce words

Stem structure	Singular	Plural	English	Suffix	No. of words with this stem structure	No. of words with this suffix
consonant final	*batik*	*batih-i*	'water-melon'	*-i*	40	39
	gamis	*gamis-i*	'shirt'			
	telifision	*telifision-i*	'television'			
CV	*ka*	*ka-si*	'car'	*-si*	5	3
CVCV	*leta*	*leta-sin*	'letter'	*-sin,*	11	9
	tivi	*tivi-sin*	'TV'	*-jin*		
CVCVCV	*kubaya natana*	*kubaya-ha natana-ha*	'cup' nonce word	*-ha*	10	5
	teroli	*teroli-ho*	'trolley'	*-ho*		3
	pomodi	*pomodi-ho*	nonce word			

4.3 Affix forms for replacement number marking

There are a number of distinct affix forms amongst nouns with replacement marking. Three patterns have been identified although only one of these (the first) is regular and predictable.

(i) Derived nouns describing human roles: Nouns which have been derived from verbs have a well-defined number marking system which involves a number marker in both the singular and plural (i.e. the replacement pattern). These nouns take *-ni* for the singular and *-k* for the plural. This appears to be predictable inflection. This is also common across Nilo-Saharan languages (Dimmendaal 2000: 243). Table 14 shows a list of nominalised verbs with examples of agentive nouns. The table also includes other nouns which are not directly derived from verbs and which are used for describing people. It is worth noting that these words are used for roles that people can take. This number marking system is not used for kinship terms (mother, child, aunt etc.).

(ii) Singular with *hi-* prefix and plural with *−i* suffix: Some examples are shown in Table 15.

This appears to be the only place where a prefix is used for marking number. The prefix is probably the Lopit version of what Greenberg (1981) called the "movable *k*" which has

Table 14: Person referent nouns in singular and plural

Semantics	Stem	Singular	Plural	English
Agentive	*bara*	*há-bárà-nì*	*há-bàrà-k*	'cow farmer'
	itiyena	*há-itíyénà-nì*	*há-itíyénà-k*	'teacher'
	toho	*há-tóhò-nì*	*há-tòhò-k*	'killer'
Other		*mérò-nì*	*méró-k*	'enemy'
		dóngiò-ní	*dòngió-k*	'Lopit (mountain) people'
		márúà-nì	*máruà-k*	'elder'

Table 15: Lexemes in the replacement pattern with prefix and suffix

Singular	Plural	English
hì-reng	*reng-i*	'goat hide'
hì-ríngò	*ringo-i*	'meat'
hí-túny	*túny-í*	'snake (sp.)'
hì-yòk	*yòh-ì*	'ear'

no meaning by itself and predates current Eastern Nilotic (and many other Nilo-Saharan) languages (see also Dimmendaal 1983: 251). There is no clear semantic or phonological commonality amongst the examples in this group.

(iii) Replacement marking involving other changes: As shown in Table 16, this can involve a different vowel for the singular versus the plural.

Table 16: Lexemes in the replacement pattern with prefix and suffix

Singular	Plural	English
hàww-ê	*hàww-â*	'arrow'
húng-ú	*hùng-à*	'knee'
hòw-ê	*hòw-à*	'sweet potato'
tórór-ì	*tórór-ò*	'sugar ant'
mún-ú	*mùn-iòk*	'snake'
lòsíng-òtí	*lòsíng-óng*	'sorghum, red'

Dimmendaal (2000: 242) reports that replacement marking may come about following the loss of the morphological unmarked form in an earlier three-way number alternation. He illustrates with examples from two related Nilo-Saharan languages of the Daju family, Shatt and Sila. The three-way distinction for 'tooth'/'(set of) teeth'/'teeth' in Shatt, given

in (4), can be compared with the two-way distinction in Sila, given in (5). It appears that Sila has lost the morphologically unmarked form of *nyir*, 'teeth'.

(4) Shatt (Dimmendaal 2000: 242)
     nyix-te nyix   nyix-ke
     'tooth' '(set of) teeth' 'teeth'

(5) Sila (Dimmendaal 2000: 243)
     nyir-te nyir-ke
     'tooth' 'teeth'

It might be possible to postulate a process for Lopit somewhat similar to that described by Dimmendaal if we examine those words which were discussed in relation to Table 8. Recall that Table 8 demonstrates different speakers' choices of singular and plural forms of particular lexemes, and shows that one speaker might use a singulative pattern for a particular lexical item, while another speaker might use a plurative pattern for that same lexical item. For example, speakers AL and VH use singulative patterns for *mari* 'rib/s', while speaker DA uses a plurative pattern for *mari*. From this variation within a community, a replacement pair *mari-ti* plus *mari-jin* could potentially develop if the simple form *mari* falls out of use.

The loss of the simple (unmarked) form might be related to speakers' differing perceptions of individuation. It appears that some speakers (e.g. DA) consider the unmarked form *mari* to indicate something individuated in contrast to others (e.g. AL and VH) who consider it to indicate a group or set. If there are significant numbers of both groups of speakers, then there may be some ambiguity in the speech community. It could be that, over time, the unmarked form is used less and less frequently to avoid ambiguity and it disappears from use. Many of the examples in Table 15 and Table 16 are things which might be seen either as groups or as individuated items (e.g. *hiyok*, 'ear'; *hungu*, 'knee'). It could be that there were unmarked forms for these lexemes which have fallen out of use.

5 The greater plural and the greater singular

5.1 The greater plural

Corbett (2000: 30) discusses a three-way distinction between singular-plural-greater plural. The greater plural can cover a number of things including a much larger number than is usually associated with the plural ('plural of abundance') and a number which refers to all the members of a particular entity ('global plural'). The examples given by Corbett cover a range of languages but all involve a distinct morphological marking for the greater plural. Marking with this meaning is seen in Lopit.

Greater plural marking sometimes follows the plurative pattern. Some examples are shown in Table 17. For example, the normal plural of *toru* 'axe' is *toruo*. When the suffix -*sen* is used, it means that there are so many axes that they are uncountable.

Table 17: Examples of singular/plural/greater plural

Singular	Plural	Greater plural	English
tórú	*tòrú-ò*	*tòrú-sèn*	'axe'
gàrì	*gárí-jìn*	*gárí-sèn*	'tracks'
kérr	*kèrr-á*	*kèrr-ítí*	'sheep'

Similarly, *kerriti* can mean a large number of sheep, as is shown in the following.

(6) dè dóngé níà   ó-lúngá   kèrr-á.
    in village that.F 3-be.many sheep-PL

    'In that village there are many sheep.'

(7) kèrr-ítí   nà   dè dóngé nià.
    sheep-GPL there.COP in village that.F

    'In that village there are many sheep.'

Other Eastern Nilotic languages also exhibit two types of plural. For example, Lotuko has a greater plural similar to Lopit (Muratori 1938: 57). In Maasai and Bari there are a plural and a collective plural (Dimmendaal 2000: 242). In Teso, the second plural is a generic plural (Hilders & Lawrance 1957: 4). Some examples are shown in Table 18.

Table 18: Singular, plural, second plural forms in Eastern Nilotic languages

Language	Singular	Plural	Second plural
Lotuko	*nenie* 'goat'	*nedye* 'goats'	*nedye-jin* 'many goats'
Maasai	*eng-ker* 'sheep' (SG)	*ing-kerr-a* 'sheep' (PL)	*ing-kerr-ai* 'herds of sheep'
Bari	*nyomot-i* 'one/a seed'	*nyomot* 'seeds'	*nyomot-an* 'different kinds of seeds'
Teso	*e-tunga-nan* 'man'	*i-tunga* 'men'	*i-tunga-sinei* 'mankind'

5.2 The greater singular

A further aspect in Lopit number marking concerns the fact that singulative noun forms can sometimes have a meaning which indicates a very large number. In Lopit, the word

lome 'millet' is singulative and shows plural agreement (8). The singular form *lometi* usually has the meaning of a grain of millet. This was discussed in relation to Table 6 above. An example is shown in (9). However, it can also be used to express a very large or unbelievable amount, as shown in (10).

(8) é-íríà   ínyéjà lómè   hùná Lòhídòng.
    3SG-grinds 3SG.NOM millet.PL.ABS of.F.PL Lohidong.ABS

    'She grinds Lohidong's millet.'

(9) ó-rùmá   Lóhìdóng   lòmé-tí   dè sáì nànyì.
    3SG-find.PFV Lohidong.NOM millet-SG.ABS in tea his

    'Lohidong found a grain of millet in his tea.'

(10) eí-ngà-bàlú   íyóhoí lòmé-tí.
     1PL-PRF-harvest 1PL.NOM millet-SG.ABS

     'We have harvested so much millet!'

The use of singular words to denote very large numbers can be considered as a distinction between marked and unmarked number. This appears to be a somewhat productive process. As discussed above, Lopit has a number of mass nouns which are either inherently singular or inherently plural. The word *hifiong*, 'water' is inherently plural, as shown in (11).

(11) ínyá hífióng   hùnáng l-è-líbà   à   lòmátàt.
     NEG water.PL this.F.PL VDM-3SG-good for drinking.PL

     'This water is not good for drinking.'

Sentence (12) shows it can be marked with the suffix *-i* which, given that *hifiong* is plural, yields a singular marking. This conveys the meaning of a very large amount of water or a flood.

(12) é-sàhí ìsàbìt nàbóìté lèfè h-o-ífuòt   Táfìfèrr à hìfióng-í.
     3SG-rain week one.F   until and-3SG-be.full Tafiferr like water-SG

     'It rained for a week and the Tafiferr plain was full of water (flooded).'

The word for cow is *hiteng* and the plural is *hisung*. It is possible to use the suffix *-i* on *hiteng* to denote a special, very large number. As the following sentence shows, the word *hitengi* takes singular agreement.

(13) á-wòló   náng hìténg-í nà   dè Kápuoítà.
     1SG-see.PFV 1SG.NOM cow-SG that.F.SG in Kapoeta

     'I saw so many cows in Kopoeta.' (lit. 'I saw that cow in Kopoeta.')

A further extension of the use of a singular lexeme to indicate greater singular meaning is found with the word *tohoni* 'person'. The usual plural for this lexeme is the suppletive *hiyo* 'people'. However, the singular *tohoni* 'person' can mean a very large number of people in a sentence such as (14).

(14) h-í-wòló   íyé   tòhòní nà   dè Naìróbí?
     Q-2SG-see.PFV 2SG.NOM person that.F.SG in Nairobi

     'Did you see all the people in Nairobi?' (lit. 'Did you see that person in Nairobi?')

The singular *tohoni* is not marked morphologically in (14). Rather, it is in semantic contrast to the expectation that a discussion about people in Nairobi would involve the plural.

The use of singular marking to indicate 'a large number' does not appear to be described in the literature. Corbett (2000: 234) describes a range of "special use" plurals. These special uses include exaggerative, intensificative or sensational plurals. These forms all involve the use of a plural or a reduplicated plural to add some kind of emphasis to a situation or activity. This is not the case in Lopit where the relevant noun has singular marking. The situation in Lopit could be regarded as somewhat similar to greater plural, but it is different in that it uses the marked singular.

6 Conclusion

Lopit has a three-way system for number marking involving plurative, singulative and replacement marking. The plurative pattern is the most common, comprising 58% of the sample. It is also quite complex with many different plural suffixes used. An experiment was carried out on loan and nonce words and revealed that there are some tendencies for plural suffix choice based on the syllable structure. With regard to singulative marking, the choice of suffix is more predictable with 75% of the singulative nouns in the sample forming a singular with *-i* after consonants and *-ti* after vowels. Person (or role) nouns follow a regular replacement pattern. For other nouns with replacement marking, the patterns are unpredictable.

The semantic nature of the tripartite system was examined. In common with other researchers, I found a relationship between the degree of individuation and singulative versus plurative patterns. Where there is a high level of individuation, such as with mammals and reptile, the plurative marking is used. Singulative marking is used for entities which typically occur in very large numbers or groups, such as stars or cereal grains. In between there are some finer distinctions. With insects, those insects which are small or found in large groups or swarms follow singulative marking, whereas insects seen more often as individuals follow plurative marking.

Replacement marking in Lopit is often used for entities which are not fully individuated (ears, arrows, knees, some vegetables). I have also observed that there is some variation amongst speakers on choosing plurative or singulative marking for some nouns that are not fully individuated (such as ribs, twins, sides). I suggested that replacement marking can diachronically develop as a result of variation within a community in how speakers perceive the boundary between individuation and grouping.

Lopit has a greater plural in addition to the normal plural. It also has special form of number marking whereby a singular affix can be used to denote very large numbers.

This form of number marking has not been observed in the literature and is here termed the *greater singular*.

The results presented here offer greater insight into the workings of Eastern Nilotic number marking, and illustrate that what may initially seem to be an irregular system is, in fact, influenced by various semantic and phonotactic factors. These findings, together with observations of phenomena like the *greater singular*, suggest that there remains much to learn about Nilo-Saharan number systems, and that detailed explorations across speakers and dialects will be of benefit in developing our understanding.

Abbreviations

ABS	absolute case	PFV	perfective
COP	copula	PL	plural
F	feminine	PRF	perfect
GPL	greater plural	Q	question maker
INF	infinitive	REL	relative clause marker
IPFV	imperfective	SBO	subordinator
NEG	negative	SG	singular
NOM	nominative case	VDM	verb displacement marker

Acknowledgements

References

Arber, H. A. 1936. A simple Lotuko grammar and Lotuko vocabulary. ms. Juba.

Billington, Rosey. 2014. Advanced tongue root in Lopit: Acoustic and ultrasound evidence. In Jennifer Hay & Emma Parnell (eds.), *Advanced tongue root in Lopit: Acoustic and ultrasound evidence*, 119–122. Christchurch: University of Canterbury.

Corbett, Greville. 2000. *Number* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

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*, 86–150. Cambridge: Cambridge University Press.

Dimmendaal, Gerrit J. 1983. *The Turkana language*. Dordrecht: Foris Publications.

Dimmendaal, Gerrit J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42(2). 214–261.

Driberg, John H. 1932. Lotuko dialects. *American Anthropologist* 34(4). 601–609.

Greenberg, Joseph H. 1981. Nilo-Saharan movable k as a stage III article. In William Leben (ed.), *Précis from the 12th Conference on African Linguistics, Stanford, April 10-12, 1981* (Studies in African Linguistics Supplement 8), 43–45. Los Angeles: University of California.

Grimm, Scott. 2012. *Number and individuation.* Palo Alto, CA: Stanford University. (Doctoral dissertation.)

Hilders, J. H. & J. C. D. Lawrance. 1957. *An introduction to the Ateso language.* 2nd edn. Kampala: Eagle Press, East African Literature Bureau.

Ladu, Valente O., Achaha S. Nartisio, Caesar O. Bong, Philip H. Odingo, John K. Fabiano & Paul A. Gilbert. 2015. *Lopit grammar book.* Juba: SIL-South Sudan.

Muratori, Carlo. 1938. *Grammatica Lotuxo.* Verona: Missioni Africane.

Stirtz, Timothy. 2014. *Phonological comparison of Lopit dialects.* Juba: SIL-South Sudan.

Tucker, Archibald N. & J. Tompo Ole Mpaayei. 1955. *A Maasai grammar: With vocabulary.* London: Longmans, Green.

Turner, Darryl. 2001. Lopit phonology: An outline of the phonology of Lopit words with a proposal for a tentative working orthography. Nairobi: SIL-Sudan.

Unseth, Peter. 1988. Majang nominal plurals, with comparative notes. *Studies in African Linguistics* 19(1). 75–91.

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

Chapter 22

Daats'íin, a newly identified undocumented language of western Ethiopia: A preliminary examination

Colleen Ahland

SIL International

Daats'íin is a heretofore unknown language spoken in western Ethiopia near the border with the Republic of Sudan. The Daats'íin people live in both Ethiopia and the Republic of Sudan but only those in Ethiopia still speak the Daats'íin language. The speakers of Daats'íin may number around 1,000 but may be as few as 300-500. This paper provides the first-ever overview of basic aspects of Daats'íin phonology, morphology and syntax. The overview documents that Daats'íin is structurally similar to the nearby Gumuz languages (of possible Nilo-Saharan affiliation) in many respects, including vocabulary, bound pronominals with a distinct tone for S versus A arguments, and incorporated nouns. However, there are a few differences, mainly in structure and certain tense-aspect categories of the verb word.

1 Introduction

Daats'íin is the autonym of a people group living in western Ethiopia and the southern part of the Republic of Sudan. The Daats'íin in Sudan have lost their traditional language but those in Ethiopia still speak it. Up until 2013, the language and people group were unknown to researchers and not included in the Ethiopian Census. I traveled to the area in 2014 in order to investigate the language and confirmed that Daats'íin (ISO dtn) is distinct from but closely related to Gumuz (ISO guk). I estimate that the Daats'íin likely number less than 1000 and that their language may be in danger of dying due to their population size and the heavy influence of Arabic and Amharic in the area.

Following is a first ever description and analysis of the Daats'íin language based on a word list of 400+ words,[1] five translated and annotated texts, and forty-four hours of targeted elicitation gathered in Ethiopia during a May-June 2014 field trip. The texts were gathered from Daats'íin adult native speakers of various ages, both male and female from

---

[1] These are the first 300 words of the SIL Comparative African Wordlist (Snider & Roberts 2006) plus words gleaned from texts which matched other items on the wordlist. Individual words were also collected for the Leipzig-Jakarta wordlist of the "least borrowable" words (Haspelmath & Tadmor 2009). See Appendix.

the village of Mahadid, Qwara wereda, Amhara Region. The word list and elicitation sessions were conducted with a male native speaker in his late 20's from Mahadid village. In §2 I present the locations of villages where Daats'íin is known to be spoken and give an overview of the similarities and differences between Daats'íin and Gumuz. In §3, I present a brief phonological analysis. In §4 I present the verbal morphology, in §5 describe the reciprocal construction, and in §6 various voice constructions. I then move on to noun morphology in §7 and noun modification in §8. In §9 I present demonstratives and in §10 describe prepositions and spatial relations. Lastly, in §11 I cover syntax. I end the grammatical description with a brief conclusion in §12.

2 Background

2.1 Location

In Ethiopia, the Daats'íin live in small villages within Mahadid K'ebele, northwest of Gelegu (formerly "Tewodros Ketema"), the capital of Qwara wereda, Amhara Region, near the edge of Alatish National Park. The Daats'íin also live in villages of Inashemsh (Emshemis?) K'ebele near Omidla in Guba wereda, Benishangul-Gumuz Region (Figure 1).[2] In Sudan, ethnic Daats'íin live in the villages of Ba'asinda and Gotihayaf.[3]

2.2 Relationship to Gumuz

Daats'íin and the Gumuz languages are very similar and would appear to be dialects of the same language if one merely inspected a wordlist (see Appendix). However, there exist quite a few differences grammatically, many of which explain the reported lack of mutual intelligibility between the languages.[4]

While verb roots in Daats'íin and Gumuz are mostly cognate, the verbal morphology is very different in some respects. For both Daats'íin and the Gumuz languages, verbs are often polysynthetic (1)-(2).[5] Much verbal morphology in these languages is presumed cognate, but certain cognate morphemes appear in different orders on the verb, such as the directional morpheme /-é/[6] (1) and incorporated prepositions like the Dative (DAT) (2).

---

[2] The location of Mahadid village is 12°17'56"N, 35°45'48"E. The location of Inashemsh is approximate, based on a description given by a Daats'íin language consultant.

[3] The village names (other than Mahadid) were given to me by a Daats'íin language consultant; I have not visited these locations to confirm whether ethnic Daats'íin live there.

[4] I witnessed Gumuz and Daats'íin neighbors using Arabic or Amharic to communicate with one another. In some cases, the Daats'íin have learned the Gumuz language. However, I know of no instances where a Gumuz has learned to speak Daats'íin. Both groups claim to not understand the other's language unless they have learned it.

[5] All language examples are written phonemically including the predictable epenthetic /a/ (in both Daats'íin and Gumuz), which also carries tone. Examples not written phonemically appear in phonetic brackets [ ].

[6] When the directional /-é/ TWRD occurs on a non-motion verb, it indicates that the speaker went to a location and came back (see also Ahland 2012).
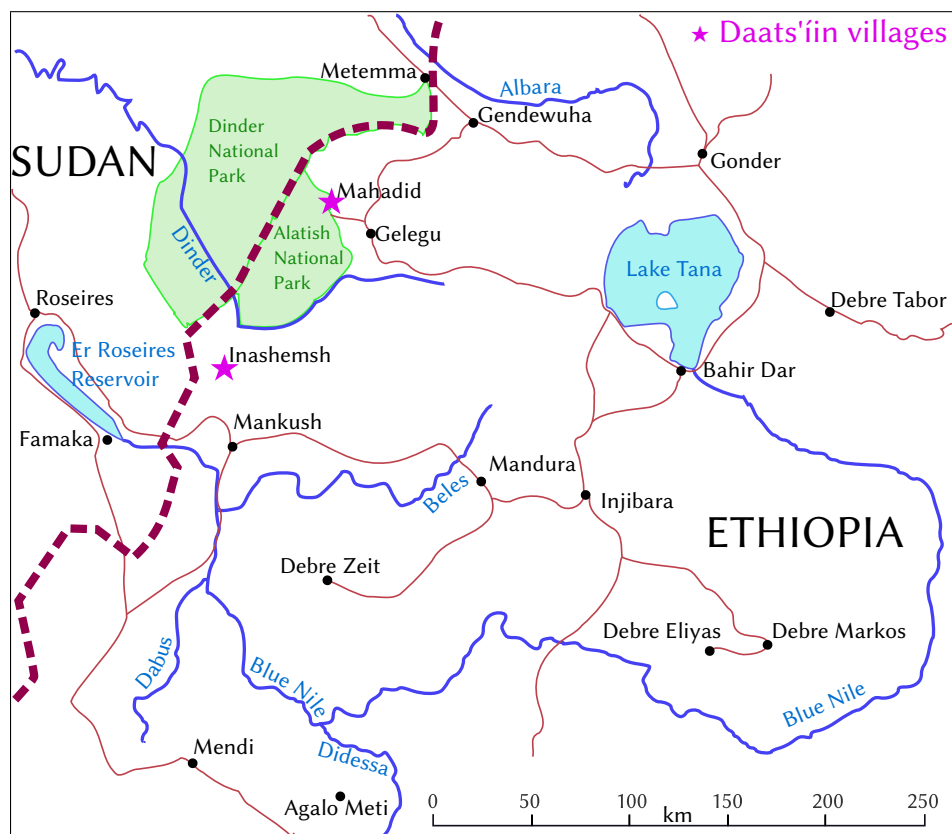
Figure 1: Daats'íin villages in Ethiopia. (Based on a map by SIL.)

(1) a. Daats'íin

ákwá   má-ba-ŋ-gáŋ-gám-ákwa-ts-é.

1PL.INCL IMP-1PL.INCL.IMP-VR-REDUP-know-1PL.INCL-BODY-TWRD

'We see each other over there.'

b. Southern Gumuz

ákwa   kâm-a-gam-ágw-é-ts.

1PL.INCL FUT-RECP-know-1PL.INCL-TWRD-BODY

'We will see each other over there.'

(2) a. Daats'íin

áljá   ká=járʔám k-ila-ca-ká.

1PL.EXCL DAT=3SG   AFF-1PL.EXCL.TR-give-DAT

'We gave to him.'

b. Southern Gumuz

áíla   b-íl-ká-cá-gá   ká=áŋ.

1PL.EXCL AFF-1PL.EXCL.TR-DAT-give-NFUT DAT=3SG

'We gave to him.'

Certain verbal morphemes only exist in Gumuz and not Daats'íin – e.g. the uncertainty prefix, perfect aspect suffix, nonfuture (NFUT) suffix (2b) (Ahland 2012) – and vice versa, i.e. a unique 1PL inclusive prefix *ba-* that co-occurs with the bound pronominal is found only in Daats'íin (1a). In addition, some verbal morphemes in Daats'íin and Gumuz that perform the same function are non-cognate, e.g. the (remote) past tense (/é-/ in Southern Gumuz and /-b/ in Daats'íin) (see §4.4), the affirmative prefix (/k-/ in Daats'íin as in (2a), but /b-/ in Southern Gumuz as in (2b), and /d-/ in Northern Gumuz), and the relativizer (§11). Also, the verbal plural (pluractional) is the result of a morphological process in Daats'íin (§4.7), but is expressed as the prefix /N-/ in the Gumuz languages.

Moreover, the Gumuz languages utilize two verbal templates based on tense: future (1b) versus nonfuture (2b); while the two Daats'íin verbal templates are based on aspect: imperfective (1a) versus perfective (2a). This would no doubt result in miscommunication as the 'future' form in Southern Gumuz (used mainly to express future time events) is similar to and presumed cognate with the 'imperfective' form in Daats'íin (which can be used to express a past event). The future template in Northern Gumuz utilizes a non-cognate future tense morpheme.[7]

Lastly, while verbs are often inflected with bound pronominals in Daats'íin, this inflection is not required. One can simply use the free pronoun without inflecting the verb for person. This is not true for either Northern or Southern Gumuz; verbs must be inflected for person.

Therefore, while Daats'íin and the Gumuz languages are very similar, there are several reasons to categorize these as distinct languages: lack of mutual intelligibility, distinct ethnolinguistic identity, and differing grammatical structures. These languages (Daats'íin, Northen Gumuz, Southern Gumuz) form their own subgroup within Nilo-Saharan or possibly stand as an independent isolate family, as proposed for Gumuz by Dimmendaal (2011). In what follows the discussion will primarily focus on Daats'íin as a hitherto undocumented language, but occasional comparisons with Gumuz will be noted.

3 Phonology

The Daats'íin phonological inventory includes 33 consonants (Table 1), and 5 vowels (Table 2) with some contrast in vowel length. Daats'íin has two tones: H (high) and L (low). In this paper, H tone is indicated with an acute accent mark while L tone is unmarked. Downstepped H tones are marked with ꜜ before the affected tone.

---

[7] Northern Gumuz has a less commonly used 'future' form which I presume is cognate to that of Southern Gumuz and Daats'íin. I tentatively label this form *immediate future* (Ahland 2012: 233).

3.1 Consonants

Daats'íin consonants span six places of articulation. Phonetically labialized consonants occur, but unlike for Gumuz these are not phonemic. Also noteworthy is the existence of a phonetic pharyngeal fricative. However, I do not consider it phonemic as the distribution is limited and predictable to some extent. Lastly, the language has both ejective and implosive stops.

Table 1: Daats'íin consonant phonemes

	Labial		Alveolar		Alveo-Palatal	Palatal		Velar		Glottal
Stop	p	b	t	d		c	ɟ	k	g	ʔ
glottalized	p'	ɓ	t'	ɗ		c'		k'		
Affricate			ts		tʃ					
glottalized			ts'		tʃ'					
Fricative	f	(v)	s	z		ʃ	(ʒ)			h
Nasal	m		n					ŋ		
Approximant										
lateral				l						
flap/trill				r						
Semi-vowel	w					j				

The palatal stop allophones [c, c', ɟ] of /c/, /c'/ and /ɟ/ are in free variation with palatalized velar stop allophones [kʲ, k'ʲ, gʲ], respectively. This is similar to Gumuz palatal stops. However, in most Gumuz dialects, the palatal stop pronunciation appears to be preferred over the palatalized velar stop pronunciation; in Daats'íin, the palatalized stop pronunciation seems to be preferred.

Marginal phonemes are in parentheses in Table 1. The consonants /v/ and /ʒ/ have highly limited distributions. The consonant /v/ is only found in one word thus far /vako/ 'bone', but exhibits contrast with similar segments in analogous environments.

(3) v   f   b
     /vako/ 'bone'   /fago/   'be drunk'   /bakaʔííba/ 's/he who had'
                      /k'ófaku/ 'navel'

Likewise, /ʒ/ is only found in the root /ʒiɟ/ 'sleep'. However, there is clear contrast with /z/ before a short /i/. With further data, it may be proven that the /ʒ/ in 'sleep' is simply an allophone of /z/ when the following consonant is palatalized or is a palatal stop. As discussed above, palatalized stops (e.g. [gʲ]) and palatal stops (e.g. [ɟ]) are in free variation.

(4) ʒ   z
     /ʒiɟ/ 'sleep'   /ázil-kʼo/ 'bend down, stoopʽ

The glottal stop /ʔ/ is realized as a voiced pharyngeal fricative [ʕ] following a sonorant and in the environment of the low vowel /a/, as in (5a). If a non-low vowel follows the glottal stop, the latter is realized as a phonetic glottal stop regardless of whether a sonorant precedes (5b). If no sonorant precedes, the glottal stop surfaces as a glottal stop [ʔ], regardless of the surrounding vowels (5c).

(5) a. [jarʕam] '3SG PRO'
     b. [jarʔii] 'be black'
     c. [baʔás] 'three'

3.2 Vowels

The vowel inventory of Daatsʼíin (Table 2) could be argued to consist of five vowels but with length contrasts realized by a difference in vowel quality. This is most noticeable with the realization of short /a/ as [ə]. Vowel quality differences are less pronounced (perhaps non-existent) for the mid-vowels /e/ and /o/. Thus, there appears to be no length contrast for the mid-vowels. Contrast between the long [i] and short [ɨ] high front vowels is similar to typical ATR contrasts found in many other African languages: [i] vs. [ɪ].

Table 2: Daatsʼíin vowel phonemes

	i		u	
		e	o	
			a	

The short round vowel /u/ is optionally realized as a labiovelar approximant followed by a high central vowel: [wɨ]. Table 3 shows posited length contrasts for the high vowels in all environments and between long and short /a/ word-medially. There is no known contrast for long and short /a/ word-initially, and only few examples word-finally. The phonetic realization of what I analyze as length contrasts are given in brackets.

While /o/ does not appear to have a contrast in length with short /o/ being realized as [wə], for example, the vowel often varies with labialization on a previous (back) consonant.

3.3 Tone

Daatsʼíin has two tones, high (H) and low (L), as represented in the near minimal pairs in (6) and (7). H tone is indicated with a acute mark ´ over the vowel, L tone is unmarked, and a HL sequence on a single vowel is indicated with a haceck ˆ over the vowel. Preliminary analysis suggests that Daatsʼíin also exhibits downstepped H tones. For example,

Table 3: Posited vowel length contrasts

	Short Vowel			Long Vowel		
/i/	/iˈl-kʼó/	[ɨlkʼó]	'head'	/íil/	[íl]	'belly'
/u/	/afútʃʼa/	[afwɨtʃʼa]	'blow'	/fúúna/	[fúnna]*a*	'smell'
/a/	/ham/	[həm]	'gall blad-der'	/ahááma/	[aháma]	'yawn'
	/baʃok/	[bəʃok]	'porcupine'	/mbaaʃia/	[m̩baʃija]	'hide (v.tr)'

---

a In this particular word, long /uu/ is realized as a short [u] followed by a lengthened nasal.

the final L tone of the HL pattern on intransitive (S)[8] bound pronominals (Table 4 in §4.2) induces a downstepped H tone on a following H-initial morpheme (see Table 5 in §4.2).

(6)
	LL	HH
	/giʃa/	/gíʃí/
	'stone(s)'	'see'

(7)
	LHL	LH
	/afága/	/fagá/
	'urinate'	'grow (up)'

4 Verb morphology

Daats'íin verbs are somewhat polysynthetic with up to eleven position classes (and possibly more). While both prefixes and suffixes are possible, verbs are more heavily suffixing. There exist two templates for Daats'íin verbs: one for perfective verbs (Figure 2) and one for imperfective verbs (Figure 3).

One of the main differences between the two templates is the position of the S/A bound pronominal. For the perfective template, the S/A bound pronominal is a prefix to the verb root (8). For the imperfective template, the pronominal is a suffix (9) save for the 1ˢᵗ person plural inclusive, which is doubly marked in the imperfective with both a prefix and a suffix.

(8)   k-**íl**-dugwa.
       aff-**1pl.excl.intr**-run
       'We ran.'

---

[8] I use the notation S, A, and P in the vein of Comrie (1989): S is the single core argument in an intransitive subject function, A is the argument of a transitive clause that correlates most closely with the notion of *Agent*, and P is the argument of a transitive clause that correlates most closely with the notion of *Patient*.

Mood	Person (S/A)	Valence Decreaser	Main Verb Root	greater plural	Past	Middle Voice	Incorporated preposition	Person (object of preposition)	Incorporated Noun	Directional
-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7
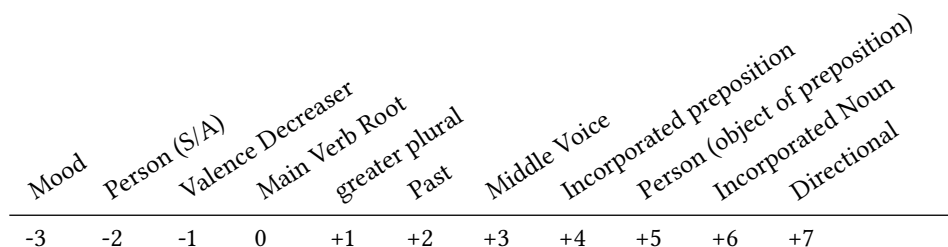
Figure 2: Position class chart for Daats'íin perfective verbs

Imperfective	1PL Inclusive	Main Verb Root	Past	greater plural	Person (S/A)	Incorporated preposition	Directional	Incorporated Noun
-2	-1	0	+1	+2	+3	+4	+5	+6
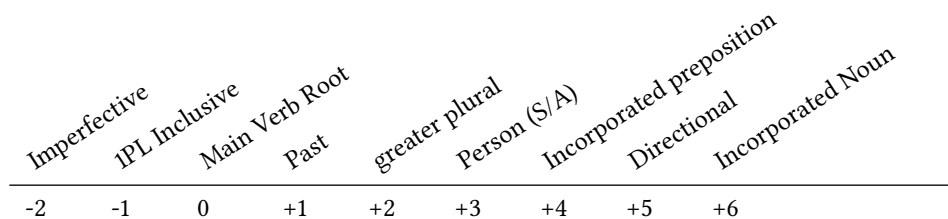
Figure 3: Position class chart for Daats'íin imperfective verbs

(9) má-dugwa-**íla**.   [mədugúla]
     IMP-run-**1PL.EXCL.INTR**

     'We run.'

Also noteworthy is that perfective is only indicated by the choice of verbal template itself; unlike for the imperfective, there is no dedicated perfective morpheme.

4.1 Complex verb stems

Daats'íin verbs often have simple roots accompanied by an historically incorporated noun (IN), which together form a complex verb stem (10).[9] This complex stem is more readily identified in the middle voice (11), in the past tense (12), or imperfective (13), where other morphemes intervene between the verb root and the incorporated noun.

(10) ɓaga   k-a-**koɗ-ás**   nkudida.
      person AFF-3SG.TR-**open-MOUTH** door

      'The person opened the door.'

---

[9] The structure of INs needs further investigation. For the most part, INs ending in L-tone /a/ in their free forms do not end in /a/ when incorporated. However, in passive and middle formations, the final /a/ is maintained, e.g. *-sa* in (11). Also, an /a/ vowel is epenthesized when a consonant-final morpheme precedes a consonant-initial IN. The epenthesized /a/ carries H tone if the preceding and following syllables are (underlyingly) L (10), with some exceptions (12). While the epenthesized /a/ and H tone appear to be somewhat predictable, I have represented them as part of the IN (i.e. *-ás* as a variant of *-sa* 'MOUTH').

(11) ʃeeʃokwá k-a-**kóɗ**-áá-**sa**.
doorway AFF-3SG.TR-**open**-MV-**MOUTH**

'The doorway opened.'

(12) ná=gats'ar   k-u-**gám**-b-**áts**   jáhú.
LOC=old.days AFF-3PL.TR-**know**-PST-**BODY** reedbuck

'In the old days, they saw reedbuck.'

(13) má-**si**-áɗá-**ííl**   alkubáj.
IMP-**wash**-1SG.TR-**BELLY** cup

'I wash the inside of the cup.'

The list of grammaticalized INs in Daats'íin is similar to that of Gumuz (semantically and structurally): /-k'ó/ 'HEAD', /-cé/ 'EYE', /-ííl(a)/ 'BELLY', /-ts(a)/ 'BODY', /-s(a)/ 'MOUTH', and /-f(a)/ 'HIP/LOINS'. Only one of these incorporated nouns does not exist as a free noun form: /-f(a)/ 'HIP/LOINS'.[10] Many of these INs clearly function as parts of wholes in an external possession (EP) construction (Payne & Barshi 1999). That is, these incorporated body part nouns often serve as the possessed 'part' of the P (object) or S (intransitive subject) argument, while the P/S argument serves as the possessor. By incorporating different body part nouns into the verb 'wash', for example, one can wash the inside of the cup (13) versus the (out)side of the cup (14).

(14) má-si-áɗá-**ts**   alkubáj.
IMP-wash-1SG.TR-**BODY** cup

'I wash the (out)side of the cup.'

Beyond the body part nouns listed above (which have commonly grammaticalized – and lexicalized – as part of complex verb stems), other body part nouns are also commonly incorporated (15), (16).[11]

(15) ɗa-ɗú-áá-**ee**.
1PL.TR-be.sick-MV-**arm**

'My arm hurts.'

(16) íi-**tʃugwa**.
be-**foot**

'Stand.'

---

[10] I have inferred this meaning of the IN /-fa/ 'HIP/LOINS' from its free cognate form in Southern Gumuz /ʃa/ (perhaps /keef/ 'loins' in Daats'íin is historically related.) The IN has grammaticalized to mean something akin to the direction 'down' or has lexicalized with verb roots like 'know' /gam-áf/ referring to cognitive function.

[11] These incorporated body part morphemes sometimes function in a more grammatical way and sometimes in a more lexical way. The grammaticalization versus lexicalization issues of one and the same morpheme must be saved for other research.

Finally, one noun that is not a body part can also be incorporated: /-gw(a)/ 'place' (17).[12]

(17) k-â-ɗá⁺b-**ágw**-é?
      AFF-3SG.INTR-find-**PLACE**-TWRD
      'Did he arrive?'

4.2 Bound pronominals

Daats'íin 1ˢᵗ and 2ⁿᵈ person bound pronominals are similar to their free forms (Table 4). The bound forms are somewhat reminiscent of an ergative pattern in that there exists a distinct difference in tonal melody for the S bound pronominals versus the A bound pronominals.[13] In general, S bound pronominals follow an H(L) pattern (17), (19) while A bound pronominals follow an L pattern (18), (20).[14] Again, these bound pronominals are prefixes in the perfective template and suffixes in the imperfective template. Table 4 provides all the free[15] and (verbal) bound pronominal forms, save the additional 1ˢᵗ plural inclusive bound form /ba-/ for the imperfective template, which co-occurs with the /-akwa/ pronominal (20).

(18) k-**aɗa**-ráfaʔ   jarʔam.
      AFF-**1SG.TR**-carry 3SG
      'I carried him.'

(19) k-**áɗa**-dugwa.
      AFF-**1SG.INTR**-run
      'I ran.'

(20) mâ-**ba**-sa-**akwa**.
      IMP-**1PL.INCL.IMP**-eat-**1PL.INCL.TR**
      'We eat.'

Daats'íin does not require a bound pronominal on the verb; it is perfectly acceptable to use a free pronoun with no bound pronominal. In such instances, the morpheme /áá-/ (INTR) is used for an S argument (21) and /a-/ (TR) for an A argument (22) in the position where the bound pronominal would occur.

---

[12] The IN /-gw(a)/ 'place' often reduces the valence of an otherwise transitive verb root. In example (17) /-gw(a)/ has lexicalized with the verb root resulting in a distinct intransitive verb, and in Table 10 (§8.3), it functions as an antipassive.

[13] The pattern is also nominative in that both the S and A pronominals can be marked on the verb and occupy the same position on the verb, whereas the P argument is not marked on the verb.

[14] There exist some variations in this pattern, which need further investigation.

[15] Tone on free pronouns needs further investigation. Tone variation appears to indicate a split system, suggesting that 1ˢᵗ person exhibits an ergative-absolutive pattern and 3ʳᵈ person a nominative accusative pattern. However, evidence is inconclusive (Kelly 2014), and just one tone form for free pronouns is given in Table 4.

Table 4: Daats'íin bound and free pronominals

| | BOUND | | FREE |
	S	A	
1SG	*áɗa*	*aɗa*	*áɗa*
2SG	*áa*	*aa*	*ám(am)*
3SG	*â*	*a*	*jarʔam*
1PL.INCL	*ákwa*	*akwa*	*ákwá*
1PL.EXCL	*íla*	*ila*	*álja*[a]
2PL	*áca*	*aca*	*áca*
3PL	*úúa*	*uua*	*máʔam*

[a] This pronoun is also pronounced [áíla] with variations in tone.

(21) áɗa k-áá-bo.
      1SG AFF-INTR-move.away
      'I moved.'

(22) áɗá k-a-rábiis-é   nágát.[16]
      1SG AFF-TR-raise-TWRD there
      'I raised him/her over there (and came back).'

4.3 Greater plural

The greater plural (GP) morpheme /-óa/ in Daats'íin is reserved for animate entities and tends to mark numbers that are unknown due to the high quantity.[17] The greater plural is often used in combination with bound pronominals. The plural number indicated in the bound pronominals generally refers to a known number, which is often 2 to 4 participants (or 'paucal'). When a plural bound pronominal combines with the greater plural verb suffix, the number of participants is understood as vast. When combining these morphemes in the 3rd person, one can indicate up to three degrees of plurality: 'some', 'many' and 'very many' (Table 5). While the greater plural can be marked on any 3rd person verb form, it can only be marked on the plural verb forms for 1st and 2nd persons.

---

[16] The verb stem /rábiis/ 'raise' may prove to be complex (verb root + incorporated noun) with further investigation.

[17] The greater plural in Daats'íin is cognate with the greater plural in Northern Gumuz. The term *greater plural* as used here is similar to that in Corbett (2000); the difference is that greater plural is marked on the verb in Daats'íin whereas Corbett (2000) uses the term to describe marking on a noun.

Table 5: 3ʳᵈ person number marking on verbs in Daats'íin

Daats'íin	Gloss
k-úú⁺-d̶-óa AFF-3PL.INTR-go-GP	'very many went'
k-á⁺-d̶-óa AFF-3SG.INTR-go-GP	'many went'
k-úú⁺-d̶á AFF-3PL.INTR-go	'some went'
k-á⁺-d̶á AFF-3SG.INTR-go	's/he went'

4.4 Past tense

Daats'íin has a past tense marker /-b/, which can occur in either the perfective (23) or imperfective (24) aspect.

(23) ná=gats'ár ɓaga bá k-a-gám-b-átsa   jáhú.
     LOC=old.days person PROX AFF-3SG.TR-know-PST-BODY reedbuck

     'In the old days, this person saw reedbucks.'

(24) ná=gats'ár   ɓaga má-sa-b-uuá   gar.
     LOC=old.days people IMP-eat-PST-3PL.TR porridge

     'In the old days, people would eat porridge.'

4.5 Incorporated prepositions

The prepositional proclitics /ká=/ 'to, for' (benefactive/dative; DAT) and /ka=/ 'with' (instrumental/comitative; INST) can be incorporated into the verb. Unlike traditional "applicatives" (Payne 1997), these incorporated prepositions do not "promote" an oblique to object grammatical status. Rather, the incorporated prepositions index an oblique referent on the verb; the external object of the preposition remains as such if it is expressed lexically, even if the incorporated preposition also occurs. Compare (25) and (26).

(25) má-nsam-d̶a-**ká**-ts   **ká**=má?am.
     IMP-speak-1SG.TR-**DAT**-BODY **DAT**=3PL

     'I tell them.'

(26) má-si-ád̶a-**ka**-tsa   **ka**=a?e.
     IMP-wash-1SG.INTR-**INSTR**-BODY **INSTR**=water

     'I bathe with water.'

In addition, the 1SG (27) and 2SG (28) pronouns can be indexed on the verb even while simultaneously expressed external to the verb as objects of a preposition. However, plural pronominals 1PL exclusive (29), 2PL (30), and 3PL (25) are not indexed on the verb when expressed as objects of a preposition.

(27) máʕám ká=áɗa k-uu-ca-ká-**áɗ**-é.
     3PL   DAT=1SG AFF-3PL.TR-give-DAT-**1SG**-TWRD
     'They gave (it) to me.'

(28) má-nsam-ɗa-ká-**áa**-tsa.
     IMP-speak-1SG.TR-DAT-**2SG**-BODY
     'I tell you.'

(29) jarʔám k-a-c-é   ká=álja.
     3SG   AFF-3SG.TR-give-TWRD DAT=1PL.EXCL
     'S/He gave to us.'

(30) má-nsam-ɗa-ká-ts   ká=áca.
     IMP-speak-1SG.TR-DAT-BODY DAT=2PL
     'I tell you all.'

4.6 Directional 'toward'

Daats'íin has a directional suffix /-é/ meaning 'toward', but no known directional meaning 'away'. The 'toward' directional can indicate: 1) action directed toward the deictic point of reference (which is normally the speaker), 2) action performed elsewhere and returning to the deictic point of reference, and 3) perfect aspect.

4.6.1 Action directed toward speaker

For motion verbs, if an entity is performing an action in the direction of the speaker, the /-é/ suffix is added (31)-(33); a motion verb unmarked for direction is typically understood as involving motion away from the speaker (34).

(31) du-ʔám   k-á-wá-é.
     child-3SG.POSS AFF-3SG.INTR-go-TWRD
     'His/her child came.'

(32) ʃeeʃokwá k-a-kó꜔ɗ-áá-s-é.
     doorway AFF-3SG.TR-open-MV-MOUTH-TWRD
     'The door opened toward me. '

(33) aíla   da=k-íl-boo-b-é   nagá cét   hogí.
     1PL.EXCL TEMP=AFF-1PL.EXCL.INTR-move-PST-TWRD here IDEO:all forest
     'When we moved here, it was all forest.'

(34) wá ká=eb-úʔ.
      go DAT=homeland-2SG.POSS
      'Go home. '

Similarly, ditransitive verbs like 'give', in which the THEME is set in motion, can also take the 'toward' directional, whether the movement is literal (27), (29) or metaphorical (35). While most instances of the directional involve motion directed toward the speaker, the meaning in example (35) involves motion directed toward a third party, presumably as the deictic reference point.

(35) ámám baʔadeén ká=jarʕám nsam-ká-ts-é.
      2SG   later   DAT=3SG   speak-DAT-BODY-TWRD
      'Tell him something later.'

4.6.2 Action performed in different location from the deictic reference point

The 'toward' directional can occur on the verb to indicate that an action was performed in a different location from the deictic reference point (normally the speaker), but then with a return to the reference point (see example 22 above). This is generally true for non-motion verbs in Daats'íin.

4.6.3 Perfect aspect

Lastly, addition of the 'toward' directional can indicate that the action was completed but retains present relevance to the speaker and the addressee. This function of the directional I refer to as 'perfect' (36)-(37).[18]

(36) k-aɗa-sá-é.
      AFF-1SG.TR-eat-TWRD
      'I have eaten. '

(37) albúna k-uu-fa-e.
      coffee AFF-3PL.IMPL-drink-TWRD
      'The coffee has been drunk.'

4.7 Verbal number (verbal plural)

According to Corbett (2000: 246-249), VERBAL NUMBER relates to events (as opposed to ENTITY NUMBER[19] which refers to nominal plurality), and can be further subdivided into EVENT NUMBER and PARTICIPANT NUMBER. EVENT NUMBER encodes whether an event

---

[18] It is not known why the verb root 'drink' /fá/ and the directional /-é/ surface with L tone in example (37). It may be that L tone is part of the 3PL Impersonal construction, outlined in §6.1.

[19] Entity number, according to Corbett (2000), is typically number encoded on nouns but can also be encoded on verbs via bound pronominals.

took place more than once, or in more than one location, which for many African languages is referred to as *pluractional* (Corbett 2000: 243). PARTICIPANT NUMBER refers to the number of participants in an event.[20] For example, a special verb form is used in some languages for one or two participants, versus many. This preliminary analysis reveals that, regarding verbal number, Daats'íin utilizes the same construction to encode EVENT NUMBER and PARTICIPANT NUMBER.[21] The verbal plural is expressed via partial ((V)C$_1$V) reduplication of the simple verb root.[22] If an event is completed *en masse* or if the action of each individual is not emphasized, the verb is not reduplicated; then only entity number (e.g.3PL) is encoded on the verb (38). In (39), by contrast, there are several drinking events and the emphasis is on the amount of coffee that was consumed. Example (40) refers to only one event, whereas (41) refers to several breaking events.

(38) ɓaga deá k-uu-fá albún.
people PROX.PL AFF-3PL.TR-drink coffee
'These people drank coffee.'

(39) ɓaga deá k-uu-fá∼fá albún.
people PROX.PL AFF-3PL.TR-PL∼drink coffee
'These people drank (a lot of) coffee.'

(40) giá k-á-n-ˈts'áda.
wood AFF-3SG.INTR-VR-break
'The wood broke.'

(41) giá k-á-nts'á∼n-ts'áda.
wood AFF-3SG.INTR-PL∼VR-break
'The wood broke into pieces.'

If the verb stem is complex, only the simple verb root is partially reduplicated and not the incorporated noun (42).

(42) k-íl-gá∼gam-b-áts áíl ííjá bac' ná=hógiá baʔ.
AFF-1PL.EXCL.TR-PL∼know-PST-BODY 1PL.EXCL every meat LOC=forest PROX
'We saw every animal in this forest.'

---

[20] Examples of 'participant number' in Corbett (2000: 247-8) appear to involve S/P arguments which are semantically PATIENTS or THEMES.

[21] In Daats'íin, there appears to be no distinction made between EVENT NUMBER and PARTICIPANT NUMBER; all are interpreted as plural events. However, the verbal plural construction is not always required when there are multiple participants (S/P PATIENT/THEMES) unless each individual event is emphasized. For this reason, one could analyze all instances of the verbal plural in Daats'íin as encoding/emphasizing multiple events, regardless of the number of participants.

[22] The valence reducer /n-/ is included in the partial reduplication (41).

4.8 Negative enclitic

The negative enclitic /=cé/ [kʲé] in Daats'íin (43) attaches to the end of the verb word and appears to be cognate with *=cê* of the Sirba Abay dialect of Southern Gumuz (44) (cf. Ahland 2012: 241-242).

(43) Daats'íin
il-tʃ'a=c   da.
1PL.EXCL.TR-have=NEG thing
'We (EXCL) don't have anything.'

(44) Southern Gumuz (Sirba Abay)
a-tʃ'á-gá=cê   lamáána.
3SG.TR-have-NFUT=NEG money
'S/he doesn't have money.'

5 Reciprocal

In order to express a reciprocal meaning in Daats'íin, one uses a transitive verb stem in the verbal plural construction (§4.7) with a plural S argument (which is doubly marked on the verb in the 1ˢᵗ plural inclusive) (45), (46). As the construction is structurally intransitive, an intransitive tonal pattern is used on the bound pronominal (see Table 4). Additionally, the construction involves a valence reducer /N-/, which is phonetically undetectable before a nasal, e.g. (45).

(45) ákwa ná=nɟertáát má-ba-ná∼n-nás-ákwa.
1PL   LOC=before IMP-1PL.INCL.IMP-PL∼VR-talk-1PL.INCL.INTR
'We were talking to each other earlier.'

(46) ákwá má-ba-ŋgá∼ŋ-gám-ákwa-tsa.
1PL   IMP-1PL.INCL.IMP-PL∼VR-know-1PL.INCL.INTR-BODY
'We see each other. '

6 Voice

Daats'íin has three voice constructions: active, passive and middle. Active voice is unmarked. Passive voice is expressed via an impersonal 3PL verbal construction. Lastly, middle voice (MV) has a distinct construction which involves the suffix /-aa/.[23] Passive and middle voices are discussed below and compared with active voice.

---

[23] This MV suffix can also surface with H tone, as in (11), (15), (32). H tone may be the underlying tone with an L tone added to valence-reducing voice constructions; the hypothesized L tone surfaces on the MV suffix itself and/or the final vowel of the IN.

6.1 Passive (3PL impersonal)

Passive voice in Daats'íin is generally expressed via the 3rd person plural impersonal (3PL.IMPL) construction. Cross-linguistically, impersonal 3rd plural subject marking differs from non-impersonal 3rd plural subject marking in two ways: 1) the impersonal construction lacks an overt antecedent in the preceding discourse; and 2) the impersonal construction is typically a phonologically (or morpho-phonologically) reduced form of the 3rd plural anaphoric form (Siewierska 2010: 75). The 3rd plural bound pronominal in the impersonal construction is non-referential in Daats'íin, but is identical in form to referential 3rd plural A argument marking. However, what distinguishes the active and passive constructions (structurally) is the position of the lexical S/P argument: in a passive construction the lexical S argument precedes the verb (47); but in the corresponding active transitive construction, the P argument follows (48).[24] Furthermore, the 3rd plural active transitive construction must have a plural agent referent; whereas in the impersonal passive, the unknown agent could be singular or plural.

(47) ɓaga   k-uu-ʃa-k'o.
      person AFF-3PL.IMPL-die-HEAD

      'A person was killed.' (The person who did it is unknown.)

(48) k-uu-ʃa-k'o   ɓaga.
      AFF-3PL.TR-die-HEAD person

      'They killed a person.'

6.2 Middle voice

Daats'íin has a construction in which the suffix -*aa* plus a transitive (complex) verb stem yields an overall intransitive clause. I refer to this as the *middle voice* (MV) construction (cf. Klaiman 1991: 44-45; Kemmer 1993: 3-4; Dixon & Aikhenvald 2000: 12; Givón 2001: 116-121). Following Kemmer (1993: 3), I use the term *middle voice* to refer to constructions that are semantically intermediate in transitivity. The Daats'íin MV construction involves verbs that depict agent-initiated actions to construe a change, potential state or resulting state of a PATIENT (cf. Givón 2001: 116). In Daats'íin, the MV construction is structurally intermediate in transitivity in that the construction has only one argument but the tonal marking of the bound pronominal on the verb indicates that the construction is transitive. Furthermore, though the verb stem is transitive, the construction does not allow expression of the AGENT in an agent phrase or otherwise. Unlike the 3rd plural impersonal construction (§6.1), a non-referential 3PL A prefix does not occur on the verb, but instead there is a prefix which agrees in person and number with the lexical S argument that precedes the verb. In addition, the Daats'íin MV construction requires an incorporated noun; the MV suffix -*aa* cannot occur on a verb stem consisting of only a

---

[24] Here I label the single argument of the 3rd plural passive as S. However, the argument also has characteristics of a transitive P argument in that the 3PL bound pronominal takes the form of an A argument.

simple verb root. Compare the active transitive voice construction in (49), in which the A argument 'fire' burned the grass, with the middle voice construction in (50) where the S argument 'grass' is the semantic PATIENT.[25]

(49) toa k-a-sa-kʼó   amfaɗea.
      fire AFF-3SG.TR-eat-HEAD grass

      'The fire (completely) burned the grass.'

(50) amfaɗeá k-a-s-aa-kʼo.
      grass   AFF-3SG.TR-eat-**MV**-HEAD

      'The grass burned.'

7 Noun morphology

Daatsʼíin exhibits minimal nominal morphology. Simple nouns can be inflected for number (via prefixation or other morphological processes) and can take a bound pronominal expressing a possessor.

7.1 Nominal number

Most nouns in Daatsʼíin are not specified for number. The language displays what Corbett (2000) calls "general number" and others have labeled as "transnumeral" (Biermann 1982; Storch & Dimmendaal 2014), in that the noun in its unmarked form can be interpreted as either general/plural or singular. Only nouns higher on the animacy hierarchy (e.g. humans, animals) can be explicitly marked for plural, most commonly with the prefix /má-/ (Table 6, Set A). A few nouns referring to humans and animals form a plural via a morphological process (Table 6, Set B). Finally, only one noun is known to have suppletive singular and plural forms (Table 6, Set C).

Set B in Table 6 is formed via the following morphological process: $c_1v_1(c_2)(v_2) \rightarrow c_1\acute{a}\acute{a}c_1\grave{v}_1(c_2)(v_2)$ (where $\grave{v}$ is a vowel that carries L tone). In addition, /o/ and /i/ in the v_1 position tend to weaken to labialization and palatalization, respectively, when following a back consonant. Therefore, v_1 in 'guest' and 'lion' is expressed as part of c_1 when /áá/ follows.

7.2 Bound possessive pronominals

Nouns in Daatsʼíin can be inflected with the bound possessive pronominal suffixes listed in Table 7. These pronominals are related to the free pronoun forms (Table 4), save for the 2nd person singular.

---

[25] It is also feasible to apply the term 'anticausative' (cf. Dixon & Aikhenvald 2000: 7) to what I call the MV construction.

Table 6: Nominal plural strategies

		Singular	Plural
Set A	'woman'	*gáf*	*má-gáf*
	'man'	*gwinzá*	*má-gwinzá*
	'mother'	*jaajó*	*má-jaajó*
	'older brother'	*aʔé*	*má-aʔé*
	'older sister'	*dadó*	*má-dadó*
Set B	'guest'	*kodar*	*kwáákodar*
	'king'	*t'ís*	*t'áát'is*
	'young man'	*síbi*	*sáásibi*
	'bird'	*mété*	*máámete*
	'reedbuck'	*jahu*	*jáájahú*
	'lion'	*hii*	*hjááhii*
	'mouse'	*bu*	*báábu*
Set C	'child'	*du*	*dííd*

Table 7: Bound possessive pronominals

	Singular	Plural	
		INCL	EXCL
1	*-máɗa*	*-ákwa*	*-múlja*
2	*-ʔú*	*-áca*	
3	*-ʔám*	*-máʔám*	

For noun roots that are consonant final, a short /a/ is epenthesized between the final consonant and any bound pronominal that is consonant initial (51).[26]

(51) rus-ʔú [rusaʔú] k-á-dugwa.
      cow-2SG.POSS   AFF-3SG.INTR-run

      'Your cow ran.'

---

[26] In Daats'íin, noun roots in citation form are typically uttered without a final /a/. However, a final /a/ is often added with varying tones (which together I suspect is marking case when the noun occurs as part of certain syntactic constructions. It could be that (51) does not involve an epenthesized /a/ but rather the final /a/ is added because the word is uttered as part of a sentence.

8 Noun modification

Daats'íin has two related constructions for when one noun modifies another: the associative construction (§8.1) and the attributive construction (§8.2). The two noun-noun (NN) constructions differ in order of head and modifying noun. Both constructions require an epenthesized /-a/ after the first noun if the first noun is consonant-final (in citation form). Two other related constructions are the relator noun construction (§10.2) and pronominal nominalizations. The last is formed with the pronouns /etá-/ 'SG.HUM', /dáá-/ 'PL.HUM', and /dá-/ (nonhuman) as the first "noun" of the construction (§8.3).

8.1 Associative construction

The associative construction is a NN construction in which the second noun modifies the first. The Daats'íin associative construction is semantically similar to other similarly-named constructions found across Africa (Welmers 1974: 275-276). The semantics of this construction include possession (mainly parts of wholes, material, contents, and function/purpose).

When a L tone noun modifies a L tone noun in the associative construction, a H tone is suffixed to the modified noun (first noun of the construction) and an /a/ is epenthesized if the first noun is consonant-final (Table 8). Tonal behavior of non-L tone nouns in this construction is yet to be determined.

Table 8: L tone nouns in the associative construction

L tone nouns	Associative construction
batʃ 'meat' + *rus* 'cow'	*batʃárus* 'cow meat'
tʃugw 'foot' + *hii* 'lion'	*tʃugóhii* 'lion leg'[a]
tʃak'o 'house' + *mfaɗe* 'grass'	*tʃak'ómfaɗe* 'grass hut'

[a] Regarding the alternating pronunciation [tʃugw] / [tʃugo-] 'foot', see §3.2 above.

8.2 Attributive construction

The attributive construction, like the associative construction, is comprised of two nouns which together may be pronounced as a single phonological word. However, in the attributive construction the first noun is the modifier and it is typically a deverbal noun (52).[27]

[27] Deverbal nouns in Daats'íin are typically formed with the nominalizers /ma-/ and /ga-/. The /ma-/ nominalizer yields a form that is more verb-like, in that the form is also used for infinitive-like constructions.

(52) ma-ráɗa-ɓag   k-á-w-é.
     NMLZ-be.tall-person AFF-3SG.INTR-go.away-TWRD
     'The tall person came.'

In order to pluralize a NN attributive construct, the first (modifying) noun is typically a nominalized verbal plural form of a stative verb (in bold), and the second noun, if animate, takes the plural form it would typically take outside of the construction. Compare the singular (53a) and plural (53b) forms of 'beautiful horse'. The plural form, as opposed to singular, does not appear to take the additional /-a/ on the nominalized noun typical of NN constructions, nor do the two nouns appear to be phonologically bound.

(53) a. bá   **ma-ʔarásá**-marta.
         PROX NMLZ-be.beautiful-horse

         'This is a beautiful horse.'

      b. deá   **ma-ʔá-ʔaras**   má-marta.
         PROX.PL NMLZ-REDUP-be.beautiful PL-horse

         'These are beautiful horses.'

The attributive construction is structurally similar to relator (inherently possessed) nouns of the associative construction in that if the second noun of the construction is not expressed, the 3ʳᵈ singular possessive bound pronominal, i.e. the inherent possession (IP) marker, is used in its place (see also §10.2). In the associative construction, a similar phenomenon occurs but any bound pronominal is optional. Table 9 compares these three NN constructions and with the 3SG POSSESSIVE/IP.

Table 9: NN Constructions with 3SG possessor/inherent possession

	NN construction	N + 3SG.POSS/IP
Associative construction	*tʃugó-hii*	*tʃugó-ʔám*
	leg-lion	leg-3SG.POSS
	'lion leg'	'its leg'
Relator noun construction	*i´íl-tʃok'o*	*íil-ʔám*
	belly-house	belly-IP
	'inside (the) house'	'inside (of it)'
Attributive construction	*ga-ʃé-ɓag*	*ga-ʃe-ʔam*
	NMLZ2-be.good-person	NMLZ2-be.good-IP
	'good person'	'good (one/thing)'

Beyond nominalized stative verbs, at least two other noun-like words can serve as the first noun in the attributive construction: 'big' (54)-(55) and 'small' (56).[28] Because

---

[28] These are considered attributive constructions because the semantic head is the second noun, e.g. *babúʃá-kambo* 'big-camel'.

these are not deverbal nouns, they are pluralized with the /má-/ prefix (55) instead of the reduplicated form found in the nominalized verbal plural (see (53) above).

(54) kaambo bá   babúʃa-ʔam.
      camel   PROX big-IP
      'This camel is big.'

(55) kaambo deá   má-babúʃa-ʔam.
      camel   PROX.PL PL-big-IP
      'These camels are big.'

(56) kaambo bá   duʃíʃa-ʔam.
      camel   PROX small-IP
      'This camel is small.'

8.3 Pronominal NN Constructions

Pronominal NN constructions have the bound pronominal forms /etá-/ 'SG.HUM', /dáá-/ 'PL.HUM', and /dá-/ (nonhuman) as the first "noun" of the construction and either a non-derived (57) or derived noun (Table 10) as the second noun of the construction. This construction type is often used for professions (57a) or for general names of things or people (as in clan names, e.g. *Daa-tsʼíin*). Those expressions with derived nouns as the second noun often function as nominalized relative clauses or, better yet, as nouns with participial modifiers.

(57) a. áɗá etá-maj.
          1SG SG.HUM-field
          'I am a farmer.'
      b. álja   dáá-maj.
          1PL.EXCL PL.HUM-field
          'We are farmers.'

Table 10: Pronominal NN constructions with derived nouns

	Nonhuman	Human singular	Human plural
N > 'sit'	*dá-ꜜm-ʔíí-f*	*etá-m-ʔíí-f*	*dáá-ꜜmá-ʔíí-f*
	NHUM-NMLZ-be-HIP	SG.HUM-NMLZ-be-HIP	PL.HUM-NMLZ-be-HIP
	'chair'	'a person who sits'	'people who sit'
N > 'eat'	*dá-m-ꜜsá-gó*	*etá-m-ꜜsá-gó*	*dáá-ꜜmá-sá-gó*
	NHUM-NMLZ-eat-PLACE	SG.HUM-NMLZ-eat-PLACE	PL.HUM-NMLZ-eat-PLACE
	'food'	'a person who eats'	'people who eat'

9 Demonstratives

Daats'íin demonstratives function as both modifiers (54)-(56) and pronouns (53). The demonstratives appear to exhibit three degrees of distance: proximal, mid-distal, and distal (Table 11). For the mid-distal and distal singular demonstratives there exists some variation. The analysis in Table 11 is tentative.[29]

Table 11: Demonstratives

	Singular	Plural
Proximal	*bá(ʔ)*	*dea*
Mid-distal	*baʔan / ha*	*dean*
Distal	*báát / taat / háát*	*deet*

10 Prepositional phrases and spatial relations

10.1 Prepositonal proclitics

Daats'íin has four prepositional proclitics: /ká=/ 'to, for' (dative) (25), (29)-(30); /ka=/ 'with' (instrumental/comitative) (26), (58); /ná=/ 'in, on, at, from' (locative/ablative) (59); and /jáá=/ 'of' (genitive). The genitive construction (60) is distinct from the associative construction (§8.1) in that with the associative the possessor names a general category (e.g. 'cow head') and with the genitive the possessor names a more specific entity (e.g. 'cow's head').

(58) alálágí-c ka=ɓaga   deán.
      meet-2PL COM=person MID.PL
      'Meet (2PL) with those people.'

(59) álja   ná=tʃók'ó
      1pl.EXCL LOC=house
      'We (EXCL) are at home.'

(60) ilk'wá jáá=rus
      head GEN=cow
      '(the) cow's head'

The genitive proclitic can also combine with the bound possessive pronominals to form genitive pronouns (61).

---

[29] Some conflicting data suggests the final /t/ on what I have called the Distal forms may not denote distance but rather specificity. The final /t/ is likely related to the final /t/ found in relative clauses (§11).

(61) dua bá   jaa=múlja.
      child PROX GEN=1PL.EXCL.POSS

      'This child is ours (EXCL).'

One can also form a genitive pronoun with the genitive proclitic + /go/ 'place' + a bound possessive pronominal (62)-(64).

(62) dua bá   jáá=gó-ákwa.
      child PROX GEN=PLACE-1PL.INCL.POSS

      'This child is ours.'

(63) dua bá   jáá=go-áca.
      child PROX GEN=PLACE-2PL.POSS

      'This child is yours (PL).'

(64) dua bá   jáá=go-máʕam.
      child PROX GEN=PLACE-3PL.POSS

      'This child is theirs.'

10.2 Relator nouns

Relator nouns in Daats'íin are grammaticalized body part terms (in addition to /go/ 'place'), which combine with nouns and the three prepositional proclitics, /ká=/ 'to, for', /ka=/ 'with', and /ná=/ 'in, on, at, from', to form complex prepositional phrases. The most commonly used preposition for this construction (in the present corpus) is the locative /ná=/ (65)-(67).

(65) giʃá a-ʔíí-f   ná=kʼó-tʃokʼwa.
      rock 3SG-be-HIP LOC=HEAD-house

      'The rock is on top of the house. '

(66) giʃá a-ʔíí   ná=étʃ-étʃé-tʃokʼwa.
      rock 3SG-be LOC=BACK-REDUP-house

      'The rock is behind the house.'

(67) giʃá a-ʔíí   ná=ganɗo-tʃokʼwa.
      rock 3SG-be LOC=FOREHEAD-house

      'The rock is in front of the house.'

11 Clausal syntax

This section presents some brief notes on clause-level matters. The major constituent order in Daats'íin clauses tends to be AVO/SV (68)-(69). In embedded clauses, the constituent order remains the same (33), (86). Constituent order in certain copular clauses, by contrast, is relatively more free (65), (79).

(68) gáfá   k-a-gám-tsá   gwinza.
       woman AFF-3SG.TR-know-BODY man

       'The woman saw the man.'

(69) gáfa   k-á-dugwa.
       woman AFF-3SG.INTR-run

       'The woman ran.'

There appears to be a nominative-accusative system of case marking on nouns, though this needs further research. S/A arguments tend to have a final High-tone /-á/ (70)-(72); and P arguments tend to have a final Low-tone /-a/ (73), or lack the suffix /-a/ (70), (74). For example, the A argument 'woman' ends with /-á/ in (68), but has no suffix as a P argument (70). Likewise, the P argument 'man' in (68) ends with /-a/, but terminates with /-á/ as an A argument (70). However, S arguments do not always follow the A argument-marking pattern (69). The noun /ɓag(a)/ 'person' always carries the same tone regardless of whether it functions as S, A, or P (73)-(74), though it often lacks the final /a/ when it functions as a P argument (74).

(70) gwinzá k-a-gám-ts   gáf.
       man   AFF-3SG.TR-find-BODY woman

       'The man saw the woman. '

(71) jahwá   k-á-dugwa.
       reedbuck AFF-3SG.INTR-run

       'The reed buck ran.'

(72) jahwá   k-á-biʔe.
       reedbuck AFF-3SG.INTR-fall

       'The reed buck fell.'

(73) ɓaga k-a-ráfaʔa   jahwa.
       person AFF-3SG.TR-carry reedbuck

       'The person carried the reed buck.'

(74) ɓaga   k-a-ʃa-k'w-é   ɓag.
       person AFF-3SG.TR-die-HEAD-TWRD person

       'A person killed (another) person. '

Tone marking for free pronouns is yet even more inconsistent, suggesting a possible split in the case marking system of pronouns (Kelly 2014).[30]

Copular clauses either link a noun phrase (NP) with another NP (predicate nominal constructions); or link a NP with a location (predicate locative constructions), usually

---

[30] Recall that the bound pronominals in Daats'íin also distinguish between S and A arguments, also suggesting some kind of split system.

expressed as a prepositional phrase (PP). Most copular clauses in Daats'íin involve juxtaposition, either NP-NP (77) or NP-PP (59). Alternatively, the copula /ʔíí/ can also be used for predicate locative constructions (66)-(67). For predicate nominals in the past, one must use the copula /káán/.[31] This copula form is fixed, regardless of person (76)-(78).

(75) áɗá etá-maj.
      1SG SG.HUM-field
      'I am a farmer.'

(76) áɗá káán   etá-maj.
      1SG COP.PST SG.HUM-field
      'I was a farmer.'

(77) jarʕám káán   etá-maj.
      3SG   COP.PST SG.HUM-field
      'He was a farmer.'

(78) álja   káán   dáá-maj.
      1PL.EXCL COP.PST PL.HUM-field
      'We (EXCL) were farmers.'

To express a predicate nominal in the future, one can use the imperfective form of the verb 'sit' /ʔíí-f(a)/ (79) or the verb 'become' /bági/ (80).

(79) áljá   dáá-maja   má-ʔíí-íla-fa.
      1PL.EXCL PL.HUM-field IMP-be-1PL.EXCL-HIP
      'We (EXCL) will be farmers.'

(80) (baʔadéén) áɗá ḿ-bágí-ɗ   etá-maj.
      later   1SG IMP-become-1SG SG.HUM-field
      '(Later) I will be a farmer.'

To express a predicate locative in the recent past, one uses the perfective form of the verb 'be' just as one would for an event in the present (81).[32] For the remote past, one uses the perfective form of the verb /ɗál-af/ marked with the past suffix and preceded by a prepositional phrase indicating distant past time (82).

(81) ʔám k-á-ʔíí   ná=tʃók'ó.
      2SG AFF-3SG.INTR-be LOC=house
      'You were in the house.'

---

[31] This past copula form was apparently borrwed from Arabic.
[32] Expressing a copular clause in the present versus recent past appears to involve the presence versus absence of the affirmative (AFF) prefix.

(82) áljá   ná=baʔadéén-át k-íl-ɗál-bá-f   ná=tʃók'ó.
     1PL.EXCL LOC=later-DIST? AFF-1PL.EXCL-dwell?-PST-HIP LOC=house
     'We were at home a while back.'

To express a future time predicate locative, one uses the imperfective form of the verb /ɗál-af/ in addition to the time adverb 'later' (83).

(83) áljá   **baʔadéén** má-ɗál-ílá-f   ná=tʃók'ó.
     1PL.EXCL **later**   IMP-dwell?-1PL.EXCL-HIP LOC=house
     'We will be at home.'

Relative clauses in Daats'íin are introduced with the relativizer /ba=/. Clauses that relativize on S and A are doubly marked. In such instances the lexical head of the relative clause is marked with the morpheme /=(a)t/. This same morpheme serves as a clitic marking the end of the relative clause, whether the final element is a noun (84) or a verb (85).

(84) ɓaga-**t**   ba=k-a-cíl   gáfa=**t**   k-á-lag-é.
     person-RELC REL=AFF-3SG.TR-hit woman=RELC AFF-3SG.INTR-return-TWRD
     'The person who hit the woman returned.'

(85) ɓaga-**t**   ba=k-á-dugu=**t**   k-á-lag-é.
     person-RELC REL=AFF-3SG.INTR-run=RELC AFF-3SG.INTR-return-TWRD
     'The person who ran returned.'

For relativization on locations, the relative clause (RELC) suffix /=(a)t/ is not used. The relative clause is introduced with /ba=/ and the verb of the relative clause has the incorporated noun /-go/ 'place'. In addition, the head of the relative clause is the noun /go/ 'place' which functions as a relative pronoun (86)-(87).

(86) ba=k-a-ɗáb-agó   ká=gó   ba=káána   ká=gát.
     REL=AFF-3SG-reach-PLACE DAT=PLACE REL=COP.PST DAT=here
     'The place s/he arrived was here.'

(87) ba=k-a-ɗáb-agó   gó   ba=káána   k-aɗa-gám-tsa.
     REL=AFF-3SG-reach-PLACE place REL=COP.PST AFF-1SG.TR-find-BODY
     'I saw the place where s/he arrived. '

12 Conclusion

This brief sketch of Daats'íin hopefully lays the groundwork for future investigation of the language and its relationship to other languages of the area. Daats'íin is clearly related to the Gumuz languages, though many questions still remain about the grammar. As Daats'íin has relatively few speakers and is spoken in an area where Arabic, Amharic

and Oromo are used as languages of wider communication, it is crucial that the language be fully described before the Daats'íin people abandon their language, as they apparently have in Sudan.

Abbreviations

Body part nouns (+ /go/ 'place') that have grammaticalized to some degree are glossed in small caps and are not included in this list, save the relative pronoun, /go/.

1,2,3	1st, 2nd, 3rd person	NMLZ2	(product) nominalizer, derives nominals which retain no verbal arguments
A	most agent-like argument of two argument verb		
AFF	affirmative mood	NoG	Northern Gumuz
C	consonant	NP	noun phrase
COM	comitative	P	most patient-like argument of two argument verb
COP	copula		
DAT	dative	PL	plural
EP	external possession	PLACE	relative pronoun for 'place, where'
EXCL	exclusive		
FUT	future	POSS	possessive
GEN	genitive	PRO	pronoun
GP	greater plural	PROX	proximal demonstrative
HUM	human	PST	past
H	high (tone)	RECP	reciprocal
IDEO	ideophone	REDUP	reduplicate
IMP	imperfective	REL	relativizer
IMPL	impersonal	RELC	enclitic marking first and last elements of a relative clause
IN	incorporated noun		
INCL	inclusive		
INSTR	instrumental	S	single argument of verb
INTR	intransitive	SG	singular
IP	inherent possession	SoG	Southern Gumuz
L	low (tone)	TEMP	temporal clause prefix
LOC	locative prepositional proclitic	TR	transitive
		TWRD	action directed towards speaker; action taking place in different location from speaker
MID	mid-distal demonstrative		
MV	middle voice		
NN	noun-noun		
NHUM	nonhuman	V	vowel; verb
NEG	negative	VR	valence reducer
NFUT	nonfuture		
NMLZ	nominalizer, verbal noun		

Acknowledgements

Appendix

The English list of words is taken from: Haspelmath & Tadmor (2009). The Gumuz column represents Southern Gumuz unless otherwise indicated.

Word meaning	Daats'íin	Gumuz (S. Gumuz)
1 fire	*too*	*máánɟa*
2 nose	*íít*	*ííta*
3 to go	*ɗa*	*ɗá*
4 water	*áʔéé*	*aja*
5 mouth	*hós*	*sa*
6 tongue	*t'at'é*	*k'ót'át'á*
7 blood	*maha*	*mahá*
8 bone	*voko*	*ʒákwá*
9 2sg pronoun	*ámam*	*áam*
10 root	*tírágí*	*tíŋáɟá*
11 to come	*wéá*	*wé*
12 breast	*kú*	*kúá*
13 rain	*dám*	*dama*
14 1sg pronoun	*áɗa*	*ára*
15 name	*ts'ê*	*ts'éa*
16 louse	*sankun*	*sakúná* (NoG: head louse)
17 wing	*p'áp'á*	*p'áp'a*
18 flesh/meat	*bátʃʾ*	*batʃʾa*
19 arm/hand	*éé*	*ela*
20 fly	–	–
21 night	*mágúŋkú*	*magókwa*

22 ear	*ts'ê*	*ts'éa*
23 neck	*ɓi*	*ɓia*
24 far	*háátí*	*háat*
25 to do/make	*háádama or dá*	*dá*
26 house	*tʃok'o*	*máts'á*
27 stone/rock	*giʃa*	*gíʃá*
28 bitter	*ántʃ'iŋ*	*nc'áŋ*
29 to say	*kéa*	*kál*
30 tooth	*k'ôs*	*k'ósa*
31 hair	*bek'o*	*béék'wá*
32 big	*baabúʃaʔam*	*babák'ómá*
33 one	*mité*	*metáam*
34 who?	*wada*	*odé*
35 3sg pronoun	*jáárʔám*	*áŋa*
36 to hit/beat	*tʃííla*	*átʃ*
37 leg/foot	*tʃugw*	*tʃogwa*
38 horn	*ɟik'o*	*k'ála* (NoG: *ɟúk'ómá*)
39 this	*baa(ʔ)*	*baaʔ*
40 fish	*wii*	*uŋa*
41 yesterday	*náfinɟígin*	*ná=mágáázig*
42 to drink	*fa*	*fá*
43 black	*jaraʔii*	*ŋíí*
44 navel	*k'ófak'u*	*k'ófagwa*
45 to stand	*íˤkátʃugwa*	*íitʃogw*
46 to bite	*k'áŋa*	*k'áɲ*
47 back	*étʃá*	*tʃítʃá*
48 wind	*zumbór*	*zibá*
49 smoke	*ɗak'u*	*súŋmaanɟa*
50 what?	*ńtsaka*	*ntsá*
51 child (kin term)	*du*	*dua*
52 egg	*íʃ*	*iʃa*
53 to give	*kʲá*	*cá*
54 new	*jáácici*	*cicá*
55 to burn (intr.)	*sakʲé too* (fire burning)	*(ɟi)súk'*
56 not	*=kʲé*	*=cê*
57 good	*gaʃeʔam*	*gaʃalamá*
58 to know	*gama*	*gam*
59 knee	*k'ókoku*	*k'úcítʃogwa*
60 sand	*sanseʔ*	*saasama*
61 to laugh	*gec'*	*alo*
62 to hear	*gásak'o*	*gásak'w*
63 soil	*ílmai*	*ńtʃók'wa*
64 leaf	*ts'éiŋgi*	*ts'ênɟá*

65 red	jaawiᵛʔe	ééc
66 liver	ndihi	andííá
67 to hide	mbaʃia	bats' (maaʃiukʼw)
68 skin/hide	be	bea
69 to suck	asámtsa ásámkʼó	ts'imúkʼw
70 to carry	ráfaʔa	bur
71 ant	bigámz (black ant)	basésíá
72 heavy	áhicʼa	hic'
73 to take	kántsa	tséets
74 old	gaz (person)	gááza
75 to eat	sa	sá
76 thigh	albáádar	ílcʼáʃa
77 thick	átaba	tab
78 long	aráɗa	zem
79 to blow	fútʃʼa	fwítsʃ
80 wood	gí	ɟá
81 to run	dugwa ranɗócʼa (many)	dugw
82 to fall	bíʔé	faat (SoG) beʔ (NoG)
83 eye	(kʼó)kéé	kʼócá
84 ash	mfoʔo	mfá
85 tail	ts'íŋts'iŋ	tsia
86 dog	kʼaw	kʼóá
87 to cry/weep	ísí	íís
88 to tie	fántʃʼá	afíntʃʼ (NoG)
89 to see	gíʃí	jîr
90 sweet	mínna	mîn
91 rope	sí	síá
92 shade/shadow	masaŋgil	masáánɟílá
93 bird	mété	mátá
94 salt	gunn	saŋa
95 small	duʃíʃaaʔam	dúganáátsámá
96 wide	áfágááʔííl	fagííl (widen)
97 star	ntʃʼékʼ	bííʒa
98 in	íílʔám	níílámá
99 hard	áhicʼa	ʒígâts
100 to crush/grind	garʃ	ganʃ

References

Ahland, Colleen. 2012. *A grammar of Northern and Southern Gumuz*. Eugene, OR: University of Oregon. (Doctoral dissertation.)

Biermann, Anna. 1982. Die grammatische Kategorie Numerus. In Hans-Jakob Seiler & Christian Lehmann (eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen, teil. 1: Bereich und Ordnung der Phänomene*, 229–243. Tübingen: Narr.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology, 2nd edn.* Chicago: University of Chicago Press.

Corbett, Greville. 2000. *Number* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Dimmendaal, Gerrit J. 2011. *Historical linguistics and the comparative study of African languages.* Amsterdam/Philadelphia: John Benjamins.

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2000. Introduction. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Changing valency: Case studies in transitivity*, 1–28. Cambridge: Cambridge University Press.

Givón, Talmy. 2001. *Syntax: An introduction, 2 vols.* Amsterdam/Philadelphia: John Benjamins.

Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the world's languages: A comparative handbook.* Berlin/New York: Mouton de Gruyter.

Kelly, Eliza. 2014. An exploration of alignment patterns in Daats'in: A first look at a recently identified Ethiopian language. (Paper presented at the Penn-York Undergraduate Research Conference. Houghton, NY, 1 November, 2014.)

Kemmer, Suzanne. 1993. *The middle voice.* Amsterdam/Philadelphia: John Benjamins.

Klaiman, M. H. 1991. *Grammatical voice* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Payne, Doris L. & Immanuel Barshi. 1999. External possession: What, how, where and why? In Doris L. Payne & Immanuel Barshi (eds.), *External possession* (Typological Studies in Language 39), 3–29. Amsterdam: John Benjamins.

Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists.* Cambridge: Cambridge University Press.

Siewierska, Anna. 2010. From third plural to passive: Incipient, emergent, and established passives. *Diachronica* 27(1). 73–109.

Snider, Keith & James Roberts. 2006. SIL comparative African wordlist (SILCAWL). https://www.sil.org/resources/publications/entry/7882.

Storch, Anne & Gerrit J. Dimmendaal. 2014. One size fits all? In Anne Storch & Gerrit J. Dimmendaal (eds.), *Number constructions and semantics: Case studies from Africa, Amazonia, India, and Oceania*, 1–32. Amsterdam: John Benjamins.

Welmers, William Everett. 1974. *African language structures.* Los Angeles: University of California Press.

Part III

Afro-Asiatic

Chapter 23

Somali gender polarity revisited

Morgan Nilsson

University of Gothenburg

The notion of GENDER POLARITY is a well established part of the description of Somali nouns. It refers to the phenomenon that in most Somali nouns a change in number is accompanied by a simultaneous change in gender, and it implies that number is actually expressed by means of the change in gender. However, from a synchronic point of view there seems to be little evidence for such an interpretation, as plural gender is realised solely through the shape of the definite article suffix on the noun itself. In this article the main arguments for the standard description are investigated and an alternative solution based on typological considerations of the data is proposed, claiming that gender is not relevant in the plural. Instead, the form of the plural definite article is predictable from the gender of the noun in the singular, together with some morpho-phonological characteristics of the stem. Additionally, many nouns traditionally claimed to be plural are argued here to be formally singular and mostly collective.

1 Gender polarity

According to the traditional point of view, Somali has two genders and two morphemes expressing definiteness, one for each gender: {k}[1] for masculine gender and {t} for feminine gender. The same morphemes are used both in the singular and in the plural, and the majority of Somali nouns change their gender when they change their number. All nouns that are feminine and take the article {t} in the singular become masculine and take the article {k} in the plural; whereas most nouns that are masculine and take the article {k} in the singular become feminine and take the article {t} in the plural. A smaller group of nouns which are masculine in the singular remain masculine also in the plural, and hence take the article {k} irrespective of their number. This standard point of view, illustrated in Table 1,[2] is presented in practically all modern works on Somali, among others the reference grammars by Moreno (1955), Saeed (1993; 1999), Puglielli & Mansuur (1999) and Berchem (2012).

---

[1] Curly brackets indicate underlying morphemes that are realised differently in different phonological contexts.

[2] The horizontal line in the middle of Table 1 is discontinuous as some nouns are masculine in both numbers.

Table 1: Gender polarity – the traditional view on gender in Somali nouns

	Singular	Plural	
Feminine	{t}	{k}	Masculine
Masculine	{k}	{t}	Feminine

Some examples of actual occurrences of these articles are given in Table 2. The ordering of suffixes added to noun stems is 1) plural morpheme, 2) definite article (marked by boldface), and 3) "case" ending. Somali also exhibits several morpho-phonological alternations. Depending on the stem-final phoneme, the masculine definite article suffix {k} is realised as one of the three posterior consonants /k/, /g/, /h/ or as zero; whereas the feminine definite article suffix {t} is realised as one of the three coronal consonants /t/, /d/ or /ʃ/. The most common plural suffix is realised as /o/ when word final, but as /a/ when non-final. For example, in the form *sababaha* 'the reasons' at the end of the first line in Table 2, /sabab/ is the stem, the following /a/ marks the plural, the /h/ is a realisation of the definite article {k}, and the final /a/ is a final vowel occurring in the unmarked base form of a definite noun, traditionally referred to as the absolute case.

Table 2: Nouns exemplifying the distribution of gender and number in Somali

	F.SG.INDEF	F.SG.DEF	M.PL.INDEF	M.PL.DEF
'reason'	*sabab*	*sabab-**t**-a*	*sabab-o*	*sabab-a-**h**-a*
'knife'	*mindi*	*mindi-**d**-a*	*mindi-yo*	*mindi-ya-**h**-a*
	M.SG.INDEF	**M.SG.DEF**	**F.PL.INDEF**	**F.PL.DEF**
'stool'	*gambar*	*gambar-**k**-a*	*gambar-ro*	*gambar-ra-**d**-a*
'wall'	*derbi*	*derbi-**g**-a*	*derbi-yo*	*derbi-ya-**d**-a*
	M.SG.INDEF	**M.SG.DEF**	**M.PL.INDEF**	**M.PL.DEF**
'table'	*miis*	*miis-**k**-a*	*miis-as*	*miis-as-**k**-a*

The first two nouns in Table 2, *sabab* 'reason' and *mindi* 'knife', have a feminine definite article, i.e. a coronal consonant, in the singular; whereas they have a masculine definite article, i.e. a back consonant, in the plural. They are therefore traditionally considered feminine in the singular and masculine in the plural. The next two examples, *gambar* 'stool' and *derbi* 'wall', have a masculine (back) definite article in the singular and a feminine (coronal) definite article in the plural. They are therefore traditionally considered masculine in the singular and feminine in the plural. The final example consists of the noun *miis* 'table', which is masculine both in the singular and in the plural.

Table 3: The interdependence between gender and tonal accent

	INDEF	DEF
MASCULINE		
'boy'	*ínan*	*ínan-**k**-a*
'stool'	*gámbar*	*gámbar-**k**-a*
FEMININE		
'girl'	*inán*	*inán-**t**-a*
'reason'	*sabáb*	*sabáb-**t**-a*

It is traditionally argued that a distinct gender value is associated with the plural form of nouns[3] and that the definite article is then assigned according to the plural gender of each specific noun (El-Solami-Mewis 1988; Saeed 1999: 54–55; Berchem 2012: 48–49; among others).[4] In this article, however, I argue that the form of the plural definite article is morphologically predictable without reference to plural gender. The principles for the distribution of the definite article suffixes in the plural will be discussed in detail in §2 below.

Another important exponent of gender is also regularly pointed out in the literature, namely the position of tonal accent, realised as high pitch, which falls on the second to last mora of the stem in most masculine singular nouns and on the last mora of the stem in most feminine singular nouns, as shown in Table 3.[5]

This gender-based distribution of tonal accent plays an important role in the traditional argumentation because of a group of nouns referred to as the 5th declension by Saeed (1993: 134; 1999: 61) and Orwin (1995: 48). In this group of nouns we find, for instance, *dibi* 'ox'. The nouns in the 5th declension lack an overt plural morpheme, and therefore they may be considered to exhibit "pure" gender polarity as they form their plural simply by changing their gender, which in the indefinite form is expressed merely by means of tonal accent shift.

The plural *dibí*, with the feminine definite form *dibída*, does not contain any overt plural morpheme. The plural is simply expressed by the change in gender, which is realised in the form of the definite article and the typical feminine tonal accent on the last mora of the stem.

Furthermore, with forms of the 5th declension such as *dibída* 'oxen', traditionally considered to be feminine plural, agreeing verbs and pronouns may be either singular or plural. Consider (1), where *dibída* occurs in the subject case form.

---

[3] Lecarme (2002), however, suggests that the gender value is associated with the individual plural morphemes, and that it is the plural morpheme that bears the gender value in the plural forms of nouns, not the nouns themselves.

[4] From a pedagogical perspective, it is also highly relevant that the prevailing description is perceived by most learners of Somali grammar (both Somali speakers and foreign learners) as a very complicated way of describing the rather simple facts of present-day Somali.

[5] This is not an exceptionless principle, but it applies to the vast majority of nouns ending in a consonant or a vowel other than *-e* or *-o*; nouns ending in these two vowels are subject to specific tonal accent assignment rules of their own.

Table 4: A noun in the 5th declension: *dibi* 'ox'

	INDEF	DEF	
SINGULAR	*díbi*	*díbi-**g**-a*	MASCULINE
PLURAL	*dibí*	*dibí-**d**-a*	FEMININE

(1) a. Dibí-d-u   waa ay[6] daaq-ay-aan.
        OX\PL-F.DEF-SBJ DECL PRO.3 graze-PROG-PRS.3PL
        'The oxen are grazing.'

     b. Dibí-d-u   waa ay   daaq-ay-saa.
        OX\PL-F.DEF-SBJ DECL PRO.3 graze-PROG-PRS.3FSG
        'The oxen are grazing.'

In (1a) *dibídu* 'the oxen' is followed by the predicate verb *daaqayaan* 'are grazing' in the plural, whereas it may equally well be followed by the verb form *daaqaysaa* 'is grazing' in the feminine singular (1b). This variation in agreement between the plural and the singular has been taken as evidence that plural forms like *dibí* 'oxen' must be feminine, as the singular agreement marker is feminine (Hetzron 1972: 259–260; Zwicky & Pullum 1983: 391–393; Lecarme 2002: 134–137). It has also been pointed out that the same type of variation occurs with plurals of Arabic origin which lack any Somali plural morpheme, e.g. the feminine plural *kutub-t-a* 'the books' of the masculine singular noun *kitaab-k-a* 'the book'; or the feminine plural *macallim-iin-t-a* 'the teachers' of the masculine singular *macallin-k-a* 'the teacher'.

The reason, however, for having singular verb agreement with a plural noun as in (1b) is not made very clear. In §3 of this article, I will argue that the reason for the variation in agreement is that these nouns should not be interpreted as plurals, but as collective nouns that are formally feminine singular. A strong general connection in Afroasiatic languages between feminine suffixation and the derivation of collective nouns was pointed out already by Speiser (1938), who at the same time was strongly opposed to the notion of gender polarity in general, but not specifically in Somali.

2 No gender distinction in the plural

Before going into the alternative analysis of the Somali data, it is crucial to take some typological considerations into account. Elaborating the claim by Hockett (1958: 231) that gender is "reflected in the behavior of associated words", Corbett (2000: 89–90) states that the "relevant 'reflection' in the associated words is agreement [...]. No amount of marking on a noun can prove that the language has a gender system; the evidence that

---

[6] The subject pronoun *ay* 'she, it, they', which is used in positive statements even if there is a subject noun, is homonymous in the feminine singular and in the plural.

nouns have gender values in a given language lies in the agreement targets which show gender." Building on Corbett's definition, I hence assume that gender is not present in a noun unless it is reflected as agreement on other associated words. If morphological differences in the noun itself would be enough to distinguish between genders, there are as many as seven different realisations of the definite article suffix in the singular of Somali nouns. However, nobody has proposed seven genders, precisely because the agreement patterns in pronouns and verbs only give evidence for two genders.

In order to establish whether Somali has a gender distinction in the plural, we will therefore have a closer look at the morphological exponents of agreement in Somali in other parts of speech than the noun itself. A practically exhaustive list of such exponents is found in Table 5.[7]

Table 5: Exponents of agreement in Somali

			SG	PL
VERBS	PRESENT INDICATIVE	M	*-aa*	*-aan*
		F	*-taa*	
	PAST INDICATIVE	M	*-ay*	*-een*
		F	*-tay*	
	PRESENT SUBJUNCTIVE	M	*-o*	*-aan*
		F	*-to*	
PRONOUNS	REDUCED PERSONAL	M	*uu*	*ay*
		F	*ay*	
	FULL PERSONAL	M	*isaga*	*iyaga*
		F	*iyada*	
	DEMONSTRATIVE	M	*kan*	*kuwan*
		F	*tan*	
	POSSESSIVE	M	*-iis*	*-ood*
		F	*-eed*	

There exist a few other verb categories and types of pronouns than those listed in Table 5, but the pattern regarding the type of forms remains just the same. The important point is that there is not a single category in which we find a gender distinction in the plural. This is typologically quite a common pattern, and is also found in languages such as Russian, Swedish and German.

---

[7] No gender distinction is made in the singular of adjectives, and plurality is optionally expressed by means of reduplication of the initial CV(C).

As it is clear that Somali does not exhibit any agreement distinctions in gender in the plural in any associated words, there is no reason to define the gender of a noun in the plural. The variation found in the morphology of plural nouns themselves, that is, in the form of the plural definite article, is instead easily predictable on morphophonological grounds. The singular gender and the syllabic structure of the noun are enough to make the right choice of the plural definite article. The principles are that 1) feminine nouns take the definite article {k} in the plural; 2) masculine nouns in which the plural morpheme is preceded by a monosyllabic stem take the definite article {k} in the plural;[8] and 3) masculine nouns with bisyllabic or longer stems take the definite article {t} in the plural.

As there is no need for a gender distinction in the plural for agreement purposes, there is consequently no need to posit a gender shift for the majority of nouns as in the traditional analysis. Instead, just as in many other languages such as Russian, Swedish or German, a noun is most conveniently interpreted as having the same gender at all times, as proposed in Table 6.

Table 6: Nouns exemplifying a simplified view of gender in Somali

	SG.INDEF	SG. DEF	PL.INDEF	PL.DEF
FEMININE NOUNS				
'reason'	*sabab*	*sabab-**t**-a*	*sabab-o*	*sabab-a-**h**-a*
'knife'	*mindi*	*mindi-**d**-a*	*mindi-yo*	*mindi-ya-**h**-a*
'mother'	*hooyo*	*hooya-**d**-a*	*hooyo-oyin*	*hooyo-oyin-**k**-a*
MASCULINE NOUNS				
MONOSYLLABIC STEM				
'table'	*miis*	*miis-**k**-a*	*miis-as*	*miis-as-**k**-a*
'tree'	*geed*	*geed-**k**-a*	*geed-o*	*geed-a-**h**-a*
'knee'	*jilib*	*jilib-**k**-a*	*jilb-o*	*jilb-a-**h**-a*
POLYSYLLABIC STEM				
'stool'	*gambar*	*gambar-**k**-a*	*gambar-ro*	*gambar-ra-**d**-a*
'wall'	*derbi*	*derbi-**g**-a*	*derbi-yo*	*derbi-ya-**d**-a*
'father'	*aabbe*	*aabba-**h**-a*	*aabba-yaal*	*aabba-yaa-**sh**-a*

---

[8] The only systematic violations of this principle are exhibited by nouns forming their plural by reduplication of a final /l/. In this group of nouns, those with a long root vowel have a definite form with {t} (which becomes /ʃ/ under influence of the preceding /l/), probably due to analogy with the very many nouns ending in /e/ that all take the plural suffix *-yaal*, e.g. *wiilal* 'boys', def. *wiilasha*, in line with *aabbayaal* 'fathers', def. *aabbayaasha*. Apart from these, there is also a very small number of high frequency words that behave in an irregular manner, e.g. *walaal* m. 'brother' and *walaal* f. 'sister', which both correspond to the plural *walaalo* 'siblings', def. *walaalaha*.

As becomes evident from Table 6, the polar syncretism is not total. Many masculine nouns have the same definite article both in the singular and in the plural.[9] Actually, such polar syncretisms are not unique. To a varying extent, they can also be observed in other languages, for instance in the Swedish definite article (Table 7), or in the nominative dual and plural suffixes of Slovene nouns (Table 8).

Table 7: Definite forms of Swedish nouns

		SINGULAR	PLURAL
COMMON GENDER	'the stone'	sten-**en**	sten-arna
NEUTER	'the house'	hus-et	hus-**en**

Table 8: Nominative forms of Slovene nouns

		SINGULAR	DUAL	PLURAL
MASCULINE	'train'	vlak	vlak-**a**	vlak-**i**
NEUTER	'window'	okn-o	okn-**i**	okn-**a**

3 Agreement variation

In this section I discuss agreement as evidence for gender polarity in Somali. It is crucial to recall, as was pointed out in Table 5, that there is no gender distinction in the exponents of agreement in the plural. The argument used by Hetzron (1972), Zwicky & Pullum (1983) and Lecarme (2002) for forms like *dibída* 'the oxen' to be feminine plural is not that they trigger agreement in the feminine plural, as no such form exists, but that they exhibit a variation in agreement between the plural form and the feminine singular form of a predicate verb. The discussion in the articles mentioned is limited to nouns in the 5th declension, such as *tuugta* 'the thieves' in (2a) and (2b), and nouns with Arabic plural forms, such as *kuraasta* 'the chairs' in (2c) and (2d), both groups showing a variation between plural and feminine singular agreement in the predicate verb.

(2) a. Si-d-ee   tuug-t-u   u feker-aan?
         manner-DEF-Q thieves-DEF-SBJ in think-PRS.3PL
         'How do thieves think?'

     b. Si-d-ee   tuug-t-u   u feker-taa?
         manner-DEF-Q thieves-DEF-SBJ in think-PRS.3FSG
         'How do thieves think?'

---

[9] Actually, according to my native speaker consultants, some Somali dialects in Ogaden only use the definite article {k} for all nouns in the plural, meaning that nouns have come to behave just like all the other parts of speech by exhibiting only one common form in the plural.

    c. Kuraas-t-u   waxa ay   ku wareeg-san   yihiin   miis weyn.
       chairs-DEF-SBJ FFOC PRO.3 in go.around-PTCP be.PRS.3PL table big

       'The chairs are placed around a big table.'

    d. Kuraas-t-u   waxa ay   ku wareeg-san   tahay   miis weyn.
       chairs-DEF-SBJ FFOC PRO.3 in go.around-PTCP be.PRS.3FSG table big

       'The chairs are placed around a big table.'

The literature does not discuss whether there are other nouns, apart from the 5th declension and Arabic-borrowing types, which exhibit similar variation in their agreement patterns. It is therefore important to point out that many other frequently-occurring nouns do exhibit the same type of agreement patterns in predicates and pronouns with regard to gender and number. In (3a) the feminine noun *carruurtu* 'the children' triggers plural agreement in the finite verb and the possessive suffix *-ood* 'their', whereas in (3b) the same noun is followed by the feminine singular finite verb *karto* 'can'. This noun, however, is traditionally not claimed to be plural, but instead to be a singular collective noun.

(3)   a. Carruur-t-u   waxa ay   jecel yihiin   waalid-k-ood.
        children-DEF.F-SBJ FFOC PRO.3 fond be.PRS.3PL parents-DEF.M-POSS.3PL

        'The children love their parents.'

    b. Kor u qaad buug-ga si   ay   carruur-t-u   u arki kar-to.
        top to take book-DEF manner PRO.3 children-DEF-SBJ for see can-SBJV.3FSG

        'Hold up the book so that the children can see it.'

    c. Maxaa   ay   samay-nay-aan dad-k-u?
        what\FOC PRO.3 do-PROG-PRS.3PL people-DEF-SBJ

        'What are people doing?'

    d. Ma sheegi kar-taa   xayawaan-no uu   dad-k-u
        Q tell   can-PRS.2SG animal-PL   PRO.3MSG people-DEF-SBJ
        dhaqd-o?
        breed-SBJV.3MSG

        'Can you say some animals that people breed?'

Even more interestingly, a substantial proportion of such nouns are masculine, and these nouns therefore trigger variation between the plural and the masculine singular. In (3c) the masculine noun *dadku* 'the people' triggers plural agreement in the subject pronoun[10] *ay* and the finite verb *samaynayaan* 'do', whereas in (3d) the very same noun is accompanied by the masculine singular subject pronoun *uu* and the masculine singular finite verb *dhaqdo* 'breeds'.

Having a larger set of data, it also becomes clear that this type of variation in predicate and pronoun agreement patterns occurs only with nouns which lack an overt synchronic Somali plural morpheme, but which have a meaning that may be perceived as plural, as

---

[10] Subject pronouns are used in positive statements even if there is a subject noun.

in (4a) and (4b). However, nouns which have an indisputable plural morpheme, i.e. an overt synchronic Somali plural ending, never trigger any such variation. Such nouns are always accompanied by verbs and pronouns in the plural, as in (4c), whereas (4d) is incorrect.

(4) a. Dibi-d-u   w-ay   daaq-ay-aan.
        oxen-DEF-SBJ DECL-PRO.3 graze-PROG-PRS.3PL
        'The oxen are grazing.'

     b. Dibi-d-u   w-ay   daaq-ay-saa.
        oxen-DEF-SBJ DECL-PRO.3 graze-PROG-PRS.3FSG
        'The oxen are grazing.'

     c. Dibi-ya-d-u w-ay   daaq-ay-aan.
        ox-PL-DEF-SBJ DECL-PRO.3 graze-PROG-PRS.3PL
        'The oxen are grazing.'

     d. *Dibi-ya-d-u w-ay   daaq-ay-saa.
        ox-PL-DEF-SBJ DECL-PRO.3 graze-PROG-PRS.3FSG
        'The oxen are grazing.'

From a typological point of view, variation between agreement in the singular and the plural is quite common. It is reported from a diversity of languages, and it occurs, for instance, in both Swedish and English with words such as 'the team', as in (5) and (6).

(5) a. Lag-et   kom   in på plan-en   inställ-**t**   på att vinn-a.
        team-SG.DEF come-PRT in on field-SG.DEF determined-SG on to win-INF

     b. Lag-et   kom   in på plan-en   inställ-**da**   på att vinn-a.
        team-SG.DEF come-PRT in on field-SG.DEF determined-PL on to win-INF
        'The team entered the playground determined to win.'

(6) a. The team **is** friendly.
     b. The team **are** friendly.

As pointed out by Corbett (2000: 187), the typologically interesting distinction is between two types of agreement, namely syntactic agreement, determined by the form of the noun; and semantic agreement, determined by the meaning of the noun. Typologically, there are certain types of nouns that typically trigger this kind of variation, and Somali agreement variation fits neatly into this typological pattern, as the Somali nouns triggering such variation belong to typologically expected categories.

First, collective nouns commonly cause this type of variation in a number of languages. Corbett (2000: 118–119) uses the term *collective* to indicate that a noun is "referring to a group of items considered together rather than a number of items considered individually. […] The primary function of collectives is to specify the cohesion of a group". Most of the Somali nouns exhibiting variation in their agreement patterns can be included in this category. Actually, Puglielli & Siyaad (1984: 82) state that the plurals of the Somali

5th declension have been recategorised as collective forms, but despite this claim they still choose to treat forms like *dibída* 'oxen' as formally plural.

It should, however, be pointed out that the collective nouns of the 5th declension in most instances also have a regular plural counterpart based on the masculine singular form of the same word; hence there are two forms expressing plural meaning, but only one form with a plural morpheme. Because the feminine collective form exhibits agreement variation in the predicate verb, allowing both the plural and the feminine singular verb forms, I claim that the feminine collective form should be treated as formally singular, as shown in Table 9.

The type of nouns found in the 5th declension are highly interesting, as these nouns exhibit a singular, a regular plural and a collective form based on the very same root. The number of such noun stems is just a couple of dozen, but interestingly enough some newer words, often not mentioned in the literature, have also found their way into this group, e.g. the Arabic[11] loanword *baabuur* 'car', illustrated by (7), as well as the English loanword *buug* 'book'.

(7) Baabuur-t-u   waa ay   nooc-yo badan yihiin.
     car\F.COLL-DEF-SBJ DECL PRO.3 type-PL many be.PRS.3PL

     'There are cars of many types.'

Table 10 presents some further nouns of the 5th declension type with both their regular plural form and their collective singular form.

Interestingly enough, a very similar pattern can be found in Swedish for the noun *mygga* 'mosquito' (Table 11).

Judging from further corpus data, it also becomes clear that borrowed Arabic plural forms behave in the same manner as indigenous Somali collective nouns. The varying agreement is also confirmed by Puglielli & Siyaad (1984: 86), but they refrain from calling these forms collectives. I will, however, argue that such feminine forms, containing Arabic plural morphemes, are not plurals in Somali, but collective forms which are grammatically singular. The argument is based on the same type of variation in agreement as that encountered with nouns of the 5th declension, and on the fact that these nouns also exhibit a regular plural with an overt Somali plural morpheme (Table 12).

Some speakers prefer plural forms over collective forms when referring to a smaller number of individualised objects, whereas the collective nouns refer to many objects as a coherent group.

Other examples of nouns that behave as collective nouns, not included within the two types already discussed, are the feminine nouns *carruur-t-a* 'the children' and *lo'-d-a* 'the cattle', as well as the masculine nouns *dumar-k-a* 'the women', *dad-k-a* 'the people', and *rag-g-a* 'the men'. None of these has a corresponding individualising singular form based on the same root. Instead, they have corresponding individualising forms based on another root, as shown in Table 13.

---

[11] Which in its turn is most probably a loan from the French *vapeur* 'steam' (Jan Retsö, personal communication).

Table 9: Typical forms of nouns with both a plural and a collective form

	SG.INDEF	SG.DEF	PL.INDEF	PL.DEF
M	*díbi* 'ox'	*díbi-**g**-a* 'the ox'	*dibi-yó* 'oxen'	*dibi-yá-**d**-a* 'the oxen'
F.COLL	*dibí* 'oxen'	*dibí-**d**-a* 'the oxen'		

Table 10: Further nouns with both a plural and a collective form

	SG.DEF		PL.DEF	
M	*baabúur-**k**-a*	'the car'	*baabuur-rá-**d**-a*	'the cars'
F.COLL	*baabuúr-**t**-a*	'the cars'		
M	*búug-**g**-a*	'the book'	*buug-ág-**g**-a*[a]	'the books'
F.COLL	*buúg-**t**-a*	'the books'		
M	*túug-**g**-a*	'the thief'	*tuug-ág-**g**-a*	'the thieves'
F.COLL	*tuúg-**t**-a*	'the thieves'		

[a] *Buugagga* is the regular form. There is, however, a general preference for an irregular plural with a long epenthetic /a:/ and the article {t}, i.e. *buugaagta* 'the books'.

Table 11: Forms of the Swedish noun *mygga* 'mosquito'

	SG.INDEF	SG.DEF	PL.INDEF	PL.DEF
INDIV	*mygg-a*	*mygg-a-n*	*mygg-or*	*mygg-or-na*
COLL	*mygg*	*mygg-en*		

Table 12: Forms of a typical noun 'chair' exhibiting a borrowed Arabic plural

	SG.INDEF	SG.DEF	PL.INDEF	PL.DEF
M	*kursi*	*kursi-**g**-a*	*kursi-yo*	*kursi-ya-**d**-a*
F.COLL	*kuraas*	*kuraas-**t**-a*		

Table 13: Different lexemes for collective and individualising meanings

	SG.DEF		PL.DEF	
F	*naag-**t**-a*	'the woman'	*naag-a-**h**-a*	'the women'
M.COLL	*dumar-**k**-a*	'the women'		
M	*nin-**k**-a*	'the man'	*nim-an-**k**-a*	'the men'
M.COLL	*rag-**g**-a*	'the men'		

Another interesting group of Somali nouns, which are indisputably singular, form a totally regular plural; and yet, in the singular, they may trigger both syntactic agreement in the singular and semantic agreement in the plural. These words are, according to Corbett (2000: 188), so-called *corporate nouns*, i.e. "nouns which are singular morphologically and (typically) have a normal plural and yet, when singular, may take plural agreement". This pattern can frequently be observed in Somali and seems to be equally common in both genders. Some examples are in Table 14.

Table 14: Somali corporate nouns

	SG.INDEF	SG.DEF	PL.INDEF	PL.DEF
'family'	*qoys*	*qoys-**k**-a*	*qoys-as*	*qoys-as-**k**-a*
'group, team'	*koox*	*koox-**d**-a*	*koox-o*	*koox-a-**h**-a*
'camel herd'	*geel*	*geel-a*[a]	*geel-al*	*geel-a-**sh**-a*

[a] This noun is a unique exception, realising the definite article {k} as zero in the singular.

In (8), we find the masculine singular noun *qoyskiisu* 'his family' with the predicate verb *dhaqdaan* 'breed' and the subject pronoun *ay*, here[12] representing the plural meaning 'they'. The verb and the pronoun could, however, just as well have been in the masculine singular, like the predicate verb *yahay* 'is' and the subject pronoun *uu* 'he, it' in (9).

(8) Faarax qoys-**k**-iis-u   waxa **ay**   dhaqd-**aan**   geel.
     Faarax family\MSG-DEF-POSS.3MSG-SBJ FFOC PRO.3 breed-PRS.3PL camel\M.COLL

     'Faarax's family breeds camels.'

(9) Qoys-**k**-iis-u   waxa **uu**   ka kooban   **yahay**   shan
     family\MSG-DEF-POSS.3MSG-SBJ FFOC PRO.3MSG of consisting be.PRS.3MSG five

---

[12] The homonymous subject pronoun *ay* 'they; she, it' represents the 3rd person plural as well as the feminine 3rd person singular. When it occurs together with a masculine subject noun, it may therefore only be interpreted as plural.

ruux.
person

'His family consists of five persons.'

Typologically, many languages also exhibit variation in number when nouns are used with a generic meaning. English, as well as Swedish, exhibits such variation between the definite singular and the indefinite plural form of the generic noun itself, as shown in (10).

(10) a. The tiger is in danger of becoming extinct.
      b. Tigers are in danger of becoming extinct.

In Somali, generic meaning is very often expressed by the singular definite form of the noun, as with *diinku* 'the turtle' in (11). The agreement of accompanying constituents is, however, often in the plural, like the verb *sameeyaan* 'make' in the same example.

(11) Diin-k-u   inta   badan ma sameey-**aan**   wax   dhaqdhaqaaq ah.
     turtle\MSG-DEF-SBJ amount much not make-PRS.3PL thing movement   being.

     'For a long period of time turtles don't make any movements.'

Based on the quite diverse categories of nouns which have been shown to trigger variation between singular and plural agreement, I claim that such variation is a typical general trait of Somali syntax, reaching far beyond the examples frequently discussed as results of the traditionally posited gender polarity.

4 Summary

Building on the data presented in this article, I would like to make two important claims.

- In Somali, gender is not a relevant category in the plural as there is no need to refer to the gender of a noun in the plural for the sake of agreement. Therefore, nouns should only be ascribed one gender value based on their behaviour in the singular.
- Variation in number in agreement patterns has nothing to do with plural gender. Instead, nouns that trigger variation in number agreement are grammatically singular. They are collective nouns, corporate nouns or common nouns used in a generic sense. Agreement in the singular is syntactically conditioned, whereas agreement in the plural is semantically conditioned.

Therefore, instead of the traditional view referred to as *gender polarity*, presented above in Table 1, I propose a far more simple interpretation of the definite articles and the gender system, as shown in Table 15. This analysis of the Somali gender system and its morphological exponents of definiteness is typologically uncontroversial. The notion of polarity may, of course, still be applied; but if so, only in order to refer to the morphological exponents (i.e. forms) of definiteness.

Table 15: Polarity of the exponents of definiteness in Somali nouns

	Singular	Plural
Feminine	{t}	{k}
Masculine	{k}	{t}

Acknowledgements

Abbreviations

2	2nd person	M	masculine
3	3rd person	MSG	masculine singular
COLL	collective	PL	plural
DECL	declarative particle	POSS	possessive
DEF	definite	PRO	pronoun
F	feminine	PRS	present
FFOC	final focus, i.e. the clause final NP is focused	PROG	progressive
		PRT	preterite
FOC	focus particle	PTCP	participle
FSG	feminine singular	Q	question marker
INDEF	indefinite	SBJ	subject
INDIV	individualising	SG	singular
INF	infinitive		

References

Berchem, Jörg. 2012. *Referenzgrammatik des Somali.* Norderstedt: Books on Demand.

Corbett, Greville. 2000. *Number* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

El-Solami-Mewis, Catherine. 1988. On classifying Somali nouns by their plural classes. In Marianne Bechhaus-Gerst & Fritz Serzisko (eds.), *Cushitic-Omotic: Papers from the international Symposium on Cushitic and Omotic Languages (Cologne, January 6-9, 1986)*, 555–563. Hamburg: Helmut Buske Verlag.

Hetzron, Robert. 1972. Phonology in syntax. *Journal of Linguistics* 8(2). 251–265.

Hockett, Charles F. 1958. *A course in modern linguistics.* New York: Macmillan.

Lecarme, Jacqueline. 2002. Gender "polarity": Theoretical aspects of Somali nominal morphology. In Paul Boucher (ed.), *Many morphologies,* 109–141. Somerville: Cascadilla Press.

Moreno, Mario M. 1955. *Il somalo della Somalia.* Roma: Istituto Poligrafico dello Stato.

Orwin, Martin. 1995. *Colloquial Somali: A complete language course.* London & New York: Routledge.

Puglielli, Annarita & Cabdalla C. Mansuur. 1999. *Barashada naxwaha af Soomaaliga.* London: HAAN Associates.

Puglielli, Annarita & Ciise M. Siyaad. 1984. La flessione del nome. In Annarita Puglielli (ed.), *Aspetti morfologici, lessicali e della focalizzazione* (Studi Somali 5), 53–114. Roma: Ministero degli affari esteri.

Saeed, John I. 1993. *Somali reference grammar.* Kensington: Dunwoody Press.

Saeed, John I. 1999. *Somali.* Amsterdam: John Benjamins.

Speiser, Ephraim A. 1938. The pitfalls of polarity. *Language* 14(3). 187–202.

Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Phonology in syntax: The Somali optional agreement rule. *Natural Language and Linguistic Theory* 1(3). 385–402.

Chapter 24

The development of finite verbs from nominalized structures in Northern Mao

Michael Ahland

California State University, Long Beach

Northern Mao, an endangered Omotic-Mao language of Ethiopia, exhibits rigid OV patterns where final verbs are the most finite structures in the language. Only final verbs carry tense markers and indicate the completion of a sentence. Some final verb constructions, however, display internal structural relics that attest to a nominalized heritage: infinitive verb stems and morphological subordinators. This paper examines infinitival and subordinate structures and explores the diachronic pathways which likely led to today's finite final verb forms.

1 Introduction

Northern Mao (NM), also known by its autonym Màwés Aats'è, is an endangered Omotic-Mao language of Ethiopia (Bender 2000) spoken by an estimated 2-3,000 speakers (M. Ahland 2012: 13). NM exhibits rigid OV patterns where final verbs are the most finite structures in the language. The final declarative and interrogative verbs are the only structures which express a morphologically marked tense distinction (future /-gà/ vs. non-future /-∅/ on irrealis verbs), and the final verbs are the only structures which indicate the completion of a full sentence (also expressed by a speech act marker on final verbs). Some final verb constructions, however, display internal structural relics which attest to a nominalized heritage. These relics include infinitive verb stems, the inclusion of a relativizer, and other subordinators. That nominalizations such as these can serve as sources for finite constructions is widely attested in the world's languages (Gildea 1993; DeLancey 2011; Givón 2009: 68; Noonan 1997: 382). This paper explores NM's infinitival and subordinate structures and offers a coherent, cross-constructional analysis of the diachronic pathways through which nominal structures were transformed into today's finite final verb forms.

The discussion begins with pertinent background information including a typological overview §1.1, a discussion of finiteness in NM §1.2, the role of tone as a marker of stem finiteness §1.3, and the nominal status of infinitive stems §1.4. The discussion then details how subordination in NM involves nominalization through an exploration of the role of

infinitive stems in complement and adverbial constructions §2.1, relativization §2.2, and the /-gàʃ/ subordinator constructions §2.3. §3 illustrates the indications of nominalization and subordination in main clause structures: the use of the infinitive stem in finite final verbs §3.1, the role of relativization in the past progressive construction §3.2, and the historical development of the irrealis future verb construction from an old periphrastic subordinate + final verb construction §3.3. The discussion concludes with a summary of the evidence supporting a nominalized source for finite structures.

1.1 Typological overview

NM, like most Omotic languages, exhibits a rigid OV pattern across syntactic constructions. Table 1 lists the NM constituent-order patterns attested relative to Greenberg (1963)'s universals.

Table 1: Constituent order typology of Northern Mao (from M. Ahland 2012: 48)

Greenberg's universal	Parameter	Northern Mao pattern
1	main clause	OV
3, 4	adposition	NP - postposition
2	genitive (possessor) and head noun	Genitive - N
17	modifier and head noun	Modifier - N
24	relative clause and head noun	Relative Clause - N
22	comparatives	Standard-Marker-Quality
16	inflected auxiliaries	sentence final
9	question particles	sentence final
12	question words	sentence initial or *in situ*
27	affixes	primarily suffixes

Examples (1)-(4) illustrate selected patterns from Table 1: the order of major constituents in the transitive construction (1), the order of NP and adpositional element (postposition) (2)-(3), and the order of relative clause and head noun (4).

(1) múnts'-ìʃ   ʃó:ʃ-nà   ha-pí-⁺á.
     woman-SBJ snake-OBJ AFF-kill-DECL
     'A woman killed a snake.'

(2) tí-ŋ   ⁺kjat'-èt   há:l-⁺á.
     1SG-GEN house-LOC sleep-DECL
     'S/he slept at my house.'

(3) bàmbàs-ʃál-èt   ha-tí-kí-⁺á.
     Bambassi-way-SOURCE AFF-1SG-come-DECL

     'I came from Bambassi.'

(4) hez-ít   mùnts'-òl-ìʃ nók-and-wé-jà.
     hit:INF-REL woman-PL-SBJ be.good:INF-NSG-NEG-COP

     'Women who hit are not good.'

1.2 Finiteness in Northern Mao

Finiteness in NM is scalar; it is best viewed as matter of degree where constructions can be arranged on a continuum. This view of finiteness is in accord with Givón, who defines *finiteness* as "an aggregate feature of clauses" which involves properties of clauses as well as verbs (i.e. the licensing of arguments and other morphological structures such as case markers, and verbal inflectional morphology: subject marking, tense, aspect, modality, etc.) (Givón 2001: 25). In NM, as is common in other OV systems, final verbs are the most finite structures.[1]

First, only final irrealis verbs can express tense morphologically in NM. Irrealis verbs mark future with /-gà/ (5)-(6), and non-future tense with /-∅/ (7).

(5) ha-tjam-gà-t-bíʃ-á.
     AFF-count-FUT-1SG-NPST:AUX-DECL

     'I will count.'

(6) tjám-á-gà-t-bíʃ-á.
     count:INF-NEG-FUT-1SG-NPST:AUX-DECL

     'I will not count.'

(7) tjám-á-tí-bíʃ-⁺á.
     count:INF-NEG-1SG-NPST:AUX-DECL

     'I did not count.'

While realis verbs do not carry any morphological marker of tense, they always express non-future temporality (8), which exhibits the same scope behaviors as the morphological tense distinction on irrealis verbs.

(8) ha-tí-tjam-á.
     AFF-1SG-count-DECL

     'I counted.'

Additionally, all final verbs take one speech-act or modality (utterance type) marker which indicates the end of the NM sentence.[2] While declarative utterances are illustrated

---

[1] Contra, for instance, the less-finite, medial verbs in clause chains which do not carry tense marking or speech act/utterance-type marking; see also M. Ahland (2012: 559).

[2] For NM's complete set of utterance markers, see M. Ahland (2012: 469).

in examples (5)-(8), the following illustrate the utterance-final marking for the hearsay (9), polar interrogative (10), content interrogative (11), imperative (12)-(13), and jussive utterances (14)-(15).

(9) rám-ìʃ   ha-pò:n-gà-m-bìʃ-w-á.
     Ram-SBJ AFF-go.out-FUT-3-EXIST:INF-HRSY-DECL
     'Rama will leave (they say).'

(10) hup'-es-ìʃ   ha-húp'-à:?
     steal:INF-person-SBJ AFF-steal-INTR
     'Did a thief steal (it)?'

(11) kí-íʃ   tí-ŋ   kà:l-là   mí-à:?
     who-SBJ 1SG-GEN porridge-OBJ eat-INTR
     'Who ate my porridge?'

(12) ki:m-na   tjám-í?
     money-OBJ count:INF-2SG:IMP
     'Count the money!'

(13) íʃ   dù:l-ná   pì-wá?
     DEF hyena-OBJ kill:INF-2DU:IMP
     'You two kill the hyena!'

(14) ki:m-na   tjám-t-í-nè?
     money-OBJ count:INF-JUSS-3-NPST:AUX
     'Let him/her count the money.'

(15) tó:ló hán-jé:ts'-tà?
     now 1DU-run-JUSS
     'Let us two run now.'

1.3 Tone and stem precategoricality

Tone plays an important role in NM as a marker not only of lexical contrast, but also of stem type as either finite or non-finite.[3] In fact, the most basic building blocks for words in NM appear to be toneless roots (M. Ahland 2012: 182). These toneless roots (marked with √ in Table 2) can be grouped into three categories based on the stem types (Ahland 2014a): 1) roots which are used to form only noun stems; 2) roots which are used to form

[3] An acute accent diacritic [á] represents H tone and a grave accent diacritic [à] represents L; M tone is unmarked [a]. NM stems exhibit underlying tone classes where H and HL tonal melodies subdivide based on their behavior in different environments. These classes are represented with subscript numerals (see also M. Ahland 2012: 146).

both finite and non-finite (infinitive/nominal) verb stems (Table 2); and 3) roots which are used to form distinct noun stem, finite verb stem and infinitive verb stems (Table 3).

The vast majority of nominal forms (which are bi-syllabic) distribute into seven nominal stem melodies: H, M, L, HL, MH, ML, LH. These melodies serve as the primary marker of nominal status of a stem and are used for prototypical nouns as well as infinitive verb stems. Most finite verb stems, on the other hand, exhibit either a H, M, or L melody. The only exceptions to these three finite verb melodies involve the less-common verb stems which are more than a single syllable; in all cases, though, the finite verb melody of a given root is distinct from the nominal stem melody for the same root.

Table 2 lists a sampling of roots which can serve as finite and infinitive verb stems. Many roots belong to this category. Clearly, some of the infinitive verb stems in Table 2 can also be used as nouns as well as infinitives (see illustrative examples below Table 2).

Table 2: Tone classes of selected roots with distinct finite and infinitive verb stems

	Root	Finite verb stem	Tone class	Infinitive verb stem			
				Used as infinitive verb		Used as noun	Tone class
1	√*kjat'*	*kját'* 'build. house'	H	*kjat'-è* build.house:INF-TV		*kjat'-è* house-TV	ML
2	√*kjal*	*kjàl* 'lay.egg'	L	*kjál-è* lay.egg:INF-TV		*kjál-è* egg-TV	HL$_2$
3	√*go:m*	*gó:m* 'think/plan'	H	*go:m-è* think/ plan:INF-TV		*go:m-è* thought/ plan/ idea-TV	ML
4	√*int'*	*int'* 'see'	M	*ínt'-é* see:INF-TV		*ínt'-é* sight-TV	H$_1$

In (16), the root √*kjat'* is tonally marked as an infinitive verb stem (used as a noun with object case marking);[4] it is also tonally marked as a finite final verb at the end of the utterance.

(16) kjat'-nà   ha-tí-kjat'-ꜛá.
     house-OBJ AFF-1SG-build.house-DECL

     'I built a house.'

---

[4] Terminal vowels are not found in many syntactic environments, including when case markers are present (cf. M. Ahland 2012: 313).

Example (17) illustrates the full nominal status of the noun 'house'. It can be possessed, marked with the PL number suffix, and take case.

(17) ham kjat'-wol-iʃ   ha-bíʃ-and-á.
      1PL house-PL-SBJ AFF-EXIST-NSG-DECL
      'There are our houses.'

Example (18) illustrates this same root used as an infinitive stem (marked tonally the same as the noun stem, but with a verbal purpose-subordinator suffix).

(18) kjat'-gàʃ   ha-tí-wóːl-⁺á.
      build.house:INF-PURP AFF-1SG-want-DECL
      'I want to build a house.'

The root √*kjal* 'egg' behaves similarly: (19) has three instances of this root. The first instance is marked as an infinitive verb form and is used as a noun with object case marking; the second instance of the same root is used as a finite purpose (subordinate) verb stem; the third instance is a finite final verb.

(19) waːk-iʃ   kjál-nà   kjàl-gàʃ   hí-in   ha-kjàl-á.
      chicken-SBJ egg:INF-OBJ lay.egg-PURP go-SS:NF AFF-lay.egg-DECL
      'A chicken went to lay an egg and (then) laid (an egg).'

Perhaps most interesting of all are those roots which take three distinct forms for the three distinct stem types: finite verb stems, infinitive verb stems, and noun stems--each marked with a different tonal melody (Table 3).

Table 3: Tone classes of selected roots with distinct finite, infinitive, and noun stems

	Root	Finite verb stem	Tone class	Infinitive verb stem	Tone class	Noun stem	Tone class
1	√*toːk*	*tóːk* 'head.carry'	H	*tòːk-è* head.carry:INF-TV	L	*toːk-è* head-TV	ML
2	√*p'iʃ*	*p'íʃ* 'give.birth'	H	*p'iʃ-é* give.birth:INF-TV	MH	*p'iʃ-e* child-TV	M
3	√*but'*	*bùt'* 'be.afraid'	L	*bùt'-é* fear:INF-TV	LH	*bút'-è* shame/fear-TV	HL₂
4	√*git*	*gìt* 'cover'	L	*gít-è* cover:INF-TV	HL₂	*gìt-é* mask/cover-TV	LH

The examples below illustrate use of the √*toːk* root as a noun stem (20)-(21), infinitive verb stem (22), and finite verb stem (21), (23).

(20) íʃ ⁺to:k-ìʃ ha-maŋk'-á.
     3SG head-SBJ AFF-hurt-DECL
     'Her/his head hurts.' (e.g. she/he has a headache)

(21) to:k-èt   ha-tó:k-⁺á.
     head-LOC AFF-head.carry-DECL
     'S/he carried it on (her/his) head.'

(22) tò:k-gàʃ   ha-tí-wó:l-⁺á.
     head.carry:INF-PURP AFF-1SG-want-DECL
     'I want to carry it (by head).'

(23) k'ò:p-ná   tó:k-gàʃ   ha-tí-wó:l-⁺á.
     basket-OBJ head.carry-PURP AFF-1SG-want-DECL
     'I want to carry a basket (by head).'

1.4 The nominal status of infinitive stems

Infinitive stems in NM can function as nouns (as illustrated in (15) and (16), above). In addition, infinitive stems are more noun-like than verb-like in terms of word-structure and constructional distribution. Prototypical nouns, which are semantically more time-stable, in NM exhibit a number of structural and distributional features which can serve as tools for identifying and determining degrees of nominalization of less prototypical nouns, such as infinitive stems.

The discussion begins with word-structure. First, two-syllable nouns, which are by far the most common shape attested in the NM lexicon, exhibit one of seven surface tone melodies: H, M, L, HL, MH, ML, or LH (Ahland 2009: 23). Second, all nominals (including pronouns, demonstratives, and the definite article) end in a terminal vowel /-e/ when in citation form and at the ends of utterances or main clauses (M. Ahland 2012: 194, 313). This same vowel is seen on infinitive stems in citation form in Tables 3 and 4. Third, prototypical nouns can take number marking (/-kuw/ DU and /-(w)ol/ PL). While the meaning of some infinitive stems does not allow for countability, other infinitive stems can be marked for number without any additional change from infinitive stem to noun, as seen in (17) and (18), above. Finally, all nominals exhibit a tonal change when syntactically modified. The modified head takes on a lexically-specified construct form/state melody (either M, ML, or L, depending on the citation melody of the noun) (M. Ahland 2012: 146). Table 4 lists the tonal pattern changes from citation to construct melody for two-syllable nouns.[5]

---

[5] Nouns with H citation melodies divide into two classes: those with an M construct melody and those with an L construct melody. Similarly, nouns with HL citation melodies divide into two classes: one with an ML and one with an L construct melody.

Table 4: Citation vs. construction form melodies of nouns

Citation melody	Construct form melody
H$_1$	> M
M, L, HL$_1$, MH, ML	> ML
H$_2$, HL$_2$, LH	> L

Examples (24) and (25) illustrate that both noun stems and infinitive stems exhibit the citation vs. construct melody behavior. In these examples, roots from Table 3 (rows 1 and 3) are used to form noun and infinitive stems. In each instance the citation melody (of the noun and infinitive stems; see Table 3) is replaced with a lexically-specified construct melody.

(24) Citation vs. construct melodies
   N stem from ML > ML   INF stem from L > ML
   nà   to:k-è   íʃ   tò:k-è
   PROX   head-TV   3SG   head.carry:INF-TV
   'this head'   'his/her carrying (by head)'

(25) Citation vs. construct melodies
   N stem from HL > L   INF stem from LH > L
   nà   bùt'-è   íʃ   bùt'-è
   PROX   shame/fear-TV 3SG   shame/fear:INF-TV
   'this shame/fear'   'his/her fearing'

NM's infinitive forms are also similar to nouns in terms of distribution. The examples below illustrate three common noun modification constructions: associative (26), genitive (27), and possessive (28).

(26)   kús-kás-è
    hand-hoe-TV
    'hand-hoe'

(27)   es-ìŋ   kas-è
    person-GEN hoe-TV
    'person's hoe'

(28)   es   kas-è
    person   hoe-TV
    'person's hoe'

Infinitive stems, as well, can occur in the position most frequently associated with nouns in these modification constructions, including associative (29), possessive (30) and

genitive (31). In each of these constructions, the infinitive stems carry subject case markers. Example (32) illustrates that infinitive stems can also be marked for dual number—further underscoring the nominal status of the stem.

(29) k'o:m-kán-ìʃ ha-hék'-ꜜá.
be.old:INF-dog-SBJ AFF-die-DECL
'The old dog died.'

(30) ham k'o:m-ìʃ ha-hék'-ꜜá.
1PL be.old:INF-SBJ AFF-die-DECL
'Our old (one) died.'

(31) íʃ-ìŋ ak-ìʃ ha-k'úm-ꜜek'-á.
3SG-GEN eat:INF-SBJ AFF-forbid-PASS-DECL
'Her/his eating was forbidden.' (i.e. as during a fast)

(32) tí-ŋ wa:r-kuw-iʃ k'o:m-kuw-e.
1SG-GEN clothing-DU-SBJ be.old:INF-DU-TV
'My two pieces of clothing are old ones.'

2 Subordination as nominalization

Subordination in NM very frequently involves various types of nominalization. These nominalizing processes are important for understanding the development of finite main clause syntax. For the purposes of this paper, *nominalization* is a derivational process whereby an element takes on properties associated with the nominal category in a language. Since at least some NM roots are precategorical, the term *nominalization* includes the formation of a nominal stem via the assignment of a nominal tone melody and the terminal vowel. *Subordination*, as it is used here, involves the embedding of a clause (minimally a verb) within another phrase or clause. Subordinate clauses may be as simple as an infinitive verb stem, a headless relative clause or complex elements which exhibit their own internal tense and aspect, subject, case and other morphological marking. Before examining relics of historical nominalization and subordination in main clause syntax, it is first important to examine properties of nominalized/subordinate clauses synchronically.

2.1 Infinitive stems in complement and adverbial constructions

The infinitive stem is the simplest deverbal nominalization structure in Northern Mao; infinitive stems are commonly used for action and participant nominalization (M. Ahland 2012: 222). Examples (33)-(35) illustrate simple subject complements with infinitive stems.

(33) ki-iʃ   ha-nok-á.
     come:INF-SBJ AFF-be.good-DECL
     'Coming is good.' (I'm glad you're coming.)

(34) ká:nd-óʃk-ak-iʃ   ha-k'úm-⁺ek'-á.
     pig-meat-eat:INF-SBJ AFF-forbid-PASS-DECL
     'Eating pork is forbidden.'

(35) tí-ŋ   kjamb-ìʃ   ha-nok-á.
     1SG-GEN hunt:INF-SBJ AFF-be.good-DECL
     'My hunting is good.'

Examples (36) and (37) illustrate object complements, formed from the finite stem (36) and the infinitive stem (37).

(36) bàmbàs-ʃál-nà   jé:ts'-nà ha-tí-inint'-⁺á.
     Bambassi-way-OBJ run-OBJ AFF-1SG-try-DECL
     'I tried running to Bambassi.'

(37) jè:ts'-nà   ha-tí-ts'é:l-⁺á.
     run:INF-OBJ AFF-1SG-finish-DECL
     'I finished running.'

A finite verb stem in a complement construction can take an object (36), (38), (39).[6] It may be that constructions like (36), (38), and (39) are nominalizations of a verb phrase (M. Ahland 2012: 625).

(38) [pàl-là   kjámb]-ìʃ ha-nok-á.
     antelope-OBJ hunt-SBJ AFF-be.good-DECL
     'Hunting antelopes is good.'

(39) [bàmbàs-ʃál-nà   jé:ts']-nà ha-tí-wó:l-⁺á.
      Bambassi-way-OBJ run-OBJ AFF-1SG-want-DECL
     'I want to run to Bambassi.'

In NM, the choice of infinitive vs. finite verb stem in subordinate structures correlates with several finiteness-related features: infinitive stems are used only with subordinate verbs which are not overtly marked for subject (i.e. which share the same subject as their matrix verb). Finite stems, on the other hand, may be used on same-subject subordinate structures, but they are required on subordinate verbs which carry overt subject markers or which include some internal structure (e.g. an object NP, oblique, or adverbial) in the subordinate clause (cf. M. Ahland 2012: 612, 630).

---

[6] Movement predications appear to be a subset of transitive constructions in NM; as a result, semantic Goals are marked with the /-na/ OBJ case marker (cf. M. Ahland 2012: 335-339).

2.2 Relativization

Relativization is a nominalizing process in NM (M. Ahland 2012: 225). Relativized verbs are marked with the suffix /-(i)t/. The simplest and most highly nominal relative clause is headless—i.e. it does not modify a head noun; this form is subject-less. It can carry case in relevant syntactic environments (40)-(41) and can be marked for number (41). Example (42) illustrates the non-relativized finite matrix verb 'come.'

(40) ki-t-ìʃ   oʃk-nà   ha-ak-á.
      come:INF-REL-SBJ meat-OBJ AFF-eat.meat-DECL

      'The one who came ate meat.'

(41) ki-t-wol-iʃ   oʃk-nà   ha-ak-and-á.
      come:INF-REL-PL-SBJ meat-OBJ AFF-eat.meat-NSG-DECL

      'The ones who came ate meat.'

(42) íʃ-kol-té ha-kí-wand-á.
      3-PL-SBJ AFF-come-NSG-DECL

      'They came.'

Relativization makes use of both infinitive and finite verb stems, depending on the degree of finiteness of the nominalized verb phrase. If the relativized verb includes an overt subject prefix or if the relativization includes more than just the verb (such as an object NP or some other element, i.e. the relativization of a clause), the finite verb stem is required (43)-(44).

(43) [íʃ   mùnts'-ìʃ   ki:m-na   hí-tà]-t   es-ìʃ   ha-kí-⁺á.
       DEF woman-SBJ money-OBJ 3SG-give-REL person-SBJ AFF-come-DECL

      'The person to whom to the woman gave money came.'

(44) [kà:l-là   mí]-t   es-ìʃ   ha-kí-⁺á.
       porridge-OBJ eat-REL person-SBJ AFF-come-DECL

      'The person who ate the porridge came.'

That relativization is a nominalizing process is shown by the fact that the resulting forms can serve as arguments, take case, and can be marked for number. They can also be marked with the terminal vowel found on all nominals in citation form and utterance final positions (45).

(45) íʃ   mùnts'-ìʃ   kwats'-ít-è.
      DEF woman-SBJ be.tall:INF-REL-TV

      'The woman is tall.'

2.3 The /-gàʃ/ subordinator

As mentioned earlier, the term *subordination* here encompasses all embedded clauses, whether relative, complement, or adverbial. Like relativization, some other forms of subordination show properties of nominalization. One subordinate clause marker /-gàʃ/ indicates highly nominalized complements (46)-(48) and purpose adverbials (49).[7]

(46) mì-gàʃ-nà ha-tí-nó:k-⁺á.
  eat:INF-COMP-OBJ AFF-1SG-begin-DECL

  'I began to eat.'

(47) mì-gàʃ-nà ha-tí-ts'é:l-⁺á.
  eat:INF-COMP-OBJ AFF-1SG-finish-DECL

  'I finished eating.'

(48) mì-gàʃ-nà ha-tí-wó:l-⁺á.
  eat:INF-COMP-OBJ AFF-1SG-want-DECL

  'I want to eat.'

(49) mì-gàʃ tí-kí-ti-á.
  eat:INF-PURP 1SG-come-PF-DECL

  'I have come in order to eat.'

If there is a different subject (relative to the matrix clause) on the subordinate verb or if the subordinate clause also contains object NPs or adverbials, the finite verb stem must be used. Even in these instances, there is no tense or aspect marking on the subordinate verb (50)-(51).

(50) kà:l-là hí-mí-gàʃ-nà ha-tí-wó:l-⁺á.
  porridge-OBJ 3SG-eat-COMP-OBJ AFF-1SG-want-DECL

  'I want her/him to eat porridge.'

(51) hí-mí-gàʃ tí-kí-ti-á.
  3SG-eat-PURP 1SG-come-PF-DECL

  'I have come in order for him to eat.'

3 Traces of nominalization and subordination in main clause syntax

The discussion thus far has focused on necessary preliminaries including evidence of synchronic subordination as nominalization. At this point, the focus turns to the primary issue: evidence of historical nominalization in main clause development. As noted

---

[7] The object case marker is optional in all instances except when objects appear in non-canonical positions (e.g. before the subject) (M. Ahland 2012: 327).

in the introduction, some final verb constructions in NM exhibit structures which attest to a history of nominalization: infinitive verb stems, old relativizers, and subordinators are footprints of earlier nominalization in today's finite structures. The NM data demonstrate yet another instance of how nominalizations or infinitive structures/less finite structures can be reanalyzed and become associated with finite functions (cf. Givón 2009; DeLancey 2011).

3.1 The use of the infinitive stem in synchronic finite final verbs

Figure 1 reflects the fact that infinitive stems are used in various types of verbal constructions, ranging from the more nominalized subordinate complements, to final verbs that do not take tense marking (e.g. imperatives/jussives), to negative declarative and interrogative final verbs, which, apart from their required infinitive stems, appear to be otherwise fully finite.
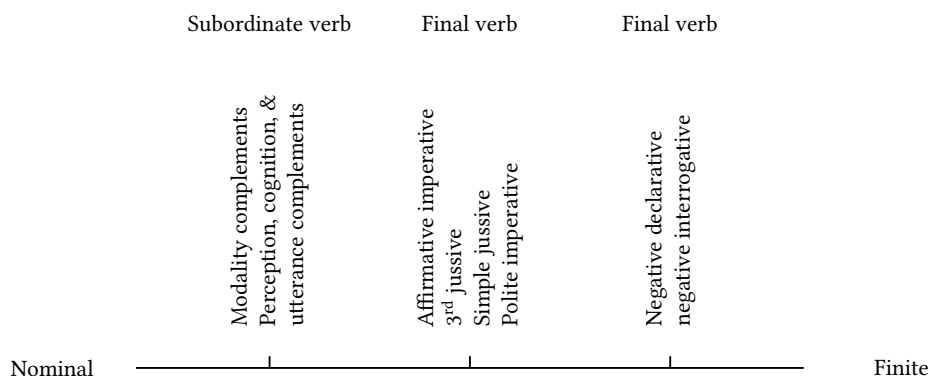
Figure 1: A nominal-finite continuum of infinitive verb stem usage

In declarative and interrogative final verbs, infinitive stems are only found in negatives (which always involve the irrealis verbal construction); see Figure 2 (reproduced from M. Ahland 2012: 266).[8] Affirmative realis (non-future) and irrealis (future) verbs require finite stems. Many affirmative imperative and jussive forms require infinitive verb stems, indicated by the boxed section in Figure 2. The reason for this likely involves differing grammaticalization pathways. There must be a reason why the infinitive stem requirement is only associated with negative constructions that utilize the irrealis verb (regardless of morphological tense marking), with certain imperative and jussive constructions, and with most highly nominalized subordinate verbs. This distribution attests to a nominalized history for these constructions.

---

[8] The realis and irrealis verb constructions are formed with different item-arrangements in terms of morphology—see §3.3, below. The left-most ellipse in Figure 2 lists features associated only with the realis verb form, while the right-most ellipse lists features associated with the irrealis verb form. The shared center lists those features which are compatible with both realis and irrealis verb forms. Additional verb constructions which do not involved the realis nor the irrealis verb form are listed at the bottom of Figure 2.
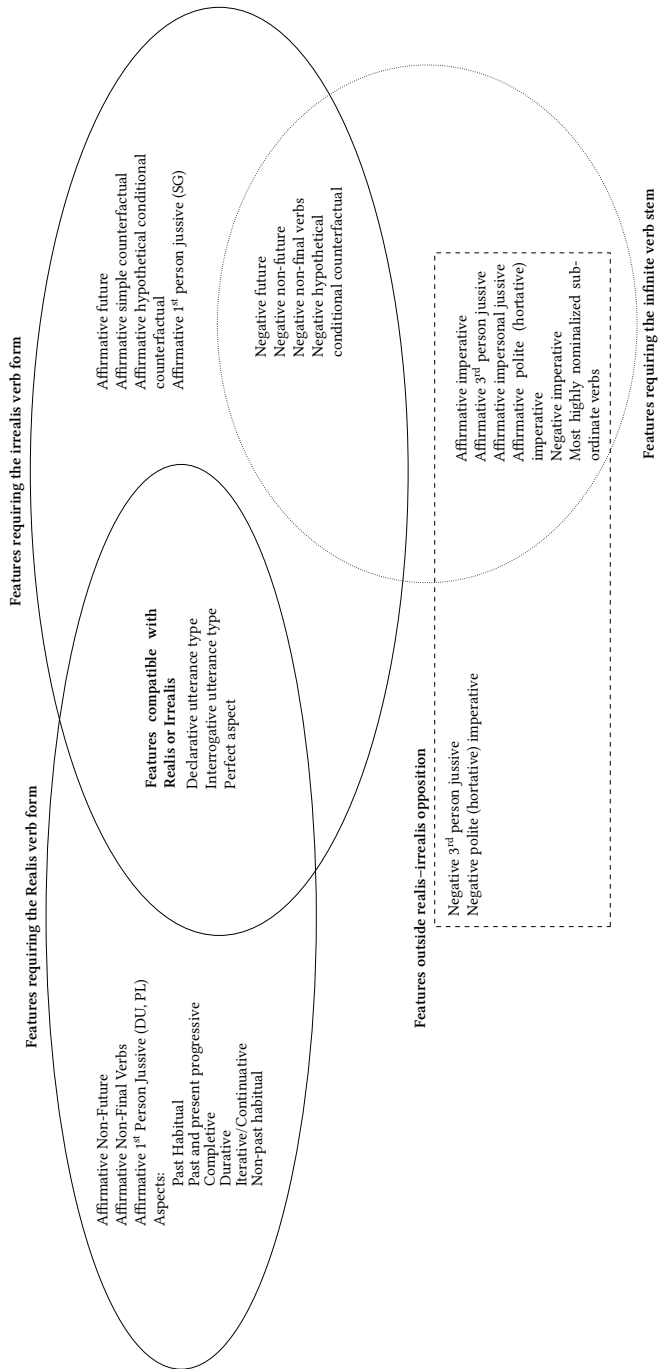
Figure 2: The distribution of infinitive verb stems and the intersection with the realis-irrealis opposition

Before turning to the use of the infinitive verb stems in finite final verbs, use of the finite verb stem is briefly illustrated: affirmative declarative non-future (realis) (52), declarative future (irrealis) (53), interrogative non-future (54) and interrogative future (55) (both irrealis).

(52) ha-tí-héz-⁺á.
      AFF-1SG-hit-DECL
      'I hit (it).'

(53) ha-héz-gà-t-bíʃ-á.
      AFF-hit-FUT-1SG-NPST:AUX-DECL
      'I will hit (it).'

(54) tí-héz-à: / ha-tí-héz-â:?
      1SG-hit-INTR   AFF-1SG-hit-INTR
      'Did I hit (it)?' 'Did I hit (it)?'
      (expecting an affirmative response)

(55) héz-gà-t-bíʃ-à:?
      hit-FUT-1SG-NPST:AUX-INTR
      'Will I hit (it)?'

3.1.1 The negative construction

Unlike the affirmative declarative and interrogative examples in (52)-(55) which require the finite verb stem, examples (56)-(60) illustrate the corresponding negative constructions (non-future and future) which require the use of the infinitive stem.

(56) hez-á-tí-bíʃ-á.
      hit:INF-NEG-1SG-NPST:AUX-DECL
      'I did not hit (it).'

(57) hez-á-gà-tí-bíʃ-á.
      hit:INF-NEG-FUT-1SG-NPST:AUX-DECL
      'I will not hit (it).'

(58) hez-á-hì-biʃ-à:?
      hit:INF-NEG-2SG-NPST:AUX-INTR
      'Did you not hit (it)?'

(59) hez-á-gà-t-bíʃ-à:?
      hit:INF-NEG-FUT-1SG-NPST:AUX-INTR
      'Will I not hit (it)?'

Apart from the requirement of the infinitive verb stem, the negative final verbs appear to be fully finite today (e.g. exhibiting a tense distinction and utterance type/speech act marking).

The organization of the internal morphology of the negative verb in (56)-(59) suggests that it developed from a periphrastic structure like (60). The bifurcation of negatives into future and non-future constructions appears to be a later development and is discussed in §3.3, following the discussion of the affirmative future's development.

(60) Hypothesized source for negatives
    Subordinate structure   Final verb
    [INF stem]+-á $_{SUB}$ SBJ-Existential.verb-DECL
    hez   -á   tí-   bíʃ   -á.
    hit:INF   -NEG   1SG-   EXIST   -DECL
    'I did not hit (it).' (< I was not hitting.) (from ex. 56)

The negative verb appears to have begun as a nominalized complement (an infinitive with a subordinating morpheme that was later reanalyzed as a negative marker) followed by a fully-finite final existential verb with subject prefixes.[9] This periphrastic structure presumably then collapsed into the single verbal word observed today (see §3.3).

3.1.2 Imperative and jussive utterances

As noted in Figure 2, infinitive stems are also required in certain imperative and jussive constructions. Unlike the other constructions we have seen, however, the infinitive stems in (61)-(63) are used in affirmative constructions.

(61)   hez-í!
        hit:INF-2SG:IMP
        'Hit (it)!'

(62)   há-hez-í!
        IMPR-hit:INF-2SG:IMP
        'Hit (it)!' (polite)

(63)   há-hez-t-í-nè.
        IMPR-hit:INF-JUSS-3-NPST:AUX
        'Let it hit (it).'

Infinitive and jussive forms do not carry tense or aspect marking in Northern Mao—which all other final verb forms do carry. That said, however, infinitive forms do all carry subject markers which also serve to indicate the utterance type. These verbs, like all other final verbs, indicate the end of a sentence, unless a coordinating conjunction is employed.

---

[9] The use of lexical verbs with infinitive stems in auxiliary-headed constructions is quite common cross-linguistically (Anderson 2006: 56).

With respect to the jussive construction in (63), the /ne/ form in the impersonal jussive (63) is a recognizable copula in Northern Mao (64) (M. Ahland 2012: 463).

(64) wè:ŋk'   nè.
      be.open:INF be.NPST
      'It is open.'

The /ʔá-/ impersonal prefix in the impersonal jussive construction in (63), as well as on the polite imperative in (62), may derive from an old demonstrative which gave rise to a 3ʳᵈ person marker in other Mao languages (Bender 2000: 206; M. Ahland 2012: 245-246; (Ahland 2015));[10] the /-t/ JUSS is likely derived from or related to the /-(i)t/ relativizer (reanalyzed as a jussive construction marker). The hypothesized source construction in (65) illustrates the demonstrative + relativized infinitive stem (nominal) as a predicate, which is then followed by a copula.

(65) Hypothesized source for jussive construction
      Modified Headless Relative (nominal)   Final verb
      DEM [[INF]+(i)t]ᵣₑₗ   SBJ-Copula
      há   hez   -t   í-   nè.
      DEM   hit:INF -REL   3-   COP
      'It is that which hit.' > 'Let it hit (it).' (similar to ex. 63)

The jussive source in (65) is represented as having been a headless relative clause (i.e. a nominal). The fact that the relativized stem is in the infinitive form suggests that the demonstrative would very likely not have been an argument of the relative clause. Perhaps the /há/ demonstrative was a modifier of the relativized predicate.

The imperative construction exhibits the only final verb in the language which uses an infinitive stem where there is no other recognizable verb stem.[11] It is not clear if an older verbal stem atrophied and disappeared or if this represents a different grammaticalization pathway.

The imperative construction (66) is unique in NM in that both the affirmative and negative polarities require the infinitive verb stem.

(66) hez-áʃ-í!
      hit:INF-NEG-2SG:IMP
      'Do not hit (it)!'

3.2 Relativization and the past progressive construction

The progressive constructions involve periphrastic constructions where lexical verbs are followed by auxiliary verbs. In the past progressive, the lexical verb obligatorily carries the /-(i)t/ relativizer (67)-(68).

---

[10] Reflexes can be found throughout the Mao subgroup and the form is still used in Seezo /hán/ (Mengistu 2015); a reflex is also found in Hozo /ʔá/ 3SG masculine (Kassa 2014).

[11] In the impersonal jussive, for example, the /nè/ form itself appears to derive from a copula verb.

(67) hí-kí-èt ha-tí-mí-t bitè.
3SG-come-TI:NF AFF-1SG-eat-REL PST:AUX

'I was eating when s/he came.'

(68) ha-jé:ts'-ìt bitè.
AFF-run-REL PST:AUX

'S/he was running.'

In fact, in the past progressive construction there are two cases of historical relativization. The past copula form /bitè/ is itself a relativization of the existential verb /biʃ/ (cf. M. Ahland 2012: 318; 461-462, and the highlighted form in (69)): /biʃ -t/ > /bi-t/ EXIST-REL. In (69)'s matrix clause, the cleft with present meaning requires no copula (i.e. a zero copula); while in (70), the past copula is used (in this instance, the form /bitè/ is not a functional auxiliary).

(69) [[nà-àt **bi-t**]ᵣᴄ ʃak'-è]ₙₚ Ø [tí-pí-t-ìʃ-é.]
here-LOC EXIST:INF-REL goat-TV be.PRES 1SG-kill-REL-SBJ-TV

'It is the goat that was here that I killed.'

(70) [[nà-àt **bi-t**]ᵣᴄ ʃak'] ₙₚ **bitè** (< bi-t-è) [tí-pí-t-ìʃ-é.]
here-LOC EXIST:INF-REL goat be.PST (< EXIST:INF-REL-TV) 1SG-kill-REL-SBJ-TV

'It was the goat that was here that I killed.'

The past copula form is a frozen form like all other Northern Mao copulas today (M. Ahland 2012: 465).

Interestingly, the non-past progressive construction appears to have developed via a slightly different pathway: 1) there is no hint of relativization and 2) the auxiliary is the non-relativized existential (71).

(71) to:lo ha-tí-mí biʃ-á.
now AFF-1SG-eat NPST:AUX-DECL

'I am eating now.'

In sum, the non-past progressive shows no overt morphological indication of nominalization in its history. But the past progressive construction does and may have developed from a nominal + nominal predication (relativization + relativization), as in (72):

(72) Hypothesized source for past progressive
Relative clause   Relative clause
[SBJ-finite] + -(i)tᵣₑₗ   [EXIST INF] + -(i)tᵣₑₗ + -TV
ha-tí-mí   -t   bi   -t   -è.
AFF-1SG-eat   -REL   EXIST:INF   -REL   -TV

'I was eating.'

Alternatively, it could be that the /bitè/ auxiliary was already grammaticalized as a past copular form — it is difficult to know whether the form was still functioning as a relative clause at the time the past progressive developed. Either way, the lexical verb was clearly expressed, as it is today, within a preceding (nominalized) relative clause.[12]

3.3 The irrealis future verb construction

Northern Mao's final future (irrealis) verb also developed from a nominal subordinate clause. The lexical verb was marked with a subordinator and was followed by a final existential verb (an AUX-headed construction; see Anderson 2006: 39ff). The subordinator was then reanalyzed as a future tense suffix. The entire periphrastic construction collapsed phonologically, resulting in a complex final verb with contained internal relics of subordinator and related subject marking. (See Ahland 2014b for further details.)

Figure 3 and example (73) illustrate the structure of the realis verbal word.

Inflectional prefixes		Finite stem		Derivational suffixes		Inflectional suffixes				
(Affirmative) $\left(\begin{array}{c}\text{Subject}\\\text{prefix}\end{array}\right)$				$\left(\begin{array}{c}\text{Valence}\\\text{decreasers}\end{array}\right)$ (Applicative)		(Perfect) $\left(\begin{array}{c}\text{Non-}\\\text{singular}\end{array}\right)$	$\left(\begin{array}{c}\text{Past}\\\text{habitual}\end{array}\right)$	(Hearsay)	$\begin{array}{c}\text{Utterance}\\\text{type}\end{array}$	

Figure 3: Realis verb item-arrangement (from Ahland 2014b: 3)

(73) kwalla   ha-tí-jé:ts'-⁺á.
      yesterday AFF-1SG-run-DECL
      'I ran yesterday.'

Figure 4 and example (74) illustrate the structure of the irrealis verbal word.

Inflectional prefix		Finite stem		Derivational suffixes		Inflectional suffixes					
(Affirmative)				$\left(\begin{array}{c}\text{Valence}\\\text{decreasers}\end{array}\right)$ (Applicative)		(Perfect) $\left(\begin{array}{c}\text{Non-}\\\text{singular}\end{array}\right)$	$\begin{array}{c}\text{Future}\\\text{suffix}\end{array}$	$\begin{array}{c}\text{Subject}\\\text{suffix}\end{array}$	Auxiliary (Hearsay)	$\begin{array}{c}\text{Utterance}\\\text{type}\end{array}$	

Figure 4: Irrealis future affirmative item-arrangement (from Ahland 2014b: 4)

Example (74) illustrates the irrealis verbal word.

(74) háts'á   ha-jé:ts'-gà-t-bíʃ-á.
      tomorrow AFF-run-FUT-1SG-NPST:AUX-DECL
      'I will run tomorrow.'

The primary differences between the realis and the irrealis verbal words involve the position of subject markers (prefixes on realis verbs and suffixes on irrealis verbs), as well as the presence of an auxiliary verb on the irrealis verbal word.

---

[12] As noted in §2.2, the use of subject marking on relativized verbs requires the use of the finite stem even though the clause is still nominalized via the relativizer.

In order to understand the history of the irrealis verb, it is important to note that today's future tense suffix /-gà/ is similar to a purposive subordinate marker /-gàʃ/ found on purpose adverbials (75), as well as on some complements (76) (cf. §2.3 above).

(75) [kà:l-là   mí]-gàʃ tí-kí-ti-á.
      porridge-OBJ eat-PURP 1SG-come-PF-DECL

      'I have come to eat porridge.'

(76) mì-gàʃ-nà   ha-tí-ts'é:l-ꜜá.
      eat:INF-COMP-OBJ AFF-1SG-finish-DECL

      'I finished eating.'

It has been argued that purpose clauses are "intrinsically future-oriented" (Schmidtke-Bode 2009: 43). In fact, purpose clauses and their matrix verbs are an attested source of future constructions across many of the world's languages (p.181). These complex constructions undergo grammaticalization such that the matrix verb becomes an auxiliary.[13] This is what appears to have taken place in NM's irrealis future verb construction.

Before discussing the reconstructed source construction for NM's irrealis future verb, it should be noted that the future tense suffix /-gà/ appears to have been /-gàm/ at an earlier stage in its development. Internal irregularities found in the future relative clause preserve what appears to be an older [m], i.e. /-gàm/ as in (77) (cf. Ahland 2014b: 10).

(77) ha-tí-kí-gàm-b-t
      AFF-1SG-come-FUT-NPST:AUX-REL

      'that I will come'

This old [m] is a significant part of the development of the irrealis future verb: the [m] has impacted the subject marking paradigm by intruding into 3rd person and 2SG marking in all constructions where today's future tense marker is found; see the shaded cells of Table 5.

The distribution of the intrusive [m] relative to subject marking is discussed following introduction of the reconstructed source construction.

As noted above, complex constructions of the type purpose clause + matrix verb are an attested source of future tense constructions. The presence of a relic subordinator positioned before an old subject-marked existential (> auxiliary verb) in the irrealis future construction suggests that irreality (and future temporality) was expressed through a periphrastic construction like the reconstruction in (78) (adapted from Ahland 2014b: 11).[14]

---

[13] Bybee, Perkins & Pagliuca (1994: 263-75) note how purpose clauses and concepts such as 'intention' and 'desire' are "compatible or harmonic" with the grammaticalization of future tense.

[14] Since affirmative future irrealis verbs exhibit only finite verb stems today, the /-gàm/ form must have attached only to the finite form. Today, however, both finite and infinitive stems can take the /-gàʃ/ subordinator, depending on the degree of finiteness of the subordinate clause (cf. M. Ahland 2012: 629).

Table 5: Free pronouns and subject marking on final verbs (Ahland 2014b: 7)

Free pronouns		Realis prefixes	Irrealis non-future (NEG) suffixes	Irrealis future (AFF and NEG) suffixes
1SG	tí-jé	tí-	-tí	-t́
1DU	han-é	hań-	-ń	-ń
1PL	hambèl-è	haṁ-	-ṁ	-ṁ
2SG	hì-jè	hì-	-hì	-èm
2DU	háw-é	háw-	-ẃ	- ´ (H Tone)
2PL	hàwèl-è	hàw-	-ẁ	- ` (L Tone)
3SG	íʃ-è	Ø-	-Ø-	-ṁ
3DU	íʃ-kuw-e	Ø- /-and/	-Ø- /-and/	-ṁ /-and/
3PL	íʃ-kol-è	Ø- /-and/	-Ø- /-and/	-ṁ /-and/

(78) Hypothesized source for future irrealis
    Subordinate structure   Final verb
    kí-gàm   ∅-bíʃ-⁺á.
    come-SUB/PURP   3SG-EXIST-DECL

    Reconstructed meaning: 'S/he is in order to come.'
    (Reconstructed effective force: 'S/he intends to come.')

The old purpose subordinator /-gàm /was reanalyzed as a future tense marker. Its final [m] was lost before all initial consonants (except for [h]) on the following subject marker (probably as the entire phrase collapsed into a verb) (79)-(80). The motivation for such a loss could have been consonant cluster simplification. Certainly three-consonant sequences of this sort in NM are not attested.

(79) ha-kí-gà-**t**-bíʃ-á.
     AFF-come-FUT-1SG-NPST:AUX-DECL
     'I will come.'

(80) ha-kí-gà-**n**-bíʃ-á.
     AFF-come-FUT-1DU-NPST:AUX-DECL
     'We (dual) will come.'

In the instances where [m] was not lost due to consonant cluster simplification, it became associated with subject person marking (81)-(82).

(81) ha-kí-gà-**m**-bìʃ-á.
     AFF-come-FUT-3-NPST:AUX-DECL
     'S/he will come.'

(82) ha-kí-g-**èm**-bìʃ-á.
      AFF-come-FUT-2SG-NPST:AUX-DECL
      'You (singular) will come.'

These facts suggest that the nominal history of the irrealis future verb impacted not only the verbal construction itself, but also the subject marking paradigm.[15] This impact includes the subject markers found on the negative future verb, which also shows the same intrusive [m] interference as the affirmative future (Table 5; and M. Ahland 2012: 385).

After the reanalysis of the /-gàm/ form (from subordinator > FUT suffix) and the formation of the new subject markers, the future tense marker, along with the intrusive [m] in the subject marking, was analogically extended to the negative periphrastic construction (60), resulting in two distinct forms: a negative future and negative non-future. The negative non-future was left unchanged.

If the analysis of the /-á/ negative marker deriving from a subordinator is accurate, there would have been no need for another subordinator such as /-gàm/ (before its reanalysis) on the lexical infinitive form in the negative source construction. This supports the hypothesis that the extension took place after reanalysis. Also, the morphological differences between the negative future and non-future verb forms suggest that the collapse of the non-future periphrastic construction to a single verbal word occurred after the future/non-future distinction developed. While the negative future matches the affirmative future in subject marking and auxiliary patterns (compare (82) with (83)), the negative non-future construction exhibits irregularities: an unreduced 1SG marker (84), an unexpected downstep on the declarative suffix (84)-(85), and an altogether different negative suffix and copular auxiliary for 3rd person forms (86) (see M. Ahland 2012: 384-384 for complete details).

(83) ki-á-g-**èm**-bìʃ-á.
      come:INF-NEG-FUT-2SG-NPST:AUX-DECL
      'You (singular) will not come.'

(84) ki-á-**tí**-bíʃ-ꜜá.
      come:INF-NEG-1SG-NPST:AUX-DECL
      'I didn't come.'

(85) ki-á-**hì**-bìʃ-ꜜá.
      come:INF-NEG-2SG-NPST:AUX-DECL
      'You (singular) didn't come.'

---

[15] This new, altered subject marking paradigm (with the intrusive [m] from the old subordinator) was even extended into new verbal constructions. Today's purpose / complement subordinator /-gàʃ/ perhaps similarly lost the old [m] as a result of cluster simplification before the initial consonant of a following relational noun / postposition /-ʃal/ 'way' (cf. Ahland 2014b: 14).

(86)   ki-wé-jà.
       come:INF-NEG-COP
       'S/he did not come.'

4   Final considerations

The NM data provide additional support to the typological and historical claim that fully-finite constructions can derive from nominalized structures (Gildea 1993; Heine 1993; Anderson 2006; 2011; DeLancey 2011). In fact, coherent explanation of NM's synchronic, finite final verb structures cannot be achieved without positing a nominal history for the negative declarative and interrogative constructions, imperatives, jussives, past progressives, and irrealis future tense constructions. Apart from such a historical analysis, various irregularities would escape explanation: why is it that some constructions require infinitive verb stems? Why do forms that clearly resemble subordinators (purpose markers and relativizers) appear internal to verbal word-forms? And, why should fused copulas and existentials occur in forms that now function as finite final verbs?

Givón has noted that in languages which place complements (including clauses) immediately adjacent to verbs, co-lexicalization and clause-union can result (Givón 2009: 74); this is the pattern that is seen repeatedly in the NM developments above. Within the languages of Ethiopia, historically nominal sources for finite structures are by no means limited to the OV pattern. Both VO and OV languages provide clear evidence of old nominalizations embedded in what are today finite structures. In Gumuz (Nilo-Saharan), which exhibits a predominantly VO pattern, the future tense construction derives from a nominalized purpose construction, with the old /mà-/ nominalizer still present in the new finite construction (C. Ahland 2012: 444). Nominal sources of now-finite verbs are also commonly found in Amharic (Ethiopian-Semitic) as well, where finite perfect verb forms carry nominal possessive suffixes, positioned before an existential (> auxiliary) (Leslau 1995: 387).

Acknowledgments

Michael Ahland

Abbreviations

Where two abbreviations are combined to form a complex gloss for a single form, the two are joined with a colon (e.g. NPST:AUX, non-past auxiliary form).

1,2,3	1st, 2nd, 3rd person		NF	non-final
AFF	affirmative		NP	noun phrase
AUX	auxiliary		NPST	non-past
COMP	complementizer		NSG	non-singular (dual/plural)
COP	copula		OBJ	object case
DECL	declarative		PASS	passive
DEF	definite		PF	perfect
DU	dual		PL	plural
EXIST	existential		PRES	present
FUT	future		PROX	proximal
GEN	genitive		PST	past
H	H tone		PURP	purposive
HRSY	hearsay		RC	relative clause
IMP	imperative		REL	relativizer
IMPR	impersonal		SBJ	subject case
INF	infinitive		SG	singular
INTR	interrogative		SOURCE	source oblique
JUSS	jussive		SS	same-subject
L	L tone		SUB	subordinate
LOC	locative		TI	temporally-integrated
M	M tone		TV	terminal vowel
NEG	negative			

References

Ahland, Colleen. 2012. *A grammar of Northern and Southern Gumuz*. Eugene, OR: University of Oregon. (Doctoral dissertation.)

Ahland, Michael. 2009. Aspects of Northern Mao phonology. *Linguistic Discovery* 7. 1–38.

Ahland, Michael. 2012. *A grammar of Northern Mao (Màwés Aas'è)*. Eugene, OR: University of Oregon. (Doctoral dissertation.)

Ahland, Michael. 2014a. Interactions between tone and morphosyntax in Northern Mao. (Paper presented at the 45th Annual Conference on African Linguistics, Workshop on Africa's Endangered Languages, University of Kansas, April.)

Ahland, Michael. 2014b. Subject marking interrupted: Perturbations from the development of Northern Mao's future tense suffix. *Studies in African Linguistics* 43(2). 1–16.

Ahland, Michael. 2015. The history of *ha*: How Ganza sheds light on the development of person marking in Northern Mao (Màwés Aas'è). (Paper presented at the 43rd Annual North American Conference on Afroasiatic Linguistics (NACAL), Georgetown University, February.)

Anderson, Gregory D. S. 2006. *Auxiliary verb constructions*. Oxford: Oxford University Press.

Anderson, Gregory D. S. 2011. Auxiliary verb constructions in the languages of Africa. *Studies in African Linguistics* 40(1-2). 1–409.

Bender, M. Lionel. 2000. *Comparative morphology of the Omotic languages*. Muenchen: Lincom Europa.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago & London: The University of Chicago Press.

DeLancey, Scott. 2011. Finite structures from clausal nominalizations in Tibeto-Burman. In Foong Ha Yap, Karen Grunow-Hårsta & Wrona Janick (eds.), *Nominalization in Asian languages: Diachronic and typological perspectives* (Typological Studies in Language 96), 343–360. Amsterdam/Philadelphia: John Benjamins.

Gildea, Spike. 1993. The rigid VS order in Panara (Cariban): A historical explanation. *International Journal of American Linguistics* 59. 44–63.

Givón, Talmy. 2001. *Syntax: An introduction, vol. 2*. Amsterdam/Philadelphia: John Benjamins.

Givón, Talmy. 2009. *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-cognition, evolution*. Amsterdam/Philadelphia: John Benjamins.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of human language*, 73–113. Cambridge, MA: MIT Press.

Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. Oxford: Oxford University Press.

Kassa, Getachew. 2014. *A grammar of Hozo*. Addis Ababa: Addis Ababa University (PhD dissertation draft.)

Leslau, Wolf. 1995. *A reference grammar of Amharic*. Wiesbaden: Otto Harrasowitz.

Mengistu, Girma. 2015. *A grammar of Seezo*. Addis Ababa: Addis Ababa University. (Doctoral dissertation.)

Noonan, Michael. 1997. Versatile nominalizations. In Joan Bybee, John Haiman & Sandra A. Thompson (eds.), *Essays on language function and language type: Dedicated to T. Givón*, 373–394. Amsterdam/Philadelphia: John Benjamins.

Schmidtke-Bode, Karsten. 2009. *A typology of purpose clauses* (Typological Studies in Language 88). Amsterdam/Philadelphia: John Benjamins.

Part IV

Cross-family studies

Chapter 25

Niger-Congo transitive reciprocal constructions and polysemy with reflexives

Ken Safir

Rutgers University

Naga Selvanathan

Rutgers University

The transitive reciprocal constructions in a large number of Niger-Congo languages are known to allow both reciprocal and reflexive meanings. Reflexive/reciprocal polysemy induced by verb affixes (or clitics) is common, but reflexive/reciprocal polysemy with full direct object (DO) anaphors appears to be relatively uncommon outside of Africa. Many other languages of Africa (e.g. Babanki, Bafut, Fe'efe'e, Ga, Ibibio, Limbum, Saari, Urhobo and Yoruba) have transitive reciprocal polysemy of this kind, but if so, the source of reciprocal interpretation remains a puzzle. We propose a novel type of reciprocal construction that can account for this pattern. In such Transitive Reciprocal Constructions (TRC), we posit that the direct object is a variable with no contribution to reciprocal meaning, but that it may have reciprocal form by virtue of shape concord with a reciprocal marker (RCM) on little v (also Bruening 2006). Otherwise it spells out as an anaphor that does not impose a specific kind of anaphoric reading. The source of reciprocal meaning is actually the RCM which can be covert in some languages. We further show that our proposal has far-reaching implications for theories of reciprocity and how identity-dependent meaning can be composed in natural language.

1 Introduction

The prevailing idea about the typology of reciprocal constructions is that there are two major types: the constructions which encode reciprocity with a periphrastic anaphor and those that encode reciprocity within the verb itself (König & Gast 2008; Siloni 2012; among others).

(1)   a. The men hit each other.
       b. The men collided.

As means of illustration, (1a) shows a construction in which *each other* is commonly assumed to be the source of reciprocal meaning (eg. Heim, Lasnik & May 1991). In (1b), however, the source of reciprocal meaning is lexicalized within the verb itself (Siloni 2012). There are, of course, further divisions within each strategy but each strategy is considered distinct from the other and each is given a separate account. A specific natural language, thus, employs either or both to convey reciprocity. In general terms, the main distinction between (1a) and (1b) comes down to a difference in verb valency. Unlike the verb in (1a) which is transitive, the verb in (1b) is intransitive.

In this paper, based on data from various African languages, it is proposed that all demonstrably transitive reciprocal sentences have the same structure, including (1a), and this analysis of *Transitive Reciprocal Constructions* (TRCs) extends to cases where the direct object appears to be a polysemic anaphor. TRCs are argued to have a reciprocal marker (RCM) hosted on little v and it is posited that it is the RCM which contributes the reciprocal meaning. The direct object in a TRC is thus simply a variable that *can* be spelled out with a reciprocal form but carries no reciprocal meaning (Safir 2014). Among other consequences, our proposal predicts a long noticed, but unexplained, aspect of reciprocals: unlike reflexives, reciprocal relationships are almost always clausal. This will be shown to arise as a natural consequence of our proposal.

The paper is organized as follows. In §2, we outline the empirical motivation for this distinct type of reciprocal strategy, the TRC. In §3, we outline the proposal in more detail and present evidence to support our analysis. In §4, we take a closer look at Lubukusu to substantiate the various aspects of the proposal. We then conclude with a list of implications that arise from our proposal.

2 Motivating the TRC: A formal account of polysemic anaphors

Our main motivation for our proposal arises out of an attempt to account for reflexive/reciprocal polysemy of argument position anaphors, which as far as we know has received little to no satisfactory formal account (although see Heine 1999 who provides a diachronic analysis of ambiguous anaphors). As an example of such polysemy, consider the following transitive constructions from Yoruba, Gungbe and Fe'efe'e.[1] The anaphors in the examples below are highlighted.

(2) Yoruba
Àwọn obìnrin náà rí **ara wọn**.
they women the see body their
'The women see themselves/each other/their bodies.' (ID1912)

---

[1] Data with an ID number have been accessed from the Afranaph Project Database (Safir, ongoing) on June, 2015, and can be found at this url: http://www.africananaphora.rutgers.edu/afranaph-database-mainmenu-130. Language data not accompanied with an ID number were personally communicated to us.

(3) Gungbe
Náwè lɔ lɛ ná hù **yé-ɖé**.
woman DET PL FUT kill 3PL-ɖé
'The women will kill themselves/each other.' (ID3597)

(4) Fe'efe'e
Pʉ-nʒwīɛ lɔfɛn **nā jɒɒ**.
C2-woman PST2-sell body POSS.DET.C1
'The women betrayed themselves/each other.' (ID8237)

Examples (2), (3) and (4) illustrate ambiguous sentences that can have a reciprocal meaning if the context permits. The Gungbe and Fe'efe'e anaphors are two-way ambiguous between a reciprocal and a reflexive reading. The Yoruba anaphor in (2) even permits a literal interpretation 'their bodies'. Such polysemy is attested in several other African languages (many represented in the Afranaph database, including Babanki, Bafut, Ga, Ibibio, Limbum, Saari and Urhobo).

What is the right formal account of this polysemy? Given the current understanding of how reciprocal constructions are formed (as illustrated in (1)), we have one of two choices. Either we have to accept that the highlighted anaphors are adjuncts and that reflexive and reciprocal meanings are mediated by a null verbal affix (under the intransitive hypothesis), or that the anaphors themselves are ambiguous with reciprocal and reflexive meaning (under what is taken to be the transitive hypothesis).

Given that the counterparts of (2–4) would be unacceptable if no direct object is lexically expressed (which would be the case if there were a null, detransitivizing affix), the sentences in question must be transitive, so we reject the former option. This means that the most plausible analysis is some form of the transitive analysis, such as an analysis where the source of reciprocity in these constructions is in each case the direct object itself. On this version of the transitive analysis, there must be (at least) two distinct lexical entries for the underlined phrases in (2–4), one of them being a reciprocal entry and another being a reflexive entry. In Yoruba, there would have to be three entries. This story quickly gets rather clumsy when we look at related languages. Consider Urhobo, a language that represents the extreme side of the scale. Aziza & Safir (2006) report that in this language, there are four different types of anaphoric direct objects and all four can be used to convey both reciprocal and reflexive meanings.

(5) Urhobo

a. Ayen mrẹ **ohwo-ohwo**.
they see.PST person-person
'They saw themselves/each other.' (ID2874)

b. Emo na fa **oma-oma-rẹ-ayen**.
children the flog.PST body-body-of-theirs-3PL
'The children flogged themselves/each other.' (ID2974)

c. Avware vwo eguọnọ kẹ **oma-rẹ-avware-e**.
We have love for body-of-ours-NEG

'We hate ourselves/each other.' (ID2828)

d. Emo na gbon **oma-obo-rẹ-ayen**.
children the criticize.PST body-hand-of-3PL

'The children criticize themselves/each other.' (Rose Aziza personal communication)

If all the highlighted phrases in (5) are the source of reciprocal/reflexive meaning, this would mean that each of them would have a reciprocal and a reflexive lexical entry – a substantial redundancy in the lexicon. Note that our consultant (Rose Aziza) reports that there is nothing determining a particular anaphor use. For example, (5) will have the same exact meaning even if one were to use the direct object forms in (5a) and (5b).

We propose an alternative version of the transitive analysis that we believe can explain the polysemy illustrated above more satisfactorily. We claim that the highlighted phrases do not contribute any reciprocal meaning at all and are just variables which have been spelled out differently. This idea is adopted from Safir (2014), who argues that there is universally only one abstract anaphoric form in natural language, but the "one true anaphor" can be spelled out differently depending on the phase in which its antecedent is found. Despite having no reciprocal/reflexive meaning themselves, these variables are merged as direct objects in order to satisfy the c-selectional needs of the verb. This is what we claim is also happening in these Niger-Congo languages. The source of reciprocal meaning in these constructions is, instead, posited to be a reciprocal marker (RCM) head that does not affect transitivity. This posited head is covert in Yoruba, Gungbe, Fe'efe'e and Urhobo but can be overt as well, as it is in Limbum, for example.

(6) Limbum
Wèe tʉr à rjà'-*(**se**) **mnyor**.
2PL have INF INF.help-PL.RCM c6-bodies

'We (PL) must help each other.' (Francis Ndi Wepngong pers. comm.)

As seen in (6), Limbum also has a direct object anaphor which does not appear to have any compositional reciprocal meaning, much like the languages in (2–5). However, Limbum differs from these other languages in having an overt RCM, realized as *-se*, which is necessary for reciprocal meaning. When this affix is absent, (6) only has a reflexive interpretation. A unified explanation for Limbum and the other languages we have seen naturally follows from our proposal. We claim that Limbum reveals what is actually happening in the languages in (2–5). In those languages, the polysemic nature of the anaphor is only illusory. There is in fact a covert head (an RCM) that determines when the sentence is interpreted as a reciprocal construction. In the absence of the RCM, the variable is simply determined as any other variable, namely, a thematic argument dependent on an antecedent, a local antecedent in this instance, to yield a reflexive reading. We,

thus, conclude that the data in (2–6) represent a distinct type of reciprocal construction, a TRC.[2] We outline the proposal in more detail in the next section.

3 Our proposal

We propose the following base structure for TRCs crosslinguistically:

(7)

vP
├── DP_antecedent
└── v'
    ├── v-RCM
    └── VP
        ├── V
        └── DP_anaphor
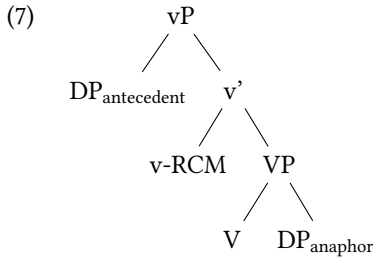
The tree in (7) shows the relevant structure in which we claim the main dependency and spell out forms are established. This is the little vP phase in which the reciprocal direct object and its antecedent are base generated. The other aspects of the proposal are as follows:

- The RCM is hosted on little v (see Bruening 2006 for a similar view).
- The subject DP (in spec, vP) binds the variable in object position and the RCM forms a mapping relation between the atoms in the subject DP. The object in TRCs does not contribute any reciprocal meaning.
- The relation between v-RCM and the DP anaphor determines the morphological shape of the anaphor (which we call *shape concord*).

In this paper, we argue that the RCM is the little v head (see also Baker, Safir & Sikuku 2013). The second aspect of the proposal is what allows us to account for reciprocal/reflexive polysemy that we saw in §2 in a straightforward way. In these sentences,

---

[2] Moysie-Faurie (2008) shows similar data in some Oceanic languages. Paicî, for example, has the following construction (Moysie-Faurie 2008: 121, her example (34)):

(i) Ru **pi-uĉa-rî ru**.
3DU PREF-look-TR 3DU

'They observe each other.'

This reciprocal construction requires an obligatory verb prefix as well as the direct object pronoun. Note that this pronoun is just a simple 3rd person pronoun, much like the subject pronoun. The fact that this is a transitive construction is indicated by the transitive verbal suffix *-rî*, which occurs when the post-verbal pronoun is included. This suffix indicates that the verb is not detransitivized by the reciprocal prefix *pi-*. TRCs with overt verb affixes and direct objects are clearly not unique to African languages. We would like to thank Patricia Cabredo Hofherr for bringing this paper to our attention.

reciprocal meaning comes from a null RCM. The various spell out forms (morphological shape) of the direct object that results are language-particular.

There is also an interesting cross-linguistic difference that our proposal can capture with a minor modification. In the languages that we have seen so far, it is difficult to see how the anaphoric form could be the source of reciprocal meaning. For example, in Limbum shown in (6), the anaphoric form *mynor* just means 'bodies'. However, we claim that our proposal can account for even constructions in which the direct object appears to be compositionally reciprocal. Consider Lubukusu.

(8) Lubukusu
Bob ne   Billi ba-a-bon-*(**an**)-a   (**ba-b-eene khu b-eene**).
Bob and Bill SM.C2-PST-see-RCM-FV c2-c2-own on   c2-own

'Bob and Bill saw each other.' (ID1377, 1379)

Like Limbum in (6), Lubukusu also has a sentence with an overt RCM and direct object. However, the RCM is obligatory for the reciprocal meaning to be realized. However, unlike Limbum, Lubukusu appears to have an anaphoric object which could plausibly encode reciprocal meaning. Nonetheless, we claim that it is only the RCM that encodes reciprocal meaning even in Lubukusu. Note that like Limbum, Lubukusu requires the presence of the RCM for the reciprocal meaning to emerge. To account for why TRCs in languages like Lubukusu have a reciprocal anaphor (RCA) as opposed to TRCs in languages like Limbum which have a non-specific anaphor (ANA), we propose reciprocal concord. We claim that Lubukusu anaphors can have an uninterpretable reciprocal feature. When such an anaphor is merged, this uninterpretable feature must be checked by the interpretable reciprocal feature of the RCM. This reverse AGREE account is similar to Zeijlstra's (2004) account of negative concord. Thus, the difference in spell-out form of the direct objects in Limbum and Lubukusu TRCs boils down to the feature difference in (9).

(9)   a. Lubukusu   b. Limbum

In (9a) and (9b), the RCM has interpretable reciprocal features. Lubukusu and Limbum differ in that only the former has an anaphoric form that has uninterpretable reciprocal features. By checking this feature, the anaphor in Lubukusu spells out as a form that

looks reciprocal. In Limbum, there is no such uninterpretable feature, thus it spells out as a non-specific anaphor.[3]

Our proposal is thus able to account uniformly for reciprocal constructions which on the surface look quite different. These surface patterns are represented schematically in (10).

(10)   a. DP$_{\text{antecedent}}$ V-RCM DP$_{\text{RCA}}$ Lubukusu
       b. DP$_{\text{antecedent}}$ V-RCM DP$_{\text{ANA}}$ Limbum
       c. DP$_{\text{antecedent}}$ V   DP$_{\text{ANA}}$ Yoruba

The (10a) pattern shows Lubukusu type languages which have an overt RCM and overt RCA. The (10b) pattern shows Limbum type languages which have an overt RCM and an overt ANA.[4] Finally, the (10c) pattern shows Yoruba type languages which have no overt RCM and an overt ANA. However, underlyingly, all three types of languages are argued to have the same structure shown in (7). They differ only in whether the RCM is overt or covert and in whether the anaphor has uninterpretable reciprocal features. In existing accounts, a uniform analysis for all these constructions is not possible and this is a further advantage of our proposal.

In the next section, we focus on Lubukusu which provides evidence for the two main aspects of the proposal: 1) that these reciprocal constructions are in fact transitive (even though the direct object can be omitted) and 2) that the position of the RCM is little v.

4 Supporting the details: Lubukusu case study

4.1 The RCA is not an adjunct

It is crucial to our analysis of TRCs that we show that the anaphoric element is not an adjunct but rather a genuine direct object. This is necessary because Mchombo (2004) (adopting Reinhart & Reuland 1993) has shown that in some Bantu languages, the RCM is a valence reducer. If the same applies to Lubukusu, then our claim that Lubukusu has TRCs is unsupported. However, we can independently show that Lubukusu RCAs are direct objects and not adjuncts. There are two pieces of evidence. The first has to do with possible interpretations in ellipsis contexts and the other has to do with anaphor positions when they are focused.

---

[3] There are alternative ways without appealing to feature checking to derive the same result. In a Distributed Morphology (DM) approach, concord with the antecedent adds a reciprocal feature to the anaphor. It will be pronounceable if there is special morphology for it, which is then inserted in the tree. If there is no morphology for it, it is not inserted as a reciprocal anaphor, but as a non-specific one. On this version of a DM account, there is always reciprocal concord, but not always morphology to show it.

[4] A reviewer asks whether an expletive can be merged as the DO in TRCs, but if so, the construction would simply not be transitive. Our claim that the DO has no reciprocal semantics does not mean that it has no thematic weight, but quite the opposite, as the DO has to be an argument of the verb. Thus, we would not expect to see an expletive in the DO position of a TRC much like we do not see sentences like the following in English, *He devoured there*, where *there* has no thematic content.

4.2 Sloppy and contrastive readings in ellipsis contexts

Mchombo (2004: 106) shows that in Chichewa, reflexive constructions without an overt object are actually transitive, and that reciprocal constructions without an overt object are intransitive. He does this using the availability of sloppy and contrastive readings in ellipsis contexts. This argumentation is taken from Sells, Zaenen & Zec (1987).

(11) Chichewa
[Aleje   á-ma-**dzi**-nyóz-á]   ku-pósá   asodzi.
  c2.hunters c2-HAB-RFM-despise-FV INF-exceed c2.fishermen

i. 'The hunters despise themselves more than the fishermen despise themselves.'
sloppy

ii. 'The hunters despise themselves more than the fishermen despise them.'
contrastive

In the reflexive construction in (11), both sloppy and contrastive readings are possible. While the sloppy reading is possible regardless of whether there is a direct object, the presence of the contrastive reading indicates that the antecedent clause must have a transitive verb.

(12) Chichewa
[Aleje   á-ma-nyóz-**án**-á]   ku-pósá   asodzi.
  c2.hunters c2-HAB-despise-RCM-FV INF-exceed c2.fishermen

'The hunters despise each other more than the fishermen despise each other.'
sloppy

By contrast, (12) shows that only the sloppy reading is possible in a Chichewa reciprocal construction. The absence of the contrastive reading is taken to indicate that the antecedent clause truly lacks a direct object. This comparison between the Chichewa RCM and reflexive marked (RFM) constructions thus provides a consistent way to test whether Lubukusu reciprocal constructions are indeed transitive. As expected under our proposal, the contrastive reading is possible.

(13) Lubukusu
[Ba-saani ba-a-biyil-**an**-a]   khu-khil-a   ba-khasi.
  c2-man   c2-PST-hate-RCM-FV c15-defeat-FV c2-woman

i. 'The men hate each other more than the women hate each other.' *sloppy*

ii. 'The men hate each other more than the women hate the men.' *contrastive*
(ID5036)

Example (13) shows that in the same type of ellipsis context as in Lubukusu, the elided clause can have the contrastive reading. Following Mchombo's conclusion for Chichewa reflexive constructions, we have to conclude that the reciprocal construction in Lubukusu is indeed transitive. The RCM, thus, cannot be a detransitivizer and the RCA (when it occurs) is a direct object.

4.2.1 Argument/Adjunct information focus asymmetries

In this section, we outline a novel argument in further support of our claim that the reciprocal construction in Lubukusu is indeed transitive. In Lubukusu, it appears that post-verbal elements that are arguments have to be *immediately* post-verbal if they have informational focus. However, post-verbal adjuncts need not be immediately post-verbal. This can be seen in the following contexts.

(14) Lubukusu
     Naanu ni-ye   ba-saani ba-a-pa   lukali?
     who   that-AGR C2-man   C2-TNS-beat fiercely

     'Who did the men beat fiercely?'

  a. A1: Ba-saani ba-a-pa   **Yohana** lukali.   (DO-Adverb)
          C2-man   C2-TNS-beat John   fiercely

      'The men beat John fiercely.'

  b. A2: #Ba-saani ba-a-pa lukali **Yohana**.   (Adverb-DO)

Example (14) shows a question in Lubukusu, seeking the identity of the person who received a beating. This question places information focus on *Yohanna.* As indicated by our informant (Justine Sikuku), as a response to the question, the adverb cannot intervene between the verb and the DO. Our consultant further reports that the order is only possible if the adverb in A2 is prosodically marked. This is quite surprising, given that in neutral contexts both the declaratives in A1 and A2 are possible without any additional prosodic marking on either the adverb or DO. This is what we see if the question asks for the manner in which the beating took place as in 'How did the men beat John?' Here, the response can have either order although information focus is on the adverb.

The difference between arguments and adjuncts also manifests with locative adjuncts versus locative arguments as shown in (15).

(15) Lubukusu
     Wahena ni-yo   Yohana a-a-kona   bwangu?
     where   that-AGR John   C1-TNS-sleep quickly

     'Where did John sleep quickly?'

  a. A1: Yohana a-kona   **mu-sikuri** bwangu.   (PP adj-Adverb)
          John   C1-TNS-sleep C18-field   quickly

      'John slept in the field quickly.'

  b. A2: Yohana a-kona bwangu **mu-sikuri**.   (PP adj-Adverb)

The answers A1 and A2 to the question in (15) illustrate information focus on an adjunct PP. In line with the pattern above, the PP adjunct can be either immediately post-verbal or have an intervening adverb. This does not hold when the PP is an argument of the verb as shown below.

(16) Lubukusu
      Wahena ni-yo   Yohana a-a-ra   si-tabu bwangu?
      where   that-AGR John   c1-TNS-put c7-book quickly

      'Where did John put the book quickly?'

    a.   A1: Yohana a-a-ra   si-tabu **khu-mesa** bwangu.   (PP arg-Adverb)
           John   c1-TNS-put c7-book c17-table   quickly

        'John put the book on the table quickly.'

    b.   A2. #Yohana a-a-ra si-tabu bwangu **khu-mesa**.   (Adverb-PP arg)

In (16), the answer to the question places information focus on an argument PP. Unlike with the PP adjunct, the adverb cannot intervene between the verb and the PP argument here. However, in a neutral context, where the locative argument is not focused, both orders in A1 and A2 are reported to be perfectly fine.

As a final illustration of this pattern, note that the restriction is not dependent on whether the intervener is an argument or adjunct. This can be seen with ditransitives.

(17) Lubukusu
      Naanu ni-ye   ba-saani ba-rum-ir-a   bi-tabu?
      who   that-AGR c2-man   c2-send-APPL-FV c8-book

      'Who did the men send the books to?'

    a.   A1: Ba-saani ba-rum-ir-a   **Maria** bi-tabu.   (IO-DO)
           c2-men   c2.TNS-send-APPL-FV Mary   c8-book

        'The men sent Mary books.'

    b.   A2. #Ba-saani ba-rum-ir-a bi-tabu **Maria**.   (DO-IO)

The IO-DO and the DO-IO orders are both possible in neutral contexts, but the IO-DO order in (17) is necessary when the IO has information focus as in this case. This shows that there is no sensitivity to whether the intervener is an argument or an adjunct. As long as some argument of the verb has information focus, it has to occur in an immediately post-verbal position and nothing can occur between it and the verb. This positional requirement does not hold of adjuncts with information focus.

This generalization in Lubukusu, which appears to be quite robust, thus gives us another reliable way to test whether the direct objects in reciprocal constructions are arguments or adjuncts. And as we will see, this test also points to the fact that the RCAs in Lubukusu reciprocal constructions are indeed arguments.

First note that when the RCA as a direct object has information focus, it must be immediately post-verbal.

(18) Lubukusu
      Naanu ni-ye   ba-saani ba-fumi-a   lukali?
      who   that-AGR c2-men   c2-praise-FV fiercely

      'Who did the men praise fiercely?'

a. A1: Ba-saani ba-fumi-an-a **ba-be-ene khu be-ene**
   c2-men c2-praise-RCM-FV c2- c2-own on c2-own

   lukali. (RCA-Adv)
   fiercely

   'The men praised each other fiercely.'

b. A2: #Ba-saani ba-fumi-an-a lukali**ba-be-ene khu be-ene**. (Adv-RCA)

Here, the adverb cannot intervene between the RCA and the adverb. This indicates that the RCA is an argument and not an adjunct. In addition, when the RCA is an IO with informational focus, only the IO-DO order is possible.

(19) Lubukusu

a. Ba-saani ba-rum-ir-an-a **ba-be-ene khu be-ene**
   c2-men c2.TNS-send-APPL-RCM-FV c2-c2-own on c2-own

   bi-tabu. (RCA-DO)
   c8-book

   'The men sent each other books.'

b. #Ba-saani ba-rum-ir-an-a bi-tabu **ba-be-ene khu be-ene**. (DO-RCA)

This evidence indicates once again that the RCA is an argument and not an adjunct of the verb. The data in this sub-section clearly converges with the conclusion based on the ellipsis facts above. The RCA in Lubukusu reciprocal constructions must be an argument. What this means is that the RCM itself cannot be a detransitivizer. The fact that the RCM and an argument RCA co-occur in Lubukusu receives a straightforward explanation in our proposal. Lubukusu reciprocal constructions are TRCs with the derivation shown in (7).

4.3 The RCM is in little v

In this section, we will reproduce some evidence from Baker, Safir & Sikuku (2013) in favor of the claim that the RCM is located in little v; although it should be noted that our claim that the RCM is hosted on little v is not a novel one (for example, see Bruening 2006). Our evidence here comes from the split comitative construction in Lubukusu.

We will now review the evidence in some detail. First, note the agreement paradigms that are evident in split comitative constructions and the corresponding readings.

(20) Lubukusu

a. O-mu-ndu ne e-m-bwa **bi**-a-rekukh-a.
   c1-c1-person and c9-c9-dog c8-PST-leave-FV

   'The man and the dog left.'

b. E-m-bwa **yi**-a-rekukh-a ne o-mu-ndu.
   c9-c9-dog c9-PST-leave-FV with c1-c1-person

   'The dog left with the man.'/'The dog and the man left.'

   c. O-mu-ndu   **a**-rekukh-a   ne   e-m-bwa.
     c1-c1-person c**1**-pst-leave-fv with c9-c9-dog

     'The man left with the dog.'/'The man and the dog left.'

Example (20a) shows an intact conjunct in the subject position of the sentence. Here, the noun class agreement is default class 8 for mixed class conjuncts, as shown in bold face. The (b) and (c) sentences show what happens when the conjunct splits. In (20b), the subject shows class 9 agreement and in (20c) the subject shows class 1 agreement. This is the case even if the sentences have the conjoined readings. This shows that the verb only agrees with whatever is in the syntactic subject position, presumably Spec, TP. Baker et al. account for the availability of the conjoined reading together with subject agreement by hypothesizing that the whole conjunct is base-generated in Spec, vP as the logical subject but that split occurs from this position. The higher conjunct moves to Spec, TP and verb movement to T results in the surface order.

The split comitative with an RCM is consistent with this general picture. Example (21) shows this.

(21)   Lubukusu

   a. O-mu-hayi   ne e-twika   bi-a-bon-**an**-a.
     c1-c1-hunter and c9-giraffe c8-tns-see-rcm-fv

     'The hunter and the giraffe saw each other.'

   b. O-mu-hayi a-a-bon-**an**-a   ne   e-twika.
     c1-c1-hunter c1-pst-see-rcm-fv with c9-giraffe

     'The hunter and the giraffe saw each other.'

Regardless of whether the conjunct is intact or split, the conjoined reading is possible and reciprocal interpretation imposed by the RCM makes it necessary in (21a,b). What this indicates is that the RCM must be able to access the conjunct before the split occurs. If it is true that the split occurs in the Spec of little v, then the RCM must not be located in a position higher than little v either. However, the RCM cannot also be lower than V, given that verb movement collects the RCM as a suffix. The most plausible location for the RCM is thus little v itself.

As a control, observe the pattern in a split comitative with a reflexive marker on the verb.

(22)   Lubukusu

   a. O-mu-hayi   ne e-twika   bi-a-**i**-bon-a.
     c1-c1-hunter and c9-giraffe c8-pst-rfm-see-fv

     'The hunter and the giraffe saw themselves.'

   b. O-mu-hayi a-a-**i**-bon-a   ne   e-twika.
     c1-c1-hunter c1-pst-rfm-see-fv with c9-giraffe

     'The hunter saw himself with the giraffe.'

     (*'The hunter and the giraffe saw themselves.')

Unsurprisingly, when the whole conjunct is present in the subject position, as in (22a), the conjoined reflexive reading is possible with default agreement on the verb. However, when split occurs, only the split reading is possible (only the surface subject is the antecedent of the reflexive relation) with corresponding agreement on the verb. What this indicates is that the RFM has to be in a position higher than little v. After all, if it was in little v itself, we would expect (22b) to have the conjoined reading like with the RCM. This analysis predicts that the RFM and RCM can co-occur given that they occur in different positions, and this is borne out as seen in (23).

(23) Lubukusu

Ba-khasi   ba-a-**i**-yeet-**an**-a   (ba-b-eene khu ba-b-eene).
c2-woman c2-PST-**RFM**-help-RCM-FV c2-c2-own on   c2-c2-own

'The women helped each other and themselves.' (ID1533)

Example (23) shows a prefixal RFM and a suffixal RCM occurring on the same verb. This is the clear evidence that the RFM and RCM cannot be located in the same position and this converges with the evidence from comitative split. Ultimately, the data we have seen here supports our claim that the RCM is in little v.

One may ask why it could not be the case that the RCM could be higher than little v, since the RCM could still "see" the whole conjunct before it splits (i.e., the conjunct is in its c-command domain and are presumably within the same phase), in which case there could be an alternative account of why conjoined readings are still possible with a split comitative like that in (21b). However, we have both empirical and theoretical reasons for ruling out this possibility. Empirically, this is not a desirable alternative, as the distinction between (21b) (with the RCM) and (22b) (with the RFM) would be eliminated. If a RCM that is higher than little v could see the whole conjunct through c-command, why can't the RFM? After all, the RFM also c-commands the little v position if it is, or is attached to, a higher head on the verb. There is also a theoretical reason against having the RCM higher than little v. We assume that the relationship between the external argument antecedent and the RCM is one of dependency, where the RCM depends on the antecedent for its interpretation. We further assume the Independence Principle (Safir 2004).

(24) Independence Principle (Safir 2004: 3)
     If x depends on y, x cannot c-command y.

Given that the RCM depends on the external argument, the Independence Principle will rule out any configuration in which the RCM c-commands the external argument. These considerations thus support the view that the RCM is in little v and not higher than it.

5 Implications and conclusion

In this paper, we have argued for a distinct reciprocal strategy which we have called the TRC in (7) above.

The main motivation for this new strategy was to provide a simple straightforward explanation for how reciprocal interpretations arise for argument position anaphors that show reflexive/reciprocal polysemy. In our analysis, the reciprocal anaphor is simply a variable that can be spelled out with a reciprocal shape in some languages. The actual reciprocal meaning comes from the RCM which can be covert in some languages. We have posited further that the RCM and/or the direct object anaphor can be covert, which generates a typology of surface patterns that we summarize in Table 1.

Table 1: Typology of TRCs. (Bolded elements are overt, unbolded elements with < > are covert.)

a.	$[_{vP}$ **DP**$_{antecedent}$ [v-**RCM** $[_{VP}$ **V DP**$_{anaphor}$]]]	Limbum, Lubukusu
b.	$[_{vP}$ **DP**$_{antecedent}$ [v-**RCM** $[_{VP}$ **V** <DP$_{anaphor}$>]]]	Lubukusu
c.	$[_{vP}$ **DP**$_{antecedent}$ [<v-RCM> $[_{VP}$ **V DP**$_{anaphor}$]]]	Urhobo
d.	$[_{vP}$ **DP**$_{antecedent}$ [<v-RCM> $[_{VP}$ **V** <DP$_{anaphor}$>]]]	–

The surface realization in line (b) of Table 1, where the non-specific anaphor is a null direct object, was argued for on the basis of ellipsis data in (13). Of the four possibilities, the pattern in line (d) is presumably ruled out on functional grounds, as there will be no evidence for the hearer that a reciprocal construction is being used as opposed to a regular intransitive sentence. Also note that the direct object DP could be an RCA or ANA. Since English reciprocal sentences with *each other* are transitive, we are committed to the analysis that such sentences in English are TRCs. Williams (1991), for instance, argues (contra Heim, Lasnik & May 1991) that *each other* is not compositional, that is, they are just non-compositional surface forms; and if this is right, then English reciprocal constructions could just be an instance of the pattern in line (c).[5] The implication here is that a reciprocal construction may appear to be intransitive because an object is missing or may appear to have a reciprocal anaphor that is contributing to semantic interpretation; but on close analysis, the construction may actually turn out to be a TRC. In Lubukusu, we showed this by demonstrating that the RCM must occur for the reciprocal meaning to emerge and by showing that the RCA is indeed an argument and not an adjunct. Although it is unrealistic to expect the same tests for transitivity to work in all languages, we call into question any claim that it is the RCA, rather than overt or covert RCM, that contributes the semantic meaning of reciprocity.[6]

---

[5] We do not defend this claim here due to space limitations, but it is worth noting that *The boys told each girl a different story* is ambiguous (as many stories as girls, or they all hear one story different from some other in the discourse); while *The boys told each other a different story* does not allow the distributed reading. This is explicable if *each other* is a non-decomposable idiom which surfaces as a result of shape concord with a null RCM on English verbs.

[6] As part of the wider cross-linguistic applicability of our proposal, a reviewer asks how our analysis could be applied to languages with "SE anaphors" (Reinhart & Reuland 1993) which also show reflexive/reciprocal polysemy which disappears in certain contexts. For example, while German SE anaphors within a preposi-

Our proposal also enables us to explain a notable difference between reflexives and reciprocals. While long distance reflexives are relatively common, long distance reciprocals are almost completely absent.[7] We believe that our proposal accounts for this quite naturally if the RCM on little v, overt or not, is the only source of reciprocal readings. When the direct object is transferred to the interfaces as the vP containing it is completed, the object's spell out form and interpretation has to be determined as well. Given that the only possible binder of the variable at this point is the subject DP in Spec, vP, it follows that any antecedent of the direct object must be locally found as well, and the relationship must be reciprocal. For the logic of this approach to go through, we must assume that reflexive interpretations do not always require the mediation of an RFM; that is, there are languages where reflexive readings are possible just because a coargument variable is bound by a local antecedent (as in typical long distance bound variable anaphora). The latter demonstration is beyond the scope of this article, however.

In conclusion, our proposal for the syntactic structure of TRCs has the following explanatory advantages. It accounts for the availability of both reciprocal/reflexive meanings with the same direct object shape, without requiring redundant stipulations in the lexicon for languages like Urhobo, which have multiple forms that allow both reflexive and reciprocal readings. It also explains how reciprocal constructions that include an overt reciprocal operator on the verb and an overt anaphor in an argument position are formed. Cross-linguistically, our analysis provides a coherent and empirically supported typology of TRCs, thus unifying the syntactic distribution of reciprocal readings across a wide range of apparently different constructions. Finally, our analysis suggests an account for the general absence of long distance reciprocals by contrast with long distance reflexives.

---

tion phrase (PP) can only be construed as a reflexive, Polish SE anaphors still retain the reflexive/reciprocal polysemy within PPs. One explanation for this contrast between German and Polish that our analysis could shed light on is by exploiting whether PP is a phase. Recall that in our analysis, the DO is spelled out as a RCA or ANA within the phase. In TRC contexts, this is possible because the vP contains the antecedent for the DO in this phase. Thus, if PPs in German were a phase as well, then the SE anaphor construed as a reciprocal within the phasal PP will be left without an antecedent. On the other hand, if Polish PPs were not phases, then the SE anaphor construed as a reciprocal will not be spelled out until the higher vP (which contains the antecedent) is completed. The inability of German PPs to allow pseudopassive formation, as opposed to English, would be consistent with the assumption that German PPs are phases and (at least some) English PPs are not. However, the reviewer also mentions another fact which we do not see how to account for, though we do not know of an account that does better. The polysemy of German SE-anaphors also disappears when the DO is stressed. In such contexts, only the reflexive meaning is possible. Thus, *Sie hassen sich* ('they hate themselves') only has the reflexive reading, whereas *Sie hassen sich* ('they hate themselves/each other') can be reflexive or reciprocal. First, we do not agree with the reviewer that this indicates that the German SE-anaphor has two functions (one as a reciprocal valence marker and one as a reflexive argument). Nonetheless, this hints at another source of parametric variation between German and English, as the latter clearly allows a stressed RCA (*They saw each other*). We will have to leave detailed extensions of our theory to other language groups for future research.

[7] That is, with the exception of the so called assisted long distance reciprocals as in *John and Paul said they saw each other*. In this sentence, the crucial reading is one where John said he saw Paul and Paul said he saw John. Although this has been analyzed as a scope phenomenon, see Dimitriadis (2000) who proposes a non-scope solution which crucially relies on local antecedenthood. For discussion of the locality of reciprocal constructions crosslinguistically, see, for example, Everaert (1999).

Acknowledgments

Abbreviations

1,2,3	1st, 2nd, 3rd person features	NEG	negation
		PERF	perfective
AGR	Agreement	PL	plural
APPL	applicative	POSS	possessive
c1, c2, c3 etc.	noun class markers	PREF	verb prefix
DET	determiner	RCA	reciprocal anaphor
DO	direct object	RCM	reciprocal marker
DU	dual	RFM	reflexive marker
FUT	future tense	SE	simplex expression
FV	final vowel	TNS	tense
HAB	habitual	TR	transitive suffix
INF	infinitive		

References

Aziza, Rose & Ken Safir. 2006. A sketch of the pattern of anaphora in Urhobo. Version 1.1. http://www.africananaphora.rutgers.edu/images/stories/downloads/casefiles/UrhoboAS.pdf, accessed 2013-09-01.

Baker, Mark, Ken Safir & Justine Sikuku. 2013. Complex anaphora in Lubukusu. In Ọlanikẹ Ọla Orie & Karen W. Sanders (eds.), *Selected Proceedings of the 43rd Annual Conference on African Linguistics*, 196–206. Somerville, MA: Cascadilla Proceedings Project.

Bruening, Benjamin. 2006. The morphosyntax and semantics of verbal reciprocals. http://udel.edu/~bruening/Downloads/Reciprocals5.pdf, accessed 2013-09-01.

Dimitriadis, Alexis. 2000. *Beyond identity: Topics in pronominal and reciprocal anaphora.* Philadelphia: University of Pennsylvania. (Doctoral Dissertation.)

Everaert, Martin. 1999. Types of anaphoric expressions: Reflexives and reciprocals. In Zygmunt Frajzyngier & Traci S. Curl (eds.), *Reciprocals: Forms and function* (Typological Studies in Language 41), 31–63. Amsterdam: John Benjamins.

Heim, Irene, Howard Lasnik & Robert May. 1991. Reciprocity and plurality. *Linguistic Inquiry* 22. 63–101.

Heine, Bernd. 1999. Polysemy involving reflexive and reciprocal markers in African languages. In Zygmunt Frajzyngier & Traci S. Curl (eds.), *Reciprocals: Forms and function* (Typological studies in language 41), 1–30. Amsterdam: John Benjamins.

König, Ekkehard & Volker Gast. 2008. Reciprocity and reflexivity – description, typology and theory. In König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations,* 1–32. Berlin: Mouton De Gruyter.

Mchombo, Sam. 2004. *The syntax of Chichewa.* Cambridge: Cambridge University Press.

Moysie-Faurie, Claire. 2008. Constructions expressing middle, reflexive, and reciprocal situations in some Oceanic languages. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, 105–168. Berlin: Mouton De Gruyter.

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24. 657–720.

Safir, Ken. 2004. *The syntax of (in)dependence.* Cambridge: MIT Press.

Safir, Ken. 2014. One true anaphor. *Linguistic Inquiry* 45. 91–124.

Sells, Peter, Annie Zaenen & Draga Zec. 1987. Reflexivization variation: Relations between syntax, semantics and lexical structure. In Masayo Iida, Stephen Wechsler & Draga Zec (eds.), *Working papers in grammatical theory and discourse structure*, 169–238. Stanford: CSLI.

Siloni, Tal. 2012. Reciprocal verbs and symmetry. *Natural Language and Linguistic Theory* 30. 261–320.

Williams, Edwin. 1991. Reciprocal scope. *Linguistic Inquiry* 22. 159–173.

Zeijlstra, Hedde. 2004. *Sentential negation and negative concord.* Amsterdam: University of Amsterdam. (Doctoral dissertation.)

Chapter 26

STAMP morphs in the Macro-Sudan Belt

Gregory D. S. Anderson

Living Tongues Institute for Endangered Languages
& University of South Africa (UNISA)

STAMP morphs —portmanteau subject-tense-aspect-mood-polarity morphs exhibiting functional and formal properties of both pronominals and auxiliary verbs— are characteristic of many of the genetic units across the Macro-Sudan Belt. STAMP morphs typically occur within a constructional frame, which may include a verb in a fully unmarked form or various construction-specific finite or non-finite forms. STAMP morphs often originate in, and synchronically may appear in, auxiliary verb constructions. Further, univerbation of STAMP morphs with following verbs has yielded various series of inflectional prefixes in a range of families and individual languages. After introducing the functional and formal/constructional properties of STAMP morphs, the paper discusses the origin of STAMP morph constructions and their subsequent developments into bound inflectional or conjugational prefix series. The paper closes by presenting a typology of STAMP morphs in three important Macro-Sudan Belt macro-groupings, viz. Chadic, Central Sudanic, and Niger-Congo languages; and for the last mentioned, particularly the Benue-Congo taxon.

1 Introduction

In this article I introduce and exemplify a curious and characteristic feature of a number of languages found across different genetic units spanning the large areal complex of equatorial Africa called the Macro-Sudan Belt (Güldemann 2008). This feature has challenged analysts pursuing descriptions of various western and central African languages, exhibiting functional properties typically associated with both subject pronouns and auxiliary verbs in many other languages; it has recently been called a STAMP morph (Anderson 2012; 2015). This is mnemonic for what these elements largely are, portmanteau morphs that encode the referent properties of semantic arguments that typically play the syntactic role of 'S[ubject]'–that is, the person, number and gender properties of such an actant–in combination with categories of T[ense], A[spect], M[ood] and P[olarity]. Such elements have also been previously called the *tense-person complex* (Creissels 2005), and *pronominal predicative markers* or *pronominal auxiliaries* (Vydrine 2011; Èrman 2002) in the Africanist literature.

Portmanteau STAMP morphs are found throughout languages representing different genetic units of the Macro-Sudan Belt. In their most basic form, they are the sole means of encoding referent properties of the syntactic subject, in addition to most of the types of categories encoded by verbal Tense, Aspect and *Aktisonsart*, Mood and Polarity morphology. An example of the simplest STAMP morph construction can be seen in (1) and (2) from Tarok, a language of Nigeria belonging to the Tarokoid Plateau genetic unit (Benue-Congo stock, Niger-Congo phylum).

(1)   Tarok (Sibomana 1981/1982: 238) [Tarokoid/Plateau]
      n   wá   ù-dɨŋ.
      1.pfv drink clsfr-water
      'I have drunk the water.'

(2)   Tarok
      mi   wá   a-tí   ipín.
      1.irr drink clsfr-tea tomorrow
      'I will drink tea tomorrow.'

In this study, I discuss some of the characteristics of STAMP morphs and STAMP morph constructions in a set of languages and genetic units of the Macro-Sudan Belt. §2 discusses the genetic units that constitute the typological sampling of African languages used in this study. §3 discusses the range of functions associated with STAMP morph constructions, and §4 explores their formal constructional features. §5 offers some thoughts on the origins of STAMP morphs. §6 discusses the subsequent role such formations have played in the development of prefixed conjugation series that are likewise found across the different genetic units of this linguistic area. Sections 7-9 briefly examine the role of STAMP morph constructions and prefix conjugations that developed out of erstwhile STAMP morphs in three macro-level groupings of genetic units (stocks) in the Macro-Sudan Belt. This includes the Chadic stock (§7), the Central Sudanic stock of Nilo-Saharan (§8), and the various genetic units found within the Benue-Congo stock of Niger-Congo (§9).

2 Typological Sampling and the Macro-Sudan Belt

In his seminal work on the Macro-Sudan Belt [MSB], Güldemann (2008) remained agnostic about the usefulness of higher order units in Africa, at least with respect to typological sampling. Instead he suggested the fourteen genera in Table 1, with Ijoid, Dogon, Chadic and Songhay being peripheral to the proposed area insofar as they show fewer of the characteristic core-features identified as definitional of the MSB, than do the majority of other identified genera.

Table 1: Genera in the Macro-Sudan Belt according to Güldemann (2008)

Adamawa	non-Bantu Benue Congo	Bongo-Bagirmi	Dogon
Kwa	Kru	Gur	Songhay
Mande	Moru-Mangbetu	Ubangian	Ijoid
Chadic	Atlantic		

Since 2008, a number of revisions to standard classifications of African languages have been offered, including statements by Dimmendaal (2008; 2011) and Sands (2009). In previous work on African linguistic typology, I have suggested a larger range of sampling units as appropriate for cross-linguistic work, which have internal diversification on the level of Germanic or Romance; I call these 'genetic units for typological sampling' (Anderson 2011), and write them in small-caps to distinguish them from names of both languages and larger taxa. Many genera in Table 1 would thus be split into multiple different genetic units; in this system, the number of such genetic units in the Macro-Sudan Belt to use in typological sampling totals over fifty. Not all these genetic units have attested STAMP morph constructions, but the vast majority does. In my corpus the genetic units that include STAMP morph constructions (or prefix conjugation series that historically derive from such constructions, see section 5) include at least those in Table 2.

3 Functions of STAMP morphs

STAMP morphs usually combine with a following verb to encode the TAM categories of the event, and the person, number and, where relevant, gender categories of the subject. Such forms can be found in as diverse an array of languages of the Macro-Sudan Belt as the Biu-Mandara (Central) Chadic language Merey (3) of Cameroon, the Kru language Wobé (4) of Côte d'Ivoire, and Bongo of the Sara-Bongo-Bagirmi family of Central Sudanic (5) spoken in South Sudan.

(3) Merey (Gravina 2007: 8) [Biu-Mandara Chadic]
na   zal.
1.PST call
'I called.'

Table 2: Genetic units w/STAMP morphs in MSB (Anderson 2011; Dimmendaal 2008; 2011; Sands 2009)

Atlantic-Congo families, isolates and stocks

Bantoid	Gbe	Kulango	Potou-Tano
Bendi	Gur	Leko-Nimbari	Senufic
Cangin	Igboid	Mbum-Day	Tarokoid
Cross River	Jen Bambukic	Na-Togo	Tenda
Edoid	Ka-Togo	Nupoid	Ukaan
Ega	Kainji	Okoid	Waja-Kam
Fali	Kru	Plateau	Wolof
Ga-Dangme			

Chadic sub-families

Biu-Mandara (Central)

East Chadic

Masa

West Chadic

Central Sudanic families

Kresh

Lendu

Mangbetu

Mangbutu-Lese/Efe

Moru-Ma'di
Sara-Bongo-Bagirmi

Mande families

East Mande

Northwest Mande

Southeast Mande

Southwest Mande

Ubangian families

Gbaya

Ngbandi

Mba

Zande
+traces in:
    Ngbaka
    Banda

(4) a. Wobé (Hofer & Link 1973/1980: Wobé 3)[1] [Kru]
         ẽ ² gyi³².
         1.PST come

         'I have come.'

     b. Wobé
         ma² gyi³².
         1.NPST come

         'I am coming.'

(5) Bongo (Tucker & Bryan 1966: 75) [Sara-Bongo-Bagirmi]
     ma   bi.
     1.INDEF give

     'I am giving.'

One of the curious and characteristic features of STAMP morph constructions is the possible encoding of negative polarity without any distinct negative polarity scope operator, the negative polarity being encoded together with TAM categories and referent properties of the subject in the STAMP morph itself. Formations of this type can be found in languages of the Macro-Sudan Belt such as Duka (Kainji) (6) or Mano (Southeast Mande) (7).

(6) Duka (Bendor-Samuel, Skitch & Cressman 1973: 13) [Kainji]
     mân   hé ò-kɔ́t.
     I.FUT.NEG go to-bush

     'I won't go to the bush.'

(7) Mano (Vydrine 2009: 226) [Southeast Mande]
     lὲέ   máá wè   gèē.
     3SG.NEG[HAB] Mano language speak

     'S/he doesn't speak Mano.'

As noted earlier, the referent categories expressed in STAMP morphs are typically restricted to subjects, but in a small number of instances one also finds portmanteau subject > object forms within a STAMP morph, as in the following sentence from Kohumono (Cross River).

(8) Kohumono (Cook 1972/1980: 355) [Cross River]
     βɔ́   fà.
     1>2.NPST bite

     'I bite you.'

---

[1] Hofer & Link's volume does not number the chapters consecutively, but rather each chapter begins anew at page 1. Thus this form comes from page 3 of the Wobé chapter. Their page-numbering convention is followed here.

In Guro (Southeast Mande), one finds portmanteau subject > object-encoding STAMP morphs that also reference polarity, as in *ɓe* vs. *yaa* in the following sentences (note the tone differences on the verbs in these constructions as well as the presence of the postverbal negator *ɗo* in (9b)).

(9) a. Guro (Vydrine 2009: 239) [Southeast Mande]
     ɓe   zuru-o.
     2sg>3sg.ipfv wash-ipfv

     '[You] wash him/her/it.'

   b. yaa   zùrù-ò   ɗo.
     2sg>3sg.ipfv.neg wash-ipfv neg

     '[You] don't wash him/her/it.'

The next section (§4) turns to the range of constructional types in which STAMP morphs participate.

4 Formal subtypes of STAMP morph constructions

STAMP morph constructions exhibit a wide-range of formal sub-types. As mentioned previously, the simplest construction consists of a STAMP morph and an unmarked or bare stem form of the verb (10).

(10) Bare Stem Construction: STAMP Verb

Some examples of this constructional type of STAMP morph may be found in Fyem (11). Wolof also has an extensive system of such STAMP morph constructions, as do a number of Mande languages. Upwards of ten separate paradigms may be found.

(11) a. Fyem (Nettle 1998: 32, 35) [Plateau]
     náá   soo Gindiríŋ.
     I.prf go   Gindiri

     'I went to Gindiri.'

   b. ín   soo dirámméka.
     I.ipfv go   farm.your.oblq

     'I will go to your farm.'

In a second STAMP morph construction the verb appears in a construction-specific tonal form, which suggests that there is an associated floating tone projecting from the STAMP morph onto the verb stem.

(12) STAMP + Floating Tone Construction: STAMP Verb_{<construction-specific tonal form>}

West Chadic Guus (also known as Sigidi) also has a highly elaborated system of STAMP morphs (Table 3), some of which have associated floating tones that they project onto the verb. The Recent Past (REC) form projects a high tone onto the following verb. The Future (FUT) does the same, and differs from the otherwise identical-looking subjunctive [SBJN] in whether it assigns this floating high tone (FUT), or does not (SBJN); the two are thus *constructionally* distinct.

Table 3: STAMP morphs in Guus [Sigidi], West Chadic (Caron 2001: 8-9)

	AOR	IMM	SBJN	FUT	HAB	PFV	REC	IRR
1	*ma*	*maa*	*mə*	*mə*$^{+H}$	*mak*	*map*	*mam*$^{+H}$	*məs*
2	*ka*	*kaa*	*kə*	*kə*$^{+H}$	*kak*	*kap*	*kam*$^{+H}$	*kəs*

(13) Guus [Sigidi] (Caron 2001: 11) [West Chadic]

ən ka   ɗuu karáŋ tʃí   məʃi.
if you.IRR beat dog   s/he.FUT die

'If you beat the dog, it will die.'

The divergent Atlantic-Congo (or Gur) isolate language Kulango is another language of the Macro-Sudan Belt that reflects floating tones associated with STAMP morphs. In Kulango (14) the first singular habitual STAMP morph projects a low tone onto the first stem syllable of the verb, while the corresponding subjunctive form projects a high tone. Both STAMP morphs bear high tones themselves, but one projects a floating low tone and the other a floating high tone.

(14) Kulango (Elders 2007: 193) [Kulango]

a. má   dɔlɪ.
   I.HAB$_{<+L>}$ sell

   'I sell.'

b. mɪ   dɔlɪ.
   1.SBJN$_{<+H>}$ sell

   'May I sell.'

The Waja-Kam language Dadiya shows similar constructions.

(15) Dadiya (Jungraithmayr 1968/1969: 196) [Waja-Kam]

ń   já.
you.PRF eat:PRF

'You have eaten.'

In a third sub-type of STAMP morph construction, the verb appears in a construction-specific co-grammaticalized aspectual form (16).

(16)   Aspect Construction: STAMP Verb_{<ASP>}

Such STAMP morph constructions are found in Tiv (Bantoid) (17), and Ndut-Falor (Cangin) (18).

(17)   a. Tiv (Arnott 1958: 114) [Bantoid/Tivoid]
            vé   pìnè.
            they.PST ask

            'They asked.'

        b. mbá   pi|ne-n.
            they.PRS ask-ASP

            'They are asking.'

(18)   Ndut-Falor (Pichl 1973/1980: Ndut-Falor 4) [Cangin]
        mi   acɛ.
        I.RLS come:PRF

        'I have come.'

A sub-type of this construction is found with the verb stem having a morphological aspect marker as well being reduplicated (19).

(19)   STAMP REDPL-Verb-ASP

This is found in the following formation from Dadiya of the Waja-Kam genetic unit (20):

(20)   Dadiya (Jungraithmayr 1968/1969: 197) [Waja-Kam]

        a. mɜ̀n   nò-lɛ.
            I.NPST drink-PROG

            'I am drinking.'

        b. mon   jà-jà-l.
            you.NPST RDPL-eat-PROG

            'You are eating.'

In some African languages this construction-determined aspectual form is itself realized tonally (21).

(21)   Tonal Morphology STAMP Verb_{<ASP=TONE>}

An example of such a realization comes from 'Bozom, a Gbaya Ubangi language (22), in which the IPFV and PFV verb forms are distinguished tonally.

(22) 'Bozom (Moñino 1995: 159) [Gbaya Ubangi]

    a. ʔầ   ré.
       he.RLS enter.IPFV

      'He enters.'

    b. má   rè.
       he.IRR enter.PFV

      'He will enter.'

Verbs in STAMP morph constructions can also appear in specific construction-dependent modal forms (23).

(23)   Modal Construction STAMP Verb$_{<MODAL>}$

Examples of this type can be found in Ndut-Falor (24) and Duka (25).

(24)   Ndut-Falor (Pichl 1973/1980: Ndut-Falor 4) [Cangin]
    ma[y] ayɛ.
    I.FUT come:MOD

    'I will come.'

(25)   Duka (Bendor-Samuel, Skitch & Cressman 1973: 96-98; 105) [Kainji]
    mɛ́ heɛ́.
    I.IRR go:COND

    'If I go...'

Split negative marking is relatively common in AVCs (Anderson 2006; 2011) where the negative is encoded on a lexical verb but other obligatory categories are on the auxiliary. In Africa, this pattern is a feature of auxiliary verb constructions in the languages of the Rashad Kordofanian genetic unit, e.g. in Rashad proper (26).

(26)   Rashad (Tucker & Bryan 1966: 297) [Rashad Kordofanian]
    ŋi fas k-eyɛ y-ɛn.
    I meat NEG-eat 1-AUX

    'I am not eating meat.'

Thus, it is not overly surprising that negative can also be encoded in STAMP morph constructions as well (27), albeit reflecting different syntactic configurations in the languages of the Macro-Sudan Belt (the lexical verb follows the functional element) than in the Nuba Hills languages (where the reverse tends to be true).

(27)   Split Negative Construction: STAMP Verb$_{<NEG>}$

Such a formation in the negative future in Gã of the Ga-Dangme Kwa genetic unit (28).[2]

---

[2] Negative can be encoded in the STAMP morph itself as well. When so, the verb is thus in a 'co-negative' form.

(28) Ga (Kropp Dakubu 1988: 105) [Ga-Dangme Kwa]
      eᴵ   bá˦-ŋ.
      he.ꜰᴜᴛ come-ɴᴇɢ
      'He will not come.'

The other common split inflectional pattern found cross-linguistically, but one that is not overly common in the languages of Africa and particularly not in those of the Macro-Sudan Belt (Anderson 2011), is a subject-object split pattern (Anderson 2006). In the split subject-object pattern, the object is encoded on the lexical verb that subcategorizes for it, and the other obligatory inflectional categories including subject information appear on the auxiliary as in the auxiliary verb construction in (29) from Gidar.

(29) Gidar (Frajzyngier 2008: 263) [Biu-Mandara Chadic]
      wá-nə̀ mpə̀r-kó.
      ꜰᴜᴛ-1 chew-2
      'I will eat you.'

This pattern also appears to be reflected in STAMP morph constructions as well (30).

(30) Split Object Construction: STAMP Verb_{<ᴏʙᴊ>}

Such a formation may be found in a small number of Chadic languages, for example, West Chadic Polci (31) or the Biu-Mandara (Central) Chadic language Mofu-Gudur (31).

(31) Polci (Caron 2008: 153) [West Chadic]
      Gǎrbà kən ndʒaŋ slo: wú ɗe kə   fũ̌:-m.
      Garba ᴄᴏᴘ cut   meat ᴀᴄᴄ ɪɴᴊ 2:ᴀᴏʀ say-1
      'If Garba slaughters a beast, tell me'

(32) Mofu-Gudur (Pohlig 1992: 4) [Biu-Mandara Chadic]
      fá   tá-ka   ɗǎf.
      ᴘʀᴏɢ.3 prepare-2.ɪᴏ food
      'She is preparing you food.'

5 Origin of STAMP morphs from auxiliary constructions

Where might STAMP morphs have come from? In other words, how did they develop? In known instances, they result from the fusing of a pronominal marker or pronoun with a following auxiliary verb, fused into a single portmanteau complex (Anderson 2006; 2011).[3] An example showing synchronic variation between a STAMP morph and its source auxiliary verb construction comes from Limbum (33).

---

[3] There are also instances where STAMP morphs have no obvious etymology, as in many Mande languages.

(33) Limbum (Fiore & Peck 1973/1980: Limbum 3-4) [Bantoid]
      mɛ:³² wɨ³ ~ mɛ³ ʃe² wɨ³ mɛ:³² < mɛ³=ʃe².
      1.ᴘʀᴏɢ come 1 ᴘʀᴏɢ come
      'I am coming.'

Researchers on individual languages and families have recognized this connection. Thus Shimizu (1983: 101), when describing STAMP morphs in the Leko-Nimbari language Zing Mumuye states that "the surface differences in subject pronouns are in fact due to the TAM markers, which are realized on them or contracted with them." Further, Babaev (2010: 35), in discussing issues in the reconstruction of Benue-Congo cautions the reader 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."

Language-specific internal evidence sheds light on this most likely source for the development of STAMP morphs. In Sara of the Sara-Bongo-Bagirmi family of the Central Sudanic stock, the so-called indefinite series subject pronouns are not used with finite, inflected forms of a following vowel-initial verb, but rather with an infinitive form.

(34) Sara (Tucker & Bryan 1966: 75) [Sara-Bongo-Bagirmi]
      ma k-usa.
      1:ᴅᴇꜰ ɪɴꜰ-eat
      'I am eating.', 'I eat (ʜᴀʙ).'

Though a finite construction that seemingly consists of a combination of a subject pronoun plus an infinitive verb form is odd typologically, such a formation is entirely consistent with typological norms if the subject pronoun actually has an auxiliary verb 'hidden' within it, or in other words is actually a STAMP morph. Sara and its sister languages (see §8 below) are far from the only languages that show constructions with infinitive verb forms in combination with specific STAMP morphs. The entirely unrelated Bantoid language Nomaande shows similar structures.

(35) Nomaande (Wilkendorf 2001: 22) [Bantoid]
      yɔ́ ɔcɔba ná-áyá.
      3.ᴘʀs ɪɴꜰ:go to-3sɢ
      'He is going to him (i.e., his home).'

Similarly, in Dott (also known as Dass or Zoɗi) of the West Chadic family (36), the continuous or progressive STAMP morph series requires a verbal noun form of the lexical verb in the predicate.

(36) Dott/Zoɗi (Caron 2002: 165) [West Chadic]
      taa   tʃét-ti.
      3ᴘʟ.ᴄᴏɴᴛ come-ᴠɴ
      'They are coming.'

Yet other languages of the Macro-Sudan Belt require dependent-marked forms of verbs to appear in combination with STAMP morphs in specific constructional subtypes. Such languages include the Kainji language Duka (37) and ʻBozom (38) of the Gbaya Ubangi family.

(37) Duka (Bendor-Samuel, Skitch & Cressman 1973: 96-98; 105) [Kainji]
mɛ́ ə̀m-hà á.
I.IRR DEP-go NEG
'I am not going.'

(38) ʻBozom (Moñino 1995: 159) [Gbaya Ubangi]
ʔà̰ rè-á.
he.RLS enter.PRF-DEP
'He has entered.'

6 STAMP morphs developing into prefixal conjugations

Having established that STAMP morph constructions often derive from AVCs, the question arises as to where these formations "go" in their subsequent historical developments. In "normal" AVCs, an auxiliary is often drawn into the verbal complex, and loses its phonological integrity as a freestanding element. Similar developments are seen with STAMP morph constructions. In the Macro-Sudan Belt one typically finds that old STAMP morphs have gotten drawn into the verb to become prefixal conjugation markers. This can be seen in cognate formations in Nchumuru varieties. Bejamso-Grubi Nchumuru (39) has freestanding STAMP morphs, while in closely related Banda Nchumuru (40), these have become a prefixed conjugation set.

(39) Bejamso-Grubi Nchumuru (Price 1975/1980: Nchumuru 5) [Potou-Tano Kwa]
màá bà.
I.PRF come
'I have come.'

(40) Banda Nchumuru (Cleal 1973/1980: Nuchumuru 4) [POTOU-TANO Kwa]
mà-ba.
1PST:PRF-come
'I have come.'

Such developments have been known in African linguistics for more than a century.[4] I now turn to brief examination of STAMP morph constructions and historically-related prefixal conjugations in Chadic, Central Sudanic, and certain Niger-Congo families.

---

[4] Seidel (1898: 211) pointed out that in Kwa languages of Togo "die verbalen Präfixe, wahrscheinlich Reste ehemahliger Hilfsverben, verschmelzen nicht selten mit den Pronominalpräfixen zu einer Silbe". What the status of STAMP morphs might be prosodically in a given language in this region bears close examination, as they may turn out to be clitics or prefixes. To be sure, Creissels et al. (2008: 93) caution that "many descriptions of African languages do not identify pronominal markers appropriately, treating them as independent words". Such is definitely the case in the East Mande Boko/Busa cluster (Jones 1998), where the orthography treats STAMP morphs as freestanding elements but phonologically they are prefixes. As

7 STAMP morphs and prefix conjugations in Chadic languages

The next three sections briefly present STAMP morphs in three of the representative stocks found in the Macro-Sudan Belt (each consisting of several genetic units for typological sampling). STAMP morphs are a characteristic of the entire Chadic macro-family.[5] Simplex formations of the most basic type (10) are characteristic of West Chadic languages like Angas (41) and Gerka (42).

(41) Angas (Burquest 1973/1980: 38) [West Chadic]
ŋâː jì.
1.COMPL come
'I have come.'

(42) Gerka (Jungraithmayr 1968/1969: 173)
kà tà ʄàm.
2.PRF drink water
'You have drunk water.'

On the other end of the spectrum, in some Biu-Mandara Chadic languages, e.g., Vamé (43), one now finds only bound prefix conjugations which have derived from STAMP morph constructions.

(43) Vamé (Kinnaird 2006: 11) [Biu-Mandara Chadic]
ə̀ŋ-lɛ.
FUT.1-go
'I will go.'

A common pattern seen throughout the Macro-Sudan Belt is for a perfective series to be bound prefixal conjugations, but an imperfective series to remain in freestanding STAMP morph constructions. An example of such a system is found in the Biu-Mandara Chadic language Mbuko (44).

(44) Mbuko (Gravina 2001: 7) [Biu-Mandara Chadic]

  a. nə̀-zlàmbál.
    1PFV/ANT-throw:ANT
    'I threw.'

---

a whole, one can find certain variation between Anglophone vs. Francophone meta-analytic traditions in the Africanist literature, with Anglophone analysts preferring more prefixes in their descriptions while Francophone scholars rather often favor analyses with pre-verbal free-standing elements. This is not an absolute.

[5] It is important to not confuse STAMP morphs with a series of intransitive copy pronouns that are also characteristic of Chadic (Frajzyngier 1977; Burquest 1986). They are unrelated phenomena.

    b. nī   zlāmbāl.
       1.ɪᴘꜰᴠ throw

       'I am throwing.'

Merey (also Biu-Mandara Chadic) shows an unusual intra-pardigmatic split between free-standing forms in the first singular past but bound prefix elements in the corresponding third person (45) (masculine singular) past form.

(45) Merey (Gravina 2007: 8) [Biu-Mandara Chadic]

    a. na   zal.
       1.ᴘsᴛ call

       'I called.'

    b. a-zal.
       3.ᴘsᴛ-call

       'He called.'

Once bound, such prefix conjugational elements can appear with new auxiliary verbs, as in the following Buduma formation (46).

(46) Buduma (Pawlak 2001: 376), (Lukas 1939: 55) [Biu-Mandara Chadic]
    a-kol   a jai-ni.
    3.ᴘʀs-be at sit-ᴠɴ

    'He is/was sitting.'

While prefix conjugations are archaic features of Afroasiatic languages (Hodge 1971; Schuh 1976; Mukarovsky 1983; Voigt 1987), the ones attested in Chadic languages do not reflect inherited structures (Schuh 1976; Voigt 1989; Jungraithmayr 2005; 2006), but rather reflect secondary developments typically derived from the univerbation of STAMP morphs with following verbs (Caron 2006; Shay 2008; Anderson 2011; 2012).

8 STAMP morphs and conjugational prefixes in Central Sudanic languages

STAMP morphs and prefixed conjugation series derived from STAMP morphs are also found in the various families of the Central Sudanic stock (Anderson 2015). Within the analytic tradition of Ma'di (47), a member of the Moru-Ma'di genetic unit Blackings & Fabb (2003) analyze the freestanding STAMP morph *ka*/*kɔ́* as a pronominal, but Tucker & Bryan (1966) analyze it as an auxiliary. Such disagreement over what traditional part-of-speech has to be assigned to these complexes underscores the unusual but characteristic features of STAMP morphs.

(47) Ma'di (Blackings & Fabb 2003: 13) [Moru-Ma'di]

 a. ká   gbándà ˋɲa.
   3SG.INDEF cassava NPST:eat

  'He eats/is eating cassava.'

 b. kɔ   ˋɲa-ʔa.
   3SG.INDEF NPST:eat-OBJ

  'He eats/is eating it.'

Various Sara-Bongo-Bagirmi languages require infinitive forms of vowel-initial verbs when used with freestanding STAMP morphs, for example, Ngambay (48), Kabba (49) and Kenga (50).

(48) Ngambay (Vandame 1963: 118) [Sara-Bongo-Bagirmi]

 a. m̄ k-ào àl ngà.
   1.FUT INF-go NEG ADV

  'I will not go again.'

 b. á k-ùsā né ngà uà.
   2.FUT INF-eat thing ADV Q

  'You are already going to eat?'

(49) Kabba (Moser 2004: 220) [Sara-Bongo-Bagirmi]
 má k-àw lò tɨ́ àáng.
 1:FUT INF-go place LOC NEG

 'I shall go nowhere.'

(50) Kenga (Neukom 2010: 15) [Sara-Bongo-Bagirmi]
 m̄ k-ɔ̀sɔ̄.
 1.FUT INF-eat

 'I will eat.'

As with Chadic, many languages in different sub-families of Central Sudanic have a bound 'definite' perfective STAMP series that contrasts with an unbound 'indefinite' imperfective series. The pattern is found in Bongo (51), Lugbara (52), Lendu (53) and Oke'bu (54), from four different genetic units of Central Sudanic.

(51) Bongo *mi- < m-i (Tucker & Bryan 1966: 75) [Sara-Bongo-Bagirmi]

 a. mi-bi.
   1.DEF-give

  'I give, I gave.'

 b. ma bi.
   1.INDEF give

  'I am giving.'

(52)　Lugbara (Tucker & Bryan 1966: 47) [Moru-Ma'di]

    a.　á-tsɔ　　　mvá.
       1.DEF-beat child

      'I beat the child.'

    b.　ma　　mvá　tsɔ.
       1.INDEF child beat

      'I am beating the child.'

(53)　Lendu (Tucker & Bryan 1966: 46) [Lendu]

    a.　má-drr　　mbí.
       1.DEF-pull rope

      'I pull the rope.'

    b.　má　　mbi　drŕ.
       1.INDEF rope pull

      'I am pulling the rope.'

(54)　Oke'bu (Tucker & Bryan 1966: 48) [Mangbutu]

    a.　l-ómá　　ùnzu.
       2.DEF-beat child

      'You beat the child.'

    b.　láà　　unzú　òma.
       2.INDEF child beat

      'You are beating the child.'

As the reader may have noticed, there is also a difference in the basic clausal syntax of the two series in these Central Sudanic languages, with VO order in the definite/perfective series and OV in the indefinite/imperfective one. The VO order in the definite series vs. OV order in the indefinite is, however, not universal in Central Sudanic and is not found, for example, in Baka (55) of the Sara-Bongo-Bagirmi family or Kresh (56).

(55)　Baka (Tucker & Bryan 1966: 75) [Sara-Bongo-Bagirmi]

    a.　m-áne　　yí.
       1.DEF-eat thing

      'I ate (something).'

    b.　má　　y-ane　yi.
       1.INDEF INF-eat thing

      'I am eating (something).'

(56) Kresh (Tucker & Bryan 1966: 75) [Kresh]

    a. m-omò nòmò.
       1-drink drink

       I drink a drink.'

    b. ma   y-òmò   nòmò.
       1.INDEF INF-drink drink

       'I am drinking.'

Note that these two languages however also have infinitive forms of the verb with vowel initial stems with the indefinite series STAMP morph, like those in (48)-(50). This feature further underscores the likely origin of STAMP morph in an auxiliary verb construction in Central Sudanic languages.

Of course, once a bound STAMP morph prefix exists in a given Central Sudanic language, it is free to attach to auxiliary verbs in new AVCs, as seen in the following Lugbara sentence (57).

(57) Lugbara (Tucker & Bryan 1966: 46, 47) [Moru-Ma'di]
    ma-ŋga   mvá tsɔ.
    1.DEF-AUX child beat

    'I shall beat the child.'

Variation in closely related varieties between bound and unbound formations can be seen in Sara-Bongo-Bagirmi languages. Compare in this regard Furu (58), where a free-standing STAMP morph is found, with the corresponding sentence in its close sister language Bagiro (59), where it is a prefix.

(58) Furu (Boyeldieu 1990: 91) [Sara-Bongo-Bagirmi]
    mí gáli   gɔ.
    I   know NEG

    'I don't know.'

(59) Bagiro (Boyeldieu 1990: 91) [Sara-Bongo-Bagirmi]
    mú-gá⁺li gɔ̀.
    1-know   NEG

    'I don't know.'

Mangbetu languages only have bound prefixes, but nevertheless reflect a possible trace of two originally distinct STAMP sets, albeit with the opposition having become phonologized in such languages as Mangbetu and Meje. Thus vowel initial stems take the second singular prefix *ni-* (60)-(61) while consonant-initial stems take *(m)ú-* (62)-(63):

(60) Mangbetu (Larochette 1958: 106) [Mangbetu]
    ni-ɛ́si-(a).
    2.PRS/FUT-do-TAM

    'You (will) do.'

(61) Meje (Larochette 1958: 106) [Mangbetu]
ni-έsí-a.
PRS/FUT-do-TAM
'You (will) do.'

(62) Mangbetu (Larochette 1958: 106-7) [Mangbetu]
mú-ta.
2.PRS/FUT-carry
'You (will) carry.'

(63) Meje (Larochette 1958: 106-7) [Mangbetu]
ú-ta.
PRS/FUT-carry
'You (will) carry.'

The Mangbutu language Lese also has only bound prefix conjugations. Again these appear to reflect two originally distinct series, such as the present progressive in the first singular in *má-* (64) and the future progressive first singular in *mɔ́-* (65).

(64) Lese (Tucker & Bryan 1966: 51) [Mangbutu]
má-dzɔ̀   kɔ̀dí̃ à-nɔ́.
1:TAM-AUX meat DEP-eat
'I am eating meat.'

(65) Lese (Tucker & Bryan 1966: 51) [Mangbutu]
mɔ́-dzá   kɔ̀dí̃ à-nɔ̀.
1:TAM-AUX meat DEP-eat
'I shall be eating meat.'

While widespread and hence tempting to reconstruct the bound perfective series and unbound imperfective series pattern all the way back to Proto-Central Sudanic, the overall frequency of such patterning in the Macro-Sudan Belt suggests that we should do so only with caution. Nevertheless, I tentatively reconstructed this to Proto-Central Sudanic (Anderson 2015). This is because the sub-families that lack this pattern have two bound series, and never the opposite patterning, such that bound imperfective series appear to be attested only if bound perfective series exist; this distribution is also found across the Benue-Congo and Chadic languages of the Macro-Sudan Belt.

9 STAMP morphs and conjugational prefixes in Niger-Congo languages

STAMP morphs and conjugational prefixes derived from STAMP morphs are found in a wide range of Niger-Congo languages. These include Klao of the Kru family (66), Dadiya

of the Waja-Kam family (67), Gã of the Ga-Dangme Kwa genetic unit (68), and Kulango (69).

(66) Klao (Marchese 1982: 3, 18) [Kru]
ɔ́ɔ̄   blē.
he:IPFV sing
'He is singing.'

(67) Dadiya (Jungraithmayr 1968/1969: 196) [Waja-Kam]
ń   já.
you.PFV eat.PFV
'You have eaten.'

(68) Ga (Kropp Dakubu 1988: 105) [Ga-Dangme Kwa]
èe   nṵ̀   nṵ̀.
3:PRS water drink
'He is drinking water.'

(69) Kulango (Elders 2007: 193) [Kulango]

  a. mí̩   dɔ́lì.
    1.SBJN sell
    'May I sell.'
  b. mí̩ì   dɔ̀lì.
    1.PROG sell
    'I am selling.'

Such formations are also widespread in Mande languages, such as the following example from Jo(wulu) of the Northwest Mande genetic unit (70).

(70) Jo (Jowulu) (Kim 2002: 11) [Northwest Mande]
kaa   ku-ki kulu nũ.
3PL.FUT *to*-DEF hot   eat
'They will eat hot *to*.'

Within Benue-Congo there is evidence for two (possibly three) sets of contrastive STAMP morphs series.[6] One is realized as a bound prefix series, while the other(s) remain(s) freestanding STAMP morphs in most but not all relevant languages. These series consist of a bound realis/perfect(ive) series marked by **ma-* (possibly with an associated low tone **mà-*) in the first singular, which contrasts with one (or two) other *m*-initial series with front vowels (or no vowel) marking irrealis/imperfective action. Thus, within Benue-Congo itself, there is some evidence for the two non-perfect(ive) series. However,

---

   [6] The reconstructions for Proto-Benue-Congo are preliminary and impressionistic. Space limitations prevent further demonstration of the details of the complex Benue-Congo situation.

the broader comparative data from Atlantic-Congo (or Volta-Congo) languages suggest that there were originally two sets of forms, and one may well have split into two contrasting series during the development of Benue-Congo, or, alternatively, in the history of individual sub-groups within Benue-Congo. However, they likely originated in a single imperfective/irrealis series from an earlier state of the language. One of these sets generally has a high-tone in the first singular and marks irrealis or future semantics. The first singular marker in the other set appears optionally without a vowel and with a non-high (mid or low) tone, and encodes present/non-past/imperfective semantics (71).

(71) Proto-Benue-Congo
       Pattern 1 *ma- I.ᴘʀꜰ/ᴘꜱᴛ/ʀʟꜱ : *mé I.ɪʀʀ/ꜰᴜᴛ
       Pattern 2 *ma- I.ᴘʀꜰ/ᴘꜱᴛ/ʀʟꜱ : *mé I.ɪʀʀ/ꜰᴜᴛ : *mi, *m(V) I.ᴘʀꜱ/ɴᴘꜱᴛ

Evidence for the *ma- (~*mà-) perfective/realis/past series in Benue-Congo (71) comes from a range of different sub-groups of the stock. Thus one finds formally and functionally cognate elements in such Benue-Congo languages as Bantoid Ndemli (72), Eleme of the Cross River family (73), and Berom of the Plateau sub-stock (74).

(72) Ndemli (Lenaka 1999: 72) [Bantoid]
       mà-tóm.
       I.ᴘꜱᴛ-send
       'I sent.'

(73) Eleme (Bond 2009: 1482), (Bond 2010) [Cross River]
       ma-ʔà.
       1.ᴀɴᴛ.ᴘʀꜰ-leave
       'I left.'

(74) Berom (Bouquiaux 1970: 299, 301) [Plateau]
       mà-ciŋ.
       1.ᴘꜱᴛ/ᴘʀꜰ-dig
       'I (have) dug.'

Evidence for the high-toned irrealis or future series marked in the first singular by *mé can be found in a range of Benue-Congo languages, e.g., the Bantoid language Tiv (75), the Kainji language Duka (76) or Berom of the Plateau stock (77).

(75) Tiv (Arnott 1967/1980: Tiv 4) [Bantoid]
       mé ꜛva.
       1ꜰᴜᴛ come
       'I will come.'

(76) Duka (Bendor-Samuel, Skitch & Cressman 1973: 17) [Kainji]
       mɛ́ róà sə á.
       I.ɪʀʀ ʀᴇᴍ.ꜰᴜᴛ drink ɴᴇɢ
       'I would not drink it.'

(77) Berom (Bouquiaux 1970: 300) [Plateau]
mé   ciŋ.
1.SBJN dig
'I must dig.'

Evidence for a third series of STAMP morphs in Proto-Benue-Congo is more tenuous and must remain beyond of the scope of the present study.

10 Summary

This study has been a preliminary survey of STAMP morph constructions in a range of genetic linguistic units spoken across the Macro-Sudan Belt. A comparative analysis of STAMP morphs across the languages of the Macro-Sudan Belt, and the prefixal conjugations that historically derived from them, yields significant insights into various layers in the history of verbal conjugation in many genetic units. Perhaps the most noteworthy insight is that there is an absolute pattern whereby, if present, a bound series encodes perfective/realis categories and an unbound series rather encodes imperfective/irrealis categories. This situation is found in Chadic, Central Sudanic and Benue-Congo languages alike. This pattern has been reconstructed to Proto-Central Sudanic (Anderson 2015). One can also reconstruct a probably bound perfective/realis conjugation series (e.g., with first singular *ma-/=) for Proto-Benue-Congo, but its immediate ancestral formation may still have been an unbound STAMP morph. The corresponding imperfective/irrealis series (e.g., with first singular *mé ~ *mI ~ *mÍ) appears to have certainly been an unbound STAMP morph in the proto-language. As in Central Sudanic, Benue-Congo languages only have bound imperfective series STAMP morphs if perfective series elements are also bound. The precise modeling of the Chadic developments is the object of current ongoing research, but the same distributional trend appears to be manifested across the Chadic languages as well.

While insights into the inflectional history of the genetic units in this macro-region are gained by such an analysis (e.g. how secondary prefix conjugations arose in Central Sudanic languages), others remain unanswered to date. The most salient of such questions is the following: Why is it that if bound/prefixal conjugations are found which appear to derive from STAMP morph constructions, in Chadic, various Niger-Congo families and Central Sudanic languages across the Macro-Sudan Belt the pattern is always one where the bound series encodes perfective/realis and the unbound series imperfective/irrealis categories, but never the reverse situation? Resolving this important and intriguing question remains a primary goal of future research.

Abbreviations and symbols

1,2,3	1st, 2nd, 3rd person		IRR	irrealis
>	acting on		+L	low tone
ACC	accusative		LOC	locative
ADV	adverb		MOD	modal
ANT	anterior		NEG	negative
AOR	aorist		NPST	non-past
ASP	aspect		OBJ	object
AUX	auxiliary		OBLQ	oblique
CLSFR	classifier		PFV	perfective
COMPL	completive		PL	plural
COND	conditional		PRF	perfect
CONT	continuative		PROG	progressive
COP	copula		PRS	present
DEF	definite		PST	past
DEP	dependent		Q	interrogative
FUT	future		REC	recent
+H	high tone		RDPL	reduplication
IO	indirect object		REM	remote
HAB	habitual		RLS	realis
IMM	immediate		SBJN	subjunctive
INDEF	indefinite		SG	singular
INF	infinitive		TAM	tense-aspect-mood
INJ	injunctive		VN	verbal noun
IPFV	imperfective			

References

Anderson, Gregory D. S. 2006. *Auxiliary verb constructions*. Oxford: Oxford University Press.

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. 2012. S/TAM/P morphs in the history of Benue-Congo and Niger-Congo conjugation. (Paper presented at Niger-Congress, Paris, September 2012.)

Anderson, Gregory D. S. 2015. STAMP morphs in Central Sudanic languages. In Angelika Mietzner & Anne Storch (eds.), *Nilo-Saharan: Models and descriptions*, 151–167. Köln: Rüdiger Köppe Verlag.

Arnott, David W. 1958. The classification of verbs in Tiv. *Bulletin of the School of Oriental and African Studies* 21(1). 111–133.

Arnott, David W. 1967/1980. Tiv. In Mary Esther Kropp-Dakubu (ed.), *West African language data sheets volume 2*, 8 pp. Accra: West African Linguistic Society.

Babaev, Kirill. 2010. Reconstructing Benue-Congo person marking II. *Journal of Language Relationship/Voprosy Jazykovogo Rodstva* 4. 1–45.

Bendor-Samuel, John, Donna Skitch & Esther Cressman. 1973. *Duka sentence, clause and phrase* (Studies in Nigerian Languages 3). Zaria: Institute of Linguistics and Centre for the Study of Nigerian Languages, Abdullahi Bayero College, Ahmadu Bello University, Kano.

Blackings, Mairi & Nigel Fabb. 2003. *A grammar of Ma'di* (Mouton Grammar Library 32). Berlin: Mouton de Gruyter.

Bond, Oliver. 2009. Multiple analytical perspectives of the Eleme anterior-perfective. *Current issues in unity and diversity of languages: Papers selected from the 18th International Congress of Linguistics, July 21-26, 2008.* 1480–1496.

Bond, Oliver. 2010. Intra-paradigmatic variation in Eleme verbal agreement. *Studies in Language* 34. 1–35.

Bouquiaux, Luc. 1970. *La langue Birom (Nigeria septentrional) – phonologie, morphologie, syntaxe.* Paris: Les Belles Lettres.

Boyeldieu, Pascal. 1990. Furu et Bagiro: L'extension oubanguienne des langues «sara». *Journal des Africanistes* 60(1). 77–106.

Burquest, Donald. 1973/1980. Angas. In Mary Esther Kropp-Dakubu (ed.), *West African language data sheets volume 1*, 35–43. Accra: West African Linguistic Society.

Burquest, Donald. 1986. The pronoun system of some Chadic languages. In Ursula Wiesemann (ed.), *Pronominal systems*, 71–101. Tübingen: Gunter Narr Verlag.

Caron, Bernard. 2001. Dott, aka Zoɗi, (Chadic, West-B, South Bauchi): Grammatical notes and vocabulary. *Afrika und Übersee* 84. 1–60.

Caron, Bernard. 2002. Guus, aka Sigidi (Chadic, West-B, South Bauchi): Grammatical notes, vocabulary and text. *Afrika und Übersee* 85. 161–248.

Caron, Bernard. 2006. South Bauchi West pronominal and TAM systems. In Bernard Caron & Peter Zima (eds.), *Sprachbund in the West African Sahel*, 93–112. Louvain/Paris: Peeters.

Caron, Bernard. 2008. La structure énonciative des subordonnées conditonnelles. In Bernard Caron (ed.), *Subordination, dependence et parataxe dans les langues africaines*, 145–158. Louvain/Paris: Peeters.

Cleal, Alizon M. 1973/1980. Nchumuru. In Mary Esther Kropp Dakubu (ed.), *West African language data sheets volume 2*, 11 pp. Accra: West African Linguistic Society.

Cook, Thomas L. 1972/1980. Kohumono. In Mary Esther Kropp Dakubu (ed.), *West african language data sheets volume 1*, 350–356. Accra: West African Linguistic Society. (1977).

Creissels, Denis. 2005. A typology of subject and object markers in African languages. In F. K. Erhard Voeltz (ed.), *Studies in African linguistic typology* (Typological Studies in Language 64), 43–70. Amsterdam: John 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*, 86–150. Cambridge: Cambridge University Press.

Dimmendaal, Gerrit J. 2008. Language ecology and linguistic diversity in Africa. *Language and Linguistics Compass* 2(5). 840–858.

Dimmendaal, Gerrit J. 2011. *Historical linguistics and the comparative study of African languages.* Amsterdam/Philadelphia: John Benjamins.

Elders, Stephan. 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), 187–200. London: University of London.

Èrman, Anna V. 2002. Sub"ektnye mestoimenija v dan-blovo i modal'no-aspektno-temporal'nye znachenija [Subject pronouns in Dan-Blowo and (their) modal-aspectual-temporal meanings]. In Valentin F. Vydrin & Aleksandr Ju Zheltov (eds.), *Juzhnye mande: Lingvistika afrikanskikh ritmakh. Materialy peterburgskoj èkspeditsii v Kot d'Ivuar (k 50-letiju Konstantina Pozdnjakov)*, 154–82. Sankt-Petersburg: Evropejskij Dom.

Fiore, Lynne & Patricia Peck. 1973/1980. Limbum. In Mary Esther Kropp Dakubu (ed.), *West African language data sheets volume 2*, 1–8. Accra: West African Linguistic Society.

Frajzyngier, Zygmunt. 1977. On the intransitive copy pronouns in Chadic. *Studies in African Linguistics, Supplement* VII. 73–84.

Frajzyngier, Zygmunt. 2008. *A grammar of Gidar* (Research in African studies 13). Frankfurt: Peter Lang.

Gravina, Richard. 2001. The verb phrase in Mbuko. Yaoundé: SIL.

Gravina, Richard. 2007. The verb phrase in Merey. Ministère de la recherche scientifique et de l'innovation. 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*, 151–185. Cambridge: Cambridge University Press.

Hodge, Carleton T. (ed.). 1971. *Afroasiatic: A survey* (Janua Linguarum Series Practica 163). The Hague: Mouton.

Hofer, Verena & Christa Link. 1973/1980. Wobé. In Mary Esther Kropp Dakubu (ed.), *West African language data sheets volume 2*, 8 pp. Accra: West African Linguistic Society.

Jones, Ross McCallum. 1998. *The Boko/Busa language cluster* (Lincom Studies in African Linguistics 30). Munich: Lincom Europa.

Jungraithmayr, Herrmann. 1968/1969. The class languages of the Tangale-Waja district. *Afrika und Übersee* 52. 161–204.

Jungraithmayr, Herrmann. 2005. Prefix and suffix conjugation in Chadic. In *Proceedings of the 10th Meeting of Hamito-Semitic (Afroasiatic) Linguistics*, 411–419.

Jungraithmayr, Herrmann. 2006. The verb in Chadic–state of the art. In Bernard Caron & Peter Zima (eds.), *Sprachbund in the West African Sahel*, 167–183. Louvain/Paris: Peeters.

Kim, Hae-Kyung. 2002. Aspect, temps et modes en jowulu. Bamako: SIL Mali.

Kinnaird, William J. 2006. The Vamé verbal system. Yaoundé: SIL.

Kropp Dakubu, Mary Esther. 1988. *The languages of Ghana.* London: Kegan Paul International for the International African Institute.

Larochette, Joe. 1958. *Grammaire des dialects 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.

Lenaka, Ngoran Loveline. 1999. *A sketch outline of the phonology of Ndemli.* Yaoundé: University of Yaoundé (Master's Thesis.)

Lukas, Johannes. 1939. *Die sprache der Buduma im Zentral Sudan. Auf Grund eigener Studien und des Nachlasses von Gustav Nachtigal.* Leipzig: F. A. Brockhaus.

Marchese, Lynell. 1982. Basic aspectual categories in Proto-Kru. *Journal of West African Languages* 12(1). 3–23.

Moñino, Yves. 1995. *Le Proto-Gbaya. Essai de linguistique comparative sur vingt-et-une langues d'Afrique centrale.* Paris: Peeters.

Moser, Rosemarie. 2004. *Kabba: A Nilo-Saharan language of the Central African Republic* (LINCOM studies in African linguistics 63). München: LINCOM.

Mukarovsky, Hans G. 1983. *Pronouns and prefix conjugation in Chadic and Hamito-Semitic.* Hamburg: Helmut Buske Verlag.

Nettle, Daniel. 1998. Materials from the southeastern Plateau languages of Nigeria. *Afrika und Übersee* 81. 253–279.

Neukom, Lukas. 2010. *Description grammaticale du kenga (langue nilo-saharienne du Tchad).* Köln: Rüdiger Köppe.

Pawlak, Nina. 2001. Diachronic typology of locative phrases in Chadic. *Sprache und Geschichte in Afrika* 16/17. 355–386.

Pichl, Walter J. 1973/1980. Ndut-Falor. In Mary Esther Kropp Dakubu (ed.), *West African language data sheets volume 2*, 8 pp. Accra: West African Linguistic Society.

Pohlig, James N. 1992. An account of Mofu-Gudur verb infixes and suffixes. (Ministry of Scientific and Technical Research). Yaoundé: SIL.

Price, Norman. 1975/1980. Nchumuru. In Mary Esther Kropp Dakubu (ed.), *West African language data sheets volume 2*, 11 pp. Accra: West African Linguistic Society.

Sands, Bonny. 2009. Africa's linguistic diversity. *Language and linguistics compass* 3(2). 559–580.

Schuh, Russell G. 1976. *The Chadic verbal system and its Afroasiatic nature.* Malibu, CA: Undena Publication.

Seidel, August. 1898. Beiträge zur Kenntnis der Sprachen im Togo. *Zeitschrift für afrikanische und oceanische Sprachen* 4. 201–286.

Shay, Erin. 2008. Coding the unexpected. Subject pronouns in East Dangla. In Zygmunt Frajzyngier (ed.), *Interaction of morphology and syntax. Case studies in Afroasiatic*, 85–105. Amsterdam: John Benjamins.

Shimizu, Kiyoshi. 1983. *The Zing dialect of Mumuye. A descriptive grammar. With a Mumuye-English dictionary and an English-Mumuye index.* Hamburg: Helmut Buske Verlag.

Sibomana, Leo. 1981/1982. 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.

Voigt, Rainer M. 1987. The two prefix-conjugations in East Cushitic, East Semitic, and Chadic. *Bulletin of the School of Oriental and African Studies* 50(2). 330–345.

Voigt, Rainer M. 1989. Verbal conjugation in Proto-Chadic. In Zygmunt Frajzyngier (ed.), *Current progress in Chadic linguistics: Proceedings of the international symposium on Chadic linguistics, Boulder, Colorado, 1-2 May, 1987. Vol. 62*, 267–284. Amsterdam: John Benjamins.

Vydrine, Valentin F. 2009. Negation in southern Mande. In Norbert Cyffer, Erwin Ebermann & Georg Ziegelmeyer (eds.), *Negation patterns in West African languages* (Typological Studies in Language 87), 223–260. Amsterdam/Philadelphia: John Benjamins.

Vydrine, Valentin F. 2011. Ergative/absolutive and active/stative alignment in West Africa. *Studies in Language* 35(2). 409–443.

Wilkendorf, Patricia. 2001. Sketch grammar of Nɔmaándɛ́: Sections 1-4. Yaoundé: SIL.

Chapter 27

Igbo-English intrasentential codeswitching and the Matrix Language Frame model

Kelechukwu Ihemere

University of Westminster, London UK

This paper uses data from Igbo-English intrasentential codeswitching involving mixed nominal expressions to test the Matrix Language Frame (MLF) model. The MLF model is one of the most highly influential frameworks used in the study of grammatical aspects of codeswitching. Three principles associated with it, the Matrix Language Principle, the Asymmetry Principle and the Uniform Structure Principle, were tested on data collected from informal conversations by educated adult Igbo-English bilinguals resident in Port Harcourt. The results of the analyses suggest general support for the three principles and for identifying Igbo-English as a "classic" case of codeswitching.

1 Introduction

It goes without saying that codeswitching (CS) is conceivably one of the most studied and discussed outcomes of language contact. Its study has been approached from two main dimensions: (i) the sociolinguistic and (ii) the linguistic. Researchers concerned with the sociolinguistic aspects of CS tend to seek to identify patterns of occurrence of CS and the impact of social-psychological factors on this bilingual behavior. Those working from a purely linguistic dimension focus on the structural aspects of CS; their aim is to uncover the syntactic and morphological characteristics of codeswitched utterances. Here, the focus is on the purely linguistic study of CS. Several constraints have been proposed by a number of researchers interested in the linguistic study of CS.

For instance, Poplack & Meechan (1998) and their associates are interested with formulating constraints on points in a sentence where CS can take place on the grounds of surface-level linear differences between the languages concerned. These researchers view restrictions on CS along the lines of dissimilarities in word order, either across clauses (intersentential) or on phrases within clauses (intrasentential). In particular, Poplack's Equivalence Constraint (EC) is based on the premise that switching is not permitted when the syntax of two languages does not match at a potential switch point.

Researchers such as Bhatt (2001) have put forward many counter-examples to such models.[1]

Another group of researchers looks for explanations at a more abstract level than linear structure. These researchers (Belazi, Rubin & Toribio 1994; Di Sciullo, Muysken & Singh 1986; MacSwan 2009) do this by structuring their explanations along the lines of what are considered generative theories of syntax. They assert that the grammatical organization of CS can be accounted for in terms of the principles of current syntactic theories, even though these theories were initially formulated to explain monolingual data. They do not recognize any theoretical (or useful) value in recognizing the asymmetry between a Matrix Language (ML) and an Embedded Language (EL). These approaches differ from the Matrix Language Frame (MLF) model (Myers-Scotton 1993; 2002; 2013).

The MLF model makes the case for a distinction between the ML and the EL. The ML plays a dominant role in shaping the overall morphosyntactic properties of codeswitched utterances. In other words, the model posits two hierarchies in reference to mixed constituents: both languages do not participate equally; only one language is the source of the abstract morphosyntactic frame. This language (and the frame) is called the ML and the other language is called the EL. This idea is formalized as the Morpheme Order Principle (MOP) and the System Morpheme Principle (SMP) of the MLF model. These are testable hypotheses referring to the existence of asymmetry between the languages implicated in CS. On the basis of these principles, only one language (the source of the frame) supplies both morpheme order and frame-building system morphemes to the frame (see §3 for more details on the MLF model).

Accordingly, this study assesses the veracity of this notion of asymmetry in the roles played by the languages participating in CS using some Igbo-English data involving mixed nominal expressions. The rationale for focusing on mixed nominal expressions for the analyses reported in this paper stems from the fact that the languages differ in the relative order of head (H) and complement (C) within the nominal argument phrase — NP (or what is now termed determiner phrase – DP, after Abney 1987). In so doing, this study contributes directly to research on the linguistic analysis of CS by showing: (i) through exemplifications from Igbo-English bilingual discourse what happens to the grammatical structures when two languages are in contact in the same clause; (ii) that CS does not crash when the surface structures of two languages do not map onto each other; (iii) that CS is possible between a functional head and its complement; and (iv) that Igbo-English is a classic case of CS. Classic CS includes elements from two (or more) language varieties in the same clause, but only one of these varieties is the source of morphosyntactic frame for the clause (Myers-Scotton 2002: 8).

The rest of the paper is structured as follows. §2 introduces the type of CS that is the focus of our analysis. §3 focuses on the MLF model and its associated principles, which are tested in this paper. §4 introduces the speakers, sampling and transcription procedures, and the bilingual data. §5-7 present the analysis and discussion of the mixed nominal expressions in Igbo-English CS. §8 is the conclusion to the paper.

---

[1] See also the examples from Igbo-English CS in §4-§6 which appear to falsify the EC.

2 Types of codeswitching

It is often the case that researchers make a distinction between intersentential and intrasentential switching (see Clyne 2003). Intersentential switching exemplifies the employment of different languages at sentence or clause boundaries, as in (1), whereas intrasentential switching involves the coding of different elements within a single clause in different languages, as in (2-5). All examples are from Igbo-English CS; the switched element is in bold font.

(1) Igbo and English intersentential switching[2]
     le   nü ihe   o me-re [PAUSE] **Did you see that**?
     look at thing he do-IND   did you see that
     'Look at what he did. Did you see that?'

In (1), the example includes full sentences in both Igbo and English. Each sentence is a single clause: *Did you see that?* is a sentence in English and it is a single clause; *Le nü ihe o mere* 'look at what he did' is a sentence in Igbo and it is a single clause. Within each clause there is no switching of languages, but there is switching between the clauses. This type of CS is not particularly interesting for researchers concerned with the purely linguistic analysis of CS because the two languages are not in contact in the same clause, unlike in the following examples:

(2) Igbo-English CS
     ha   ga-anö   **week** abüö   na   Abuja.
     they AUX-stay week NUM/D PREP Abuja
     'They will stay for two weeks in Abuja.'

(3) Igbo-English CS
     eze   anyï bïa-ra   gbaghee **hospital** ahü.
     chief our come-IND open   hospital DEM/D
     'Our chief came and opened that hospital.'

(4) Igbo-English CS
     ï-ga-abia   ï-hü   **councilor** ha   na   **opening ceremony** a?
     CL-AUX-come INF-see councilor PRN/D PREP opening ceremony DEM
     'Will you come to meet their councillor at the opening ceremony?'

(5) Igbo-English CS
     ha   küda-ra   **booth** dum na   **polling station** ahü.
     they break-IND booth Q   PREP polling station DEM/D
     'They broke all (the) booths at that polling station.'

In this study, we shall focus on mixed nominal expressions within a single clause, as illustrated in (2-5), and shall take as our unit of analysis the bilingual clause. This will be defined as a clause containing one or more morphemes from more than one language.

---

[2] A list of abbreviations used in the codeswitching glosses is placed at the end of the paper.

3 The Matrix Language Frame model

The MLF model was first articulated by Myers-Scotton in her book *Duelling languages* in 1993. The model posits that the key to understanding feature mismatches in CS is to recognize one of the asymmetries in language that is especially evident in CS: structural conflicts are resolved in favor of only one of the participating languages. According to Jake, Myers-Scotton & Gross (2002: 72), the model captures this generalization in its theoretical assumptions about the nature of linguistic competence and also about operations involved in language production. This view is conceptualized under what they term a Uniform Structure Principle (USP) and its corresponding two hierarchies that indicate how the model relates to linguistic competence.

The USP: A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this type must be observed whenever the constituent appears. In bilingual speech, the structures of the ML are always preferred, but some embedded structures are allowed if ML clause structure is observed (Myers-Scotton 2002: 8–9).

When this principle is applied to bilingual speech, it gives rise to the first hierarchy, which states that in bilingual speech, the languages involved do not participate equally: one language uniformly sets the morphosyntactic frame and this frame is referred to as the ML. The second of the two hierarchies of the USP is the distinction in the MLF model between the roles of content morphemes (similar to lexical elements) and system morphemes (similar to functional elements).

Content morphemes (e.g. nouns and verbs) are those that either assign or receive thematic roles; they refer to relations within the sentence such as whether a noun is the Agent or the Patient of the verb. Under this model, content morphemes along with one type of system morpheme called an early system morpheme, are specifically characterized as conceptually activated. Myers-Scotton (2002) explains that *conceptually activated* means that speaker pre-linguistic intentions activate (or select) content morphemes and any early system morphemes that may accompany them on the surface. This activation occurs at the first level of what is termed *the mental lexicon.*

The mental lexicon is said to consist of elements called *lemmas*[3] that are tagged for specific languages; the speaker's intentions call up language-specific lemmas, which contain the information necessary to produce surface-level forms. Furthermore, lemmas in the mental lexicon that underlie content morphemes are directly activated through the speaker's intention. In turn, these lemmas activate the lemmas underlying early system morphemes. Early system morphemes flesh out the meaning of the lemmas of the content morphemes that call them. These system morphemes are called "early" because of their early activation in the language production process. Examples of early system morphemes (Myers-Scotton 2006: 268) include plural markings, determiners (e.g. the definite article *the* and the indefinite articles *a, an* in English), and those prepositions

---

[3] Lemmas are defined as the morphological and syntactic properties which a word is said to inherently possess, which determine its co-occurrence and selectional restrictions, after Levelt's (1989) Speech Production model.

(also called satellites) that change the meanings of phrasal verbs in certain contexts (e.g. *out* as in *Alice **looks out for** her little brother* or *through* in *the actor **ran through** his lines before the performance*).

The other two types of system morphemes (bridge late system morphemes and outsider late system morphemes) are called "late" because the model claims that they are not activated until a later production level, at a second abstract level that is called *the formulator*. According to the model, the formulator is viewed as an abstract mechanism that receives directions from lemmas in the mental lexicon (those underlying content morphemes); the directions from the lemmas underlying content morphemes tell the formulator how to assemble larger constituents, such as combinations of noun phrases (NPs)/determiner phrases (DPs) and inflection (I)/verb phrases (VPs), resulting in a full clause.

Regarding bridge late system morphemes, they occur between phrases that make up a larger constituent. An excellent example of a bridge is the associative or possessive element that occurs between a possessor noun (N) and the element that is possessed in a number of languages. For instance, *of* is a bridge, as in *the house of Gina*. Also, the model considers the possessive *-s* in English to be a bridge morpheme, as in *Gina's house*. A bridge morpheme depends on the well-formedness conditions of a specific constituent in order for it to appear; such a constituent is not well-formed without the bridge morpheme.

Outsider late system morphemes like bridge late system morphemes also satisfy well-formedness conditions. However, they are said to differ from bridges in that the presence and form of an outsider depends on information that is outside the element with which it occurs and therefore outside its immediate constituent. That information is said to come from an element in another constituent or from the discourse as a whole. Myers-Scotton (2002; 2006; 2013) gives a clear example of an outsider late system morpheme as the element that shows subject-verb agreement on the verb in English. She explains that the form of the agreement marker depends on the subject. Thus, English speakers would say *the dog like-s chewing bones,* but *dogs like-Ø chewing bones.* The suffix *-s* only occurs when there is a third person singular content element in the present tense to call that suffix; otherwise, in English, there is no suffix (Ø = 'zero' marker).

Crucially, the model claims that the frame-building system morphemes in mixed constituents come from only one language, the source of the ML. This theoretical notion is formalized as three testable hypotheses, claimed to be universally applicable in cases involving classic CS:

- **The Matrix Language Principle (MLP)** states that it is always possible in cases of classic CS to identify the ML in a bilingual clause. The ML will be the language supplying the morphosyntactic frame of the clause (Myers-Scotton 2002: 8).

- **The Asymmetry Principle (AP)** states that bilingual speech is characterized by asymmetry in terms of the participation of the languages involved in CS (Myers-Scotton 2002: 9).

- **The Uniform Structure Principle (USP)** states that in bilingual speech, the structures of the ML are always preferred (Myers-Scotton 2002: 8-9).

We shall offer more specific details about how the principles apply to the Igbo-English data in subsequent sections of this paper.

4 Methodology

4.1 Sampling procedure

The sampling was done through personal contacts among speakers who knew each other well and shared the same friendship network in Port Harcourt, Nigeria. Port Harcourt is the capital of Rivers State, and the metropolitan area has a population of over a million, composed of people from different parts of Nigeria, including a large Igbo population. Hence, through pre-existing contacts in the city it was possible to recruit 50 educated adult Igbo-English bilinguals of both sexes for the study. Thirty-eight out of the fifty speakers were university graduates, while twelve were undergraduates at the time of the fieldwork in the summer of 2011. All fifty speakers speak Nigeria Standard English (NSE) as their second language.

4.2 The speakers

The speakers involved ranged in age from 20 to 55, including secondary school teachers (N = 15), engineers (N = 10), physicians (N = 4), nurses (N = 5), business owners (N = 4), and undergraduate students (N = 12). Table 1 summarizes the occupation and sex of the speakers.

Table 1: Distribution of speakers according to occupation and sex

Occupation	Male	Female	Total
Teachers	6	9	15
Engineers	9	1	10
Physicians	4	-	4
Nurses/midwives	-	5	5
Business owners	4	-	4
Undergraduates	6	6	12
Total	29	21	50

All the speakers were born and grew up in Nigeria and have Igbo as their first language. They are bilingual, being native speakers of Igbo and proficient in English. The second language was learned at school age. Having studied in Nigeria, where English

is the official language and the language of all advanced education, the speakers have all been educated almost entirely in English. To ensure their anonymity, the speakers' real names are not used in the examples; where names appear in the examples, these are pseudonyms.

4.3 Data collection and transcription procedure

A number of studies on the grammatical aspects of CS still rely on data sourced from grammaticality judgements of native speakers (see MacSwan 2009). The problem with such studies is that the judgements of two native speakers may vary with respect to a particular utterance; and as Treffers-Daller (1991) noted much earlier, such judgements tend to reflect attitudes towards language mixture rather than grammaticality or otherwise. This is mostly the case in communities where CS is stigmatized – most speakers have been known to judge CS as ungrammatical or to deny the very utterances they have been captured using, documented in recordings (MacSwan 1999: 99). Nevertheless, most studies on the grammar of CS continue to be based mainly on corpora of authentic everyday speech (as in Myers-Scotton & Jake 2014). The present study is based primarily on a corpus of naturally occurring speech involving the fifty men and women described in §4.1 and §4.2. Specifically, the corpus consists of naturally occurring group conversations among members of a friendship network. From twenty minutes to one hour of tape-recorded informal conversations involving each of the speakers were recorded by this researcher in the summer of 2011.

The transcriptions use the normal orthography of Igbo and English, respectively. However, after Echeruo (1998), instead of using subscript dots (.) for the three Igbo closed vowels *i, o*, and *u*, we will use umlauted symbols (ï; ö; ü), which are readily available on standard word processing software. Also, since there is no instance in which "ch" is in complementary distribution with "c" in Igbo, we will use "c" in all Igbo words with a sound similar to the voiceless palato-alveolar affricate [ʧ].

4.4 The bilingual data

The resulting corpus contains substantial examples of different types of CS. However, this paper is concerned with the following CS structures attested in the Igbo-English data:

- Singly occurring embedded language (EL) nouns (Ns)/noun phrases (NPs) in mixed determiner phrases (DPs) (N = 1057/1599), as in (2–5) above.
- Multi-word nominal sequences framed by a Matrix Language (ML) element (N = 192/1599), as in (4–5) above.
- Singly occurring English EL Ns/NPs + Igbo ML Ns/NPs in genitive/associative constructions (N = 165/1599), as in (6).

(6) Igbo-English CS

a-si   na a-ga-eme **wedding** Ngozi ma ö   gbakee.
CL-said C   CL-AUX-do wedding Ngozi C   she recovers

'They said that they will hold Ngozi's wedding when she recovers.'

- Singly occurring English EL Ns/NPs + Igbo ML adjectives (N = 73/1599), as in (7).

(7) Igbo-English CS

ö   na-cö   ï-zü   **portmanteau** öhüü.
she AUX-want INF-buy portmanteau new

'She wants to buy (a/the) new portmanteau.'

- English EL single Ns that occur as bare forms in otherwise Igbo utterances (N = 112/1599), as in (8).

(8) Igbo-English CS

ha   fe-re   **exam** na   Abuja.
they pass-IND exam PREP Abuja

'They passed (the) exam in Abuja.'

In the sections that follow, we test the application of the three principles of the MLF model outlined in §3. This will involve exemplification and illustration of the principles, followed by the results of a quantitative analysis relating to each principle.

5 Testing the MLP

If the principle holds, it will be possible to identify a ML in all bilingual clauses in the Igbo-English data by specific criteria (rather than just assume that Igbo is necessarily the ML). Two specific criteria will be employed to identify the ML of each bilingual clause: (i) the morpheme order criterion; and (ii) the system morpheme criterion. These two criteria follow from two additional principles, the AP and the USP, which are tested in §6 and §7.

5.1 The morpheme order criterion

The morpheme order criterion follows the Morpheme Order Principle (MOP), which predicts that in ML+EL constituents consisting of singly occurring EL lexemes and any number of ML morphemes the surface morpheme order will be that of the ML (Myers-Scotton 2002: 59).

To operationalize this criterion, we will interpret it to mean that it will be applicable wherever the two languages involved in CS have distinct surface orders. This is true of Igbo and English, since they differ in the relative order of head (H) and complement (C) within the nominal argument phrase (NP/DP). The usual order in Igbo is C followed by H rather than the H–C order of English. To illustrate this difference in the configuration of the NP/DP in both languages, consider the monolingual Igbo sentence in (9):

(9) Igbo
　　ha bi na ülö öhüü ahü.
　　they live PREP house new that
　　'They live in that new house.'

In (9), we observe what is typical within the Igbo DP: both the adjective (A) *öhüü* 'new' and determiner (D) *ahü* 'that' are post-posed to the nominal element (N) *ülö* 'house'; the reverse order is usually the case in English.

　　Also, in Igbo a N can follow another N to form a genitival relationship, as in (10):

(10) Igbo
　　ö na-agba igwe Kanye.
　　3SG AUX-ride bicycle Kanye
　　'He rides Kanye's bicycle.'

The situation in (10) is different from that of a language like English, where usually only the N in the genitive case is inflected. Igbo Ns are neither declined for case nor inflected for number like those of English. Therefore, in constructions like (10), it is the genitival N which comes second in the Igbo NP (see Emananjo 1978; Uwalaka 1997). We can illustrate how the morpheme-order criterion would apply to utterance (3), repeated here as (11):

(11) Igbo-English CS
　　eze anyï bïa-ra gbaghee **hospital** ahü.
　　chief our come-IND open　　hospital DEM/D
　　'Our chief came and opened that hospital.'

In this example the EL noun complement *hospital* precedes its Igbo D head *ahü* 'that', reflecting Igbo complement-head (C-H) order and we would thus identify the mixed DP *hospital ahü* as following Igbo order. A similar conclusion would be reached in the case of (7).

　　In (7) we observe that the Igbo true adjective *öhüü* 'new' is post-posed to the EL noun *portmanteau*. This configuration is in sharp contrast with the situation in English, where the order is reversed. Therefore, we would identify the mixed constituent as following Igbo order.

　　A few of the examples involve two Ns in genitival relationship, as in (6) with *wedding Ngozi* 'Ngozi's wedding'.

　　Firstly, we note that English also allows an analytic type of genitive (e.g. 'the wedding of Ngozi') alongside the synthetic type. However our two languages differ in the following ways: (i) in Igbo, the N+N genitive construction does not make use of a bridge morpheme (like *of*) to link the two Ns/NPs; and (ii) Igbo N+N genitive constructions do not include the use of overt determiners; if determiners are used at all, they are always post-posed to the nominal elements. Secondly, looking at the bilingual genitive construction in (6), we observe that unlike what obtains in English, where usually only the N in

the genitive case is inflected, in Igbo, the preceding N is said to be in a pre-genitival position (Uwalaka 1997), while the second N is the possessor. On the strength of this evidence, we will conclude that the word order in (6) reflects that of Igbo.

At first glance, the pre-posed Igbo N *nnukwu* 'big/bigness' in (12a) appears to pose a problem for the morpheme order criterion:

(12) a. Igbo-English CS
obodo   anyï nö na   nnukwu **trouble**.
country our   BE PREP big   trouble

'Our country is in big trouble'

b. Igbo
[NP[N nnukwu]   [N nsogbu]] ~ [NP[N nsogbu] [N nnukwu]]
big/bigness   trouble   trouble   big/bigness

'big trouble'

It is important, however, to point out that the Igbo word *nnukwu* is described by Emananjo (1978: 47-8) and Maduka-Durunze (1990: 237) as a 'qualifactive' noun. These Igbo grammarians argue that the Igbo true adjectives occur only post-nominally, as in (7). Notably, while the Igbo qualifactive nouns functioning as adjectives can occur pre-/post-nominally as in (12b), in English, adjectives typically occur pre-nominally within DP. Therefore, we can submit that when Igbo Ns are used as adjectives, as in (12a), they behave like the adjectives found in English which typically occur pre-nominally because they are in what may be termed associative constructions. Since the surface word order of the mixed NP in (12a) is compatible with that of both languages, we have coded all instances (N = 37/1599) in the data corpus represented by this example as 'either' according to the morpheme order criterion.

Another seemingly problematic case for identifying morpheme order in the Igbo-English data involves English NP compounds framed by a post-posed Igbo functional element; example (5) is repeated here as (13):

(13) Igbo-English CS
ha   küda-ra   **booth** dum na   **polling station** ahü.
they break-IND booth Q   PREP polling station DEM/D

'They broke all (the) booths at that polling station.'

The EL NP *polling station* shows a structural dependency relation that makes it well-formed in English. For instance, *station* heads the nominal sequence pre-modified by the N *polling* that denotes the type of station. Myers-Scotton (2002) argues that such examples do not pose a problem for the MLF model since the other elements surrounding the EL materials follow the MOP. That is, we agree with Myers-Scotton (2002: 139) that such phrases do not pose a problem for the MLF model because the EL multi-word nominal sequence is part of a full DP headed by the post-posed Igbo demonstrative determiner *ahü* 'that' in (13). Thus, with the postposed Igbo functional element, the full DP now has a C-H surface word order in-line with Igbo grammar.

Lastly, we consider the case of English Ns which occur in Igbo utterances with zero (Ø) determiners, as in '(the) exam' of example (8).

In (8), the NP *exam* seems to express some kind of specific reference but without using any determiner. In other words, the NP appears in a context that requires the use of an overt determiner obligatorily in English, but not in Igbo. This claim is supported by the presence of a pre-posed determiner in the monolingual English translation of (8). According to Myers-Scotton & Jake (2001: 106), EL bare forms are content morphemes that occur in a mixed constituent frame prepared by the ML, but that is missing some or all of the required ML system morphemes. Therefore, a compromise strategy is activated and used with the result that the EL content morpheme is not placed in a slot projected by its ML counterpart; rather, it is realized as a bare form or as a part of an EL island.

We, however, disagree with this explanation for why EL bare forms occur with respect to Igbo-English CS. Firstly, the EL N *exam* in (8) is the direct equivalent of its Igbo counterpart which occurs as a bare form in similar contexts. For instance, the English N *exam* is congruent with its Igbo counterpart *ule* in *Ha fe-re ule na Abuja* 'They passed (the) exam in Abuja'. Secondly, the EL N is not inserted with any noticeable compromise strategy either as suggested by Myers-Scotton & Jake (2001: 106). Instead, the N occurs in exactly the same position as its Igbo counterpart. In fact, Igbo already permits 'null determiners' in its grammar (see Obiamalu 2013: 64–65). Given this state of affairs, an alternative explanation can be proffered for the occurrence of EL bare forms in Igbo-English CS.

As a first step, we must account for the variation observed in the bilingual determinate DPs in (2)–(5) which follow Igbo C-H order. By adopting the DP-analysis of mixed nominal expressions in Igbo-English CS, which assumes that the NP is headed by a functional element, the structures where the N/NP precedes the D seem problematic for a theory that assumes that the functional head is higher in the structure and has scope over the NP which it c-commands. According to Kayne (1994), heads must always precede their associated complement position, even though the surface word order in some languages may be H-C (e.g. English) and in some others C-H, as in Igbo. Under this view, in languages like Igbo, the C is said to undergo left adjunction to the specifier (Spec) position. The claim is that the universal ordering between a head and its dependents is Spec-Head-Complement, as represented in (14):

(14)
```
            XP
           /  \
        Spec   X'
              /  \
            X⁰    YP
```
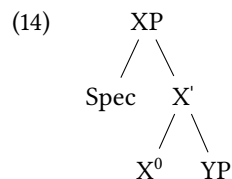
If so, then the bilingual determinate DPs discussed earlier could be analyzed as having the structure in (15). The structure in (15) says that the bilingual determinate DPs are headed by a functional head that takes an NP as C. The NP complement moves to the Spec position in surface syntax giving rise to the C-H order. For instance, the mixed DP *hospital ahü* in (11) will have the structure in (16).

(15)

```
            FP
           /  \
        Spec   F'
         ↑    /  \
         ┆   F    NP
         └ ─ ─ ─ ─ ┘
```

(16)

```
              DP
             /  \
          Spec   D'
           |    /  \
           |   D    NP
           |   |    |
      hospital ahü   tᵢ
```
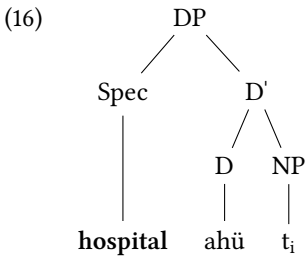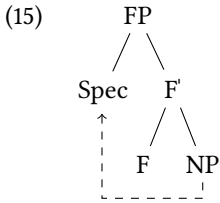
In (16), the N 'hospital' is shown to move into its surface position where it appears before the demonstrative D *ahü* 'that'; thus, creating two possibilities: first, the N head could move to the head of the functional category in a head to head movement; or second, the NP could move to the Spec position of the functional projection (FP). Given that there is no agreement morphology between the N and the associated functional category in Igbo, we assume the latter for the bilingual determinate DPs, as illustrated in (15) and (16). Consequently, to maximize structural symmetry between determinate and indeterminate nominals, we shall assume that the latter are DPs headed by a following null determiner, in line with Igbo grammar. If our supposition about the determinate DPs is correct, then the bare EL form in (8) will have the structure in (17).

(17)

```
             DP
            /  \
         Spec   D'
          |    /  \
          |   D    NP
          |   |    |
        exam  Ø    tᵢ
```
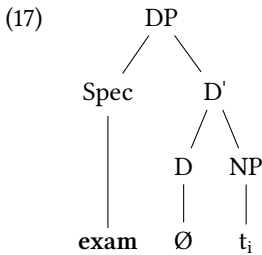
Bare nominals can be interpreted as definite, indefinite or generic, which are features associated with the functional category D (Radford 2004). Therefore, we conclude that in languages like Igbo where there are bare nominals, there is a related null D head which carries the D-features. Moreover, the same analysis can be applied to account for the variation observed in all the mixed nominal expressions presented in this study.

Next, we consider our second criterion for identifying the ML of a bilingual clause.

5.2 The system morpheme criterion

The system morpheme criterion follows from the System Morpheme Principle (SMP), which predicts that in ML+EL constituents, all system morphemes which have grammatical relations external to their head constituent will come from the ML (Myers-Scotton 2002: 59). It is immediately apparent from the way the system morpheme criterion is stated that all system morphemes are not the same in terms of whether they have grammatical relations external to their heads. Recall that in §3 we pointed out that the MLF model makes a distinction between content versus system morphemes. This distinction is on the grounds that content morphemes, such as the lone English origin noun *hospital* in (11) and the Igbo verb *bïa* 'come' in the same example, can both assign or receive thematic roles, whereas system morphemes do not assign or receive thematic roles. Furthermore, system morphemes subdivide into EARLY versus LATE system morphemes, according to whether or not they are conceptually activated or directly linked to the speaker's intentions (see already §3).

The EARLY versus LATE distinction is predicated on assumptions about how early or late in the language production process the relevant morphemes are accessed. Hence, Myers-Scotton suggests that the lemmas underlying early system morphemes, like the post-posed Igbo demonstrative D *ahü* 'that' in (11), are activated when the lemmas supporting content morphemes (such as the EL N *hospital*) point to them. These indirectly elected lemmas further realize the conceptual content of the semantic/pragmatic feature bundles. For example, in (11) *ahü* adds definiteness/specificity to its complement *hospital*. In other words, the same semantic/pragmatic feature bundle activates both *ahü* and *hospital.*

Late system morphemes are then divided into bridge and outsider late system morphemes, the latter being co-indexed with forms outside the head of their maximal projection (Myers-Scotton 2002: 75), while the former are not. That is, outsider late system morphemes are the system morphemes mentioned in §3 which have grammatical relations external to their head. Examples are subject-verb agreement, clitics and case affixes. For our purposes we shall re-define outsider late system morphemes in terms of such grammatical categories as auxiliary verb, tense, aspect, mood, and sentence negation, which are associated with the verb in both languages; rather than in terms of relations outside a morpheme's maximal projection. This is because the concept of maximal projection tends to be theory-specific (cf. Fukui 2001), and also in the case of an analytic language like Igbo, there is no agreement morphology between verb and subject. Additionally, these grammatical categories are perhaps the most frequent kind of "outsider late morphemes" that one can find in both languages, since they occur in most clauses which have finite verbs. Thus, the language source of the earlier mentioned grammatical categories in bilingual clauses containing the mixed nominal expressions should enable us to identify the ML, and this criterion should lead to the same result for each clause as the morpheme-order criterion discussed above. To take an example, the ML of (5/13) was identified as Igbo in §5.1 according to the morpheme order criterion.

We can see that (5/13) is also Igbo according to the system morpheme criterion, since the verbal inflectional morpheme *-ra* (the affirmative indicative past tense suffix) on

'break' also comes from Igbo. The same conclusion can be reached in the additional examples from Igbo-English CS below:

(18) ndï   INEC wepüta-ra   **election results** dum na   TV.
      people.of INEC bring.out-IND election results all   PREP TV

      'The INEC released all the election results on TV.'

(19) **returning officer** ahü a-bü-ghï onye   iberibe.
      returning officer DEM V-BE-NEG person stupid

      'That returning officer is not a stupid person.'

(20) ma **ceremony** ahü fu-ru   ṇnukwu ego.
      but ceremony D   cost-IND big   money

      'but that ceremony cost a lot of money'

(21) **election** afö   a a-dï-ghï   mfe m'ölï.
      election year D V-BE-NEG easy at.all

      'This year's election is not easy at all.'

As can be observed in (18)-(21), all the verbal inflectional morphology (defined as outsider late system morphemes in this study) come from the language determined as the matrix language in §5.1, namely Igbo. In addition, the nominal compounds in (5/13) *polling station*, (18) *election results*, and (19) *returning officer* are termed internal EL islands. According to Myers-Scotton (2006: 265), internal EL islands are part of a larger phrase framed by the ML. All the EL elements qualify as islands because they are phrases within bilingual clauses (they are NPs), and their words show structural dependency relations that make them well-formed in the EL (English). However, in all the examples, the EL morphemes are part of full DPs dominated by Igbo determiners: the demonstrative *ahü* 'that' (in 5/13, 19 and 20), and the quantifier *dum* 'all' (in 18). Crucially, all the mixed constituents support the morpheme order criterion because with the post-posed Igbo determiners the full DPs now have a complement-head order.

Also, we saw in §5.1 that it was sometimes not possible to apply the morpheme order criterion where the word order of a mixed nominal expression was compatible with both Igbo and English (see examples (12), and (22)).

(22) a. Igbo-English CS
          üfödü **students** a-na-ghï a-bïa   na   oge.
          Q   students V-BE-NEG V-come/arrive PREP time

          'Some students do not arrive on time.'

      b. Igbo
          [DP[Q üfödü] [N ümü_akwükwö]] ~ [DP[N ümü_akwükwö] [Q üfödü]]
               some   students   students   some

          'some students'

In examples like (12) and (22) where there is no word order conflict between Igbo and English as far as the switched element is concerned (see already discussion in §5.1 above), we shall consider the source language of the verbal inflectional morphology to identify the ML according to the system morpheme criterion only. It is equally important to underline that *üfödü* 'some' and the numeral *otu* 'one' are the only quantifiers in Igbo that may pre-modify their nouns. *Üfödü* is unique in that, unlike *otu*, it can pre- or post-modify its noun. This is not the case in English, where *some* always occurs in pre-position to the element it modifies. In fact, Maduka-Durunze (1990: 239) observes that the Igbo words *nnukwu* and *üfödü* are qualifactive nouns which, when they precede their nouns, become emphatic in their descriptive meaning or ambiguously suggest an inherent as opposed to a descriptive meaning. Crucially, since the surface word order of the mixed constituents in (12) and (22) are compatible with that of both languages, we have coded all instances in the data corpus represented by these examples as 'either' according to the morpheme order criterion. However, examining the source language of the verbal inflectional morphology, we observe that Igbo and not English is in charge of the outsider late system morphemes in the examples (e.g. the bridge late system morpheme *nö* BE in (12) and the negative suffix *-ghi* in (22)).

As it turns out, there is no case in the sample data under consideration where English supplies the morpho-syntax and Igbo the EL elements. Therefore, if the MLP is correct, then all system morphemes in the same bilingual clause will come from the same source language; also, this source language will be the same as that identified by the first criterion, that of morpheme order (§5.1).

5.3 Quantitative analysis of the MLP

In this section we present the results of a quantitative analysis conducted in order to test the MLP. The two criteria for identifying the ML were applied to each bilingual clause in order to test the MLP, according to which it is always possible in classic CS to identify the ML in a bilingual clause. The combined results are given in Table 2.

Table 2 reveals that the source language of the system morphemes utilized to mark the grammatical categories of tense, aspect and mood is Igbo (100%). 78.1% (N = 1249/1599) of all English Ns/NPs occur with post-posed Igbo determiners. This finding seems to violate the Functional Head Constraint (FHC: Belazi, Rubin & Toribio 1994) and similar CS frameworks (§1), which predict that the language feature of the complement f-selected by a functional head, like all other relevant features, *must* match the corresponding feature of that functional head. As is clear from the analysis in §5.1, switching is not blocked between a functional head (D) and its complement (N/NP) in Igbo-English CS.

According to Myers-Scotton (2002: 59), only if the terms of morpheme order and one type of system morpheme (an outsider late system morpheme) are satisfied by one and the same language, can the ML be identified as that language. On this basis, we determine that the ML of 97.7% of the bilingual clauses is Igbo unequivocally; whereas, the ML of only 2.3% of the bilingual clauses is Igbo according to just the system morpheme criterion. The finding of one language as the overwhelming source of ML in

Table 2: Identification of ML according to morpheme order and system morpheme criteria; English nouns/NPs (N = 1599)

Nominal expressions	Source of morpheme order	Source of outsider late system morpheme	Number	%
English Ns/NPs + Post-posed Igbo Ds	+ Igbo	+ Igbo	1249	(78.1%)
English Ns/NPs + Igbo Ns in genitival relationship	√	√	128	(8%)
English Ns/NPs + Post-posed Igbo As	√	√	73	(4.6%)
English Ns/NPs with zero determiner	√	√	112	(7%)
Pre-posed Igbo Ns + English Ns in associative constructions	Either	√	37	(2.3%)

Igbo-English CS parallels the findings in Hungarian-English CS (Hungarian is the ML: Bolonyai 2005), Ewe-Kabiye CS (Ewe is the ML: Essizewa 2007), and Welsh-English CS (Welsh is the ML: Davies & Deuchar 2010). That is, the evidence from Igbo-English CS confirms that the two languages do not contribute equally in the creation of mixed utterances. In the abstract interaction between the two grammars, the matrix language (Igbo) is more activated than the embedded language (English), resulting in Igbo contributing the morpho-syntactic structure into which English elements are inserted.

6 Testing the AP

As we highlighted in §3, the Asymmetry Principle (AP) states that bilingual speech is characterized by asymmetry in terms of the roles played by the languages involved in CS (Myers-Scotton 2002: 9). This asymmetry is evident in the foregoing analysis reported in Table 2.

- Asymmetry in source of verb inflections (outsider late system morphemes): As the analysis reported in Table 2 clearly indicates, all verb inflections (100%) of the bilingual clauses come from only one of the participating languages, namely Igbo.
- Asymmetry in the resolution of conflict in word order: Again, analysis of the sample reveals that wherever there is a conflict in word order between the two languages, as in 97.7% (N = 1562/1599) of the examples, the order of Igbo (the ML) prevails over that of English (the EL).

- Asymmetry in the supply of content morphemes: The asymmetry in the roles played by both languages in the Igbo-English data extends even to the supply of content morphemes (words). A morpheme count indicates that over 90% of all the content words in the sample under consideration come from Igbo.

These findings illustrate the overwhelming influence of just one language, Igbo, as the source of the morphosyntactic frame in the CS examples.

7 Testing the USP

Recall that the Uniform Structure Principle (USP) states that in bilingual speech, the structures of the ML are always preferred (§3). We have already seen that outsider late system morphemes can only come from the ML of a clause. However, the USP goes further to predict that other system morphemes, such as early system morphemes, which can come both from the ML and the EL, will be drawn preferentially from the ML of a bilingual clause (Myers-Scotton 2002: 8-9). As it turns out, Igbo (ML) contributes the overwhelming majority (98.7%) of all the early system morphemes (e.g. demonstratives, pronominal modifiers, quantifiers, numerals, pronouns, and so on); whereas English (EL) contributes only 1.3% of the early system morphemes in the form of the plural marker *-s* on the EL N *tractor+s* in (23):

(23) Igbo-English CS
      tractors ndï   ahü   emebi-ca-la.
      tractors PRN/D DEM/D damage-ENCL-PERF
      'Those tractors have been damaged completely'

It is important to underline that there is no instance in the Igbo-English data where a lexical noun is in Igbo and the determiner is from English. Thus, the role of the EL is limited to supplying nearly only content morphemes. Broadly, the analyses clearly show that the structures of Igbo (the ML) are preferred in the mixed constituents discussed in this paper.

8 Conclusion

In this study, we have evaluated the MLF model of CS with some Igbo-English data and have concluded that the data can indeed be considered a classic case of CS in that a ML can be clearly identified in bilingual clauses. We have established this by a qualitative and quantitative analysis, uncovering overwhelming supportive evidence for the MLP, the AP and the USP. The findings also confirm that CS is not blocked when the surface structures of two languages do not map onto each other; additionally, CS is possible DP-internally between a functional head and its complement.

The overall implication for the MLF model is that its predictive power lies in its recognition that there will be an asymmetry between the ML and the EL in their roles in setting

the morphosyntactic frame of the bilingual clause. The regularity with which Igbo supplies both the frame building elements (system morphemes) and sets morpheme order wherever there is a conflict in word order in Igbo-English CS bears this out.

Nevertheless, further research on a much larger corpus might reveal more problematic examples than we have identified in the present study. Also, an aspect of Igbo-English intrasentential CS that this study has not touched concerns uncovering what motivates the speakers to codeswitch in the first place. A number of reasons have been adduced in the literature as motivating factors for codeswitching. These range from social-pragmatic to grammatical considerations. Future research on Igbo-English CS will seek to ascertain the particular motivations why Igbo-English bilinguals engage in codeswitching.

Abbreviations

AUX	auxiliary	INF	infinitive
BE	copular verb	NEG	negation
CL	clitic	PERF	perfective
C	complementiser	PL	plural
D	determiner	PREP	preposition
DEM	demonstrative	PRN	pronoun
FUT	future	Q	quantifier
IND	affirmative indicative		

References

Abney, Stephen. 1987. *The English noun phrase in its sentential aspect.* Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation.)

Belazi, Heidi, Edward Rubin & Almeida Jacqueline Toribio. 1994. Code-switching and X-bar theory: The functional head constraint. *Linguistic Inquiry* 25(2). 221–237.

Bhatt, Mohan. 2001. Codeswitching: Structural models. In Rajend Mesthrie (ed.), *Concise encyclopedia of sociolinguistics,* 456–461. Amsterdam: Elsevier Science.

Bolonyai, Agness. 2005. English verbs in Hungarian-English codeswitching. In James Cohen, T. Kara McAlister, Kellie Rolstad & Jeff MacSwan (eds.), *Proceedings of the 4th International Symposium on Bilingualism*, 317–327. Somerville, MA: Cascadilla Press.

Clyne, Michael. 2003. *Dynamics of language contact: English and immigrant languages.* Cambridge: Cambridge University Press.

Davies, Peredur & Margaret Deuchar. 2010. Using the Matrix Language Frame model to measure the extent of word order convergence in Welsh-English bilingual speech. In Anne Breitbarth, Christopher Lucas, Sheila Watts & David Willis (eds.), *Continuity and change in grammar,* 77–96. Philadelphia: John Benjamins.

Di Sciullo, Anne-Marie, Pieter Muysken & Rajendra Singh. 1986. Code-mixing and government. *Journal of Linguistics* 22. 1–24.

Echeruo, Michael J. C. 1998. *Igbo-English dictionary.* New Haven: Yale University Press.

Emananjo, Nolue. 1978. *Elements of Modern Igbo grammar.* Ibadan: Oxford University Press.

Essizewa, Komlan. 2007. Language contact phenomena in Togo: A case study of Kabiye-Ewe codeswitching. In Doris L. Payne & Jaime Peña (eds.), *Selected Proceedings of the 37th Annual Conference on African Linguistics,* 30–42. Somerville, MA: Cascadilla Proceedings Project.

Fukui, Naoki. 2001. Phrase structure. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 374–406. Oxford: Blackwell.

Jake, Janice, Carol Myers-Scotton & Steven Gross. 2002. Making a minimalist approach to codeswitching work: Adding the matrix language. *Bilingualism: Language and Cognition* 5(1). 69–91.

Kayne, Richard. 1994. *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

MacSwan, Jeff. 1999. *A Minimalist approach to intrasentential codeswitching.* New York: Garland Publishing.

MacSwan, Jeff. 2009. Generative approaches to codeswitching. In Barbara Bullock & Almeida J. Toribio (eds.), *Cambridge handbook of linguistic codeswitching*, 309–335. Cambridge: Cambridge University Press.

Maduka-Durunze, Omen. 1990. Igbo adjectives as morphologized relatives. *Studies in African Linguistics* 21(2). 237–250.

Myers-Scotton, Carol. 1993. *Duelling languages: Grammatical structure in codeswitching.* Oxford: Clarendon Press. (Paperback 1997).

Myers-Scotton, Carol. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes.* Oxford: Oxford University Press.

Myers-Scotton, Carol. 2006. *Multiple voices: An introduction to bilingualism.* Malden, MA: Blackwell Publishers.

Myers-Scotton, Carol. 2013. Cost makes a difference: Codeswitching frequencies in bilingual language production. In Kelechukwu Ihemere (ed.), *Language contact: A multidimensional perspective,* 29–51. Newcastle upon Tyne: Cambridge Scholars Publishing.

Myers-Scotton, Carol & Janice Jake. 2001. Explaining aspects of codeswitching and their implications. In Janet Nicol (ed.), *One mind, two languages: Bilingual language processing*, 84–116. Oxford: Blackwell.

Myers-Scotton, Carol & Janice Jake. 2014. Nonfinite verbs and negotiating bilingualism in codeswitching: Implications for a language production model. *Bilingualism: Language and Cognition* 17(3). 511–525.

Obiamalu, Greg. 2013. The functional category D in a language without determiners: The case of Igbo. *Journal of Theoretical Linguistics* 10(2). 50–67.

Poplack, Shana & Marjory Meechan. 1998. How languages fit together in codeswitching. *International Journal of Bilingualism* 2(2). 127–38.

Radford, Andrew. 2004. *Minimalist syntax: Exploring the structure of English.* Cambridge: Cambridge University Press.

Treffers-Daller, Jeanine. 1991. Towards a uniform approach to codeswitching and borrowing. In ESF Network on Codeswitching & Language Contact (eds.), *Papers for the workshop on constraints, conditions and models*, 259–279. Strasbourg: European Science Foundation.

Uwalaka, Mary Angela. 1997. *Igbo grammar*. Ibadan: The Pen Services.

Chapter 28

Humor in Kenyan comedy

Martha Michieka
East Tennessee State University

Leonard Muaka
Winston Salem State University, North Carolina

In multiethnic and multilingual communities of Africa, speakers claim certain ethnic affil-
iations through their speech. As an identity marker, language further compartmentalizes
speakers into certain groups and leads to attitudes, labeling, stereotyping and perceptions
that go beyond language itself. Based on data drawn from Kenyan media discourse, this pa-
per focuses on Kenyan ethnic humor to analyze absurdities, incongruity, and stylized speech
by Kenyan humorists and their role in identity marking, boundary construction, performa-
tivity and personal experiences. The paper combines several approaches to examine how
contemporary Kenyan humorists such as Eric Omondi and Byron Otieno style their Swahili
and English discourses to index Kenyanness and Kenya's diverse ethnolinguistic identities.
The paper examines how this ethnic humor is perceived and what comedians do to ensure
that their humor is not offensive to implied ethnic groups. The use of ethnic flavored dis-
course styles that strategically include and exclude different participants brings humor to
the audience and helps downplay the stigma associated with certain styles of speech, ethnic
practices, and marginalized ethnic groups. Given how ethnicity is intertwined with lan-
guage, this paper argues that the use of incongruity, absurdities, and stigmatized speech
forms by comedians in their jokes helps create humor and neutralizes the ethnic tensions
that exist amongst different ethnic groups.

1 Introduction: An overview of Kenyan humor

Humor is understudied not only in Kenya, but in Africa in general. However, with a
new generation of humorists driven by liberalized media, technology, and globalization,
African humor is beginning to draw attention from researchers. In Kenya, comedy has
been made possible due to several factors, but more importantly due to the political and
socioeconomic changes in the last two decades. This paper seeks to highlight how the
aforementioned changes have impacted Kenyan comedy and uniquely positioned it as
Kenyan humor.

In the following subsections, a short overview of humor in the Kenyan context is offered.

1.1 Kenyanness and Kenyan humor

What is considered "Kenyan" and consequently what is Kenyan humor? The Kenyan community consists of over 42 million people belonging to more than 42 different ethnic groups. Generally, to be Kenyan is to belong to, or to identify with at least one of these ethnic groups. Most Kenyans are multilingual and can speak at least one ethnic language and either one or both of the Kenyan official languages (English and Swahili). When Kenyans of the same ethnic group meet, they are immediately connected through the shared language and cultural information. Most Kenyans can identify members of their ethnic groups through their easily identifiable accents as well as their shared names. Names such as Kamau, Wanjiku, Wamalwa, Kiprono, and Mogaka typically belong to particular Kenyan ethnic groups, and Kenyans, whether at home or in the diaspora, identify with those names because they indicate Kenyanness. There are also other unique ethnic group characteristics such as the choice of foods, lifestyles, and common economic, religious, and traditional activities that specific groups generally engage in. The shared knowledge Kenyans have about various ethnic groups and each group's unique characteristics make the idea of "Kenyanness" stand out.

Although ethnicity is clearly a mark of identity in Kenya, at the larger Kenyan community level, Kenyans have shared practices that extend across ethnic subgroups as well, that distinguish them from other African nationalities. This is the uniqueness that Tumusiime (2013: 5) says Kenyan humorists tap into in order to indigenize their humor and make a career. Most Kenyans, for instance, are proficient in the Kenyan varieties of English and Swahili – varieties which are distinct from those spoken in neighboring countries. These language varieties unify Kenyans, as well as distinguish them from speakers of English and Swahili in neighboring countries. Unless one has been to Kenya and learned the Kenyan ways, he/she may not, for example, understand the typical Kenyan *kitu kidogo* 'something small', an expression that every Kenyan understands to mean a bribe. Apart from the unifying lingua franca(s), Kenyans also share traditions, practices and foods such as chapati, ugali, and nyama choma. These cross-Kenyan shared characteristics are also what Kenyan humor draws from and which, when performed, resonate with Kenyan audiences. As Chiaro (1992) points out, the appreciation of jokes requires, among other things, shared cultural backgrounds.

1.2 Locating Kenyan humor

Humor is a universal phenomenon and it may even address universal themes and functions (Norrick & Chiaro 2009: xiii). However, certain humor trends unique to each society are influenced by sociopolitical and historical factors. Before being shared out in public, humorous acts can be said to have occurred within in-group members the same way riddles and stories are told. In the Kenyan case, humor has come of age and gone public largely due to a liberalized media sector that has broadened the platform of tapping into

the local talent market. During the one-party political system period in Kenya, media waves were largely state controlled and therefore only a few comedians could display their potential and talent. However, following the introduction of multiparty politics that reduced media censorship and the emergence of entrepreneurship among young people, many media houses emerged and ended up creating opportunities for upcoming and talented comedians. Without a doubt, globalization had exposed many Kenyans to foreign humor that was no longer "magical" and many wanted authentic Kenyan humor. This mindset has inspired local artists, who are adored by Kenyans happy to embrace Kenyan pride and culture both locally and abroad. Consequently this trend has given local artists a ready market that allows them to reflect on their own society and tell their own stories.

In essence, Kenyan humor appeals to many Kenyans because of the shared values between the humorists and their audience and because the humor mirrors the lives of the audience. It finds its true meaning and identity in the multifaceted linguistic and cultural landscape of the Kenyan people. Kenyan accents, styles, cultural practices, and other practices, lead to identity marking that not only in-group members acknowledge, but also allows out-group members to formulate stereotypes about them. Kenya's in-group identity is dynamic and available at different levels and among different groups. Therefore, although humor in Africa shares certain similarities, Kenyan humor is unique to its people's ways of life.

As a genre, the continued growth of Kenyan humor owes its present state to some of the early artists of the 1980s such as Amka Twende (Benjamin Otieno), Othorong'ong'o (Joseph Anyona), Mutiso (Kimunyo Mbuthia), Mzee Ojwang Hatari (Benson Wanjau), Mama Kayai (Mary Khavere), Baba Zero, Masanduku arap Simiti (Sammy Muya), Wariahe (Said Mohammed Said) and Otoyo Obambla (Samuel Mwangi) who featured prominently on *Vitimbi*, a local Swahili comedy show (*Passion For Life Magazine* July 2011). Although most of these comedians performed in Swahili, contemporary comedians tend to code switch between English and Swahili or utilize the urban variety of Sheng. Aside from television humorists, the Kenyan audiences were also entertained through print humor by artists such as Wahome Mutahi (Whispers), James Tumusiime (Bogi Benda), Cheka na Baraza,[1] and Juha Kalulu (Edward Gicheri Gitau), the last two being in Swahili. These artists were particularly humorous because many Kenyans could identify with them. Mzee Ojwang, for example, though not Luo by ethnicity, performed on stage by stylizing his Swahili to project a Swahili speaker of Luo ethnicity entertaining all Kenyans by his parody of the Luo community.

1.3 The state of Kenyan humor

In the 1990s a trio of Kenyatta University students emerged on the entertainment scene impersonating President Moi and his cabinet to entertain Kenyans at a time when ridiculing the president was almost unheard of. The group known as Redykyulass was composed of Walter Mongare, Tony Njuguna, and John Kiarie. Political changes on the conti-

---

[1] See Rhoades (1977).

nent at the time favored multi-party politics and freedom of speech, consequently allowing political satire into the Kenyan media. Kenyans embraced this new way of entertainment, ushering in a new generation of comedians such as Churchill (Daniel Ndambuki), Eric Omondi and Byron Otieno that are now known regionally and internationally.

Ndambuki's entry onto the Kenyan comedy scene in the mid-1990s brought a kind of comedy that artists and the society took seriously, and it has since given rise to young talent that draws heavily on Kenya's diverse ethnic groups. Ndambuki has nurtured many aspiring artists seeking to follow in the footsteps of the aforementioned pioneering comedians. Beyond the Kenyan scene and owing to globalization and regional collaboration, Kenyan humorists partner with their contemporaries in other African countries to entertain and build what can be referred to as African standup comedy. Kenyan comedians thrive on the diversity that exists in the country and as stated in §1.1 and §1.2, the themes that they focus on revolve around ethnolinguistic, cultural, socioeconomic, and political aspects of the society. While they may not come from the ethnic community they project, these comedians are able to capitalize on shared knowledge in their performative acts. As Ndambuki notes in an interview, Kenya's middle class has converged toward Kenyan material as Kenyan artists seek to tell their own story and celebrate their different cultures, devoid of foreign material (Kimani 2014). This is what has propelled Kenya's showbiz, unlike in the past when Kenyans bypassed their local culture and entertainment for something new and foreign.

2 Conceptual framework

This study focuses on Kenyan humor by analyzing absurdities, performativity, incongruity, and stylized speech in Kenyan humor and the role these elements play in identity marking and boundary construction. However, before we analyze Kenyan ethnic humor and how it is perceived, it is important to define two key terms, namely, *ethnicity* and *ethnic humor* as used in this paper.

2.1 Ethnicity and ethnic humor

Humor is a part of who we are as human beings, and although not everyone may be considered to be humorous, we all tend to understand what humor is in our culture. How do comedians ensure that their humor is socially acceptable, especially when it is ethnic centered? Davies (1990: 1) notes that "the term ethnic tends to be used in a broad way about a group that sees itself and is seen by others as a 'people' with a common cultural tradition, a real or imagined common descent and a distinctive identity." A group has to recognize in its members several factors that distinguish it from the rest of the communities around it. These factors should also be clear enough that outsiders notice the difference(s) and classify those who subscribe to this group as different. Davies (1990: 1) further states that "an ethnic identity may be chosen or changed either by individuals or by the collective decision of a group, whether suddenly and deliberately or as a gradual result of other decisions [...] and the members of an ethnic group may collectively choose

to assert or neglect the factors that set them apart from others". Ethnic humor, on the other hand, has been defined as a "type of humor in which fun is made of the perceived behavior, customs, personality, or any other traits of a group or its members by virtue of their specific sociocultural identity" (Apte 1985: 108). The humorist can be either an outsider expressing superiority or an insider poking fun at self.

This paper seeks to analyze Kenyan humor and establish what makes ethnic humor unique and funny. The term *ethnic* will be used in reference to the slightly over 42 different ethnic groups mentioned above. Prior to addressing these questions, a cursory overview of the frameworks used in this paper is provided.

2.2 Theories of humor

According to Raskin (1985), there are three major theories useful in the analysis and understanding of humor: superiority, incongruity, and release. In our analysis we apply these three theories as well as focus on how Kenyan comedians achieve humor without being offensive in a context that could otherwise be ethnically sensitive. Along with these concepts, we also take note of concepts such as performativity, stylization, and identity.

2.2.1 Superiority theory

Proponents of the superiority theory view humor as an avenue through which one group pokes fun at a marginalized group and by so doing, laughs at the group's folly, perhaps because it is not their own. Carroll (2014: 8) notes that for superiority theorists such as Hobbes, people's laughter results from perceiving infirmities in others which [then] reinforce our own sense of superiority. Such kind of humor tends to portray the targeted groups as "stupid, ignorant or unclean" (Apte 1985: 115). However, superiority does not necessarily entail enmity. We laugh because we experience the joy of getting at the enemy and defeating them in some way. Davies (1990: 7) maintains that "a further factor that underlies much humor and is brought out particularly well in ethnic jokes is the sense of sudden vicarious superiority felt by those who devise, tell or share a joke. Ethnic jokes "export" a particular unwanted trait to some other group and we laugh at their folly, perhaps glad or relieved that it is not our own." Whether listeners laugh at such humor or find it offensive will depend on their level of group identity. As Apte (1987: 34) claims, "if ethnic humor is initiated in a small group interaction where a listener's ethnic group is the target, the listener may choose to ignore it or to affirm his or her identity and to protest, thereby either embarrassing the teller of the joke or putting that person on the defensive." Labrador's (2004) work shows that when humor is directed at others, it can be offensive. He explores humor in Hawaii where local jokes poke fun at "the Manong" (recently arrived Filipino immigrants), who are ridiculed not just for their language and perceived heavy Filipino accent but also for the foods they eat and how they socialize in their homes. As Apte (1987: 35) notes, fury against negative humor is not unusual. For instance, "Politicians and other public figures who either intentionally

or inadvertently engage in ridicule at the expense of some ethnic group have suffered serious consequences for their action."

However, humor, even the superiority type, can serve various positive functions such as that of correcting social behavior. Ziv (1988: x) points out that we can achieve this social function by laughing "at forms of behavior or thought that are contrary to what is socially expected and accepted. Therefore laughter can have a punitive effect aimed at correcting behavior." In this article, the superiority approach allows for an analysis of Kenyan humor by seeking to show how some ethnic groups may laugh at the folly of a disempowered group. The power relations involved are not just political but they may be linguistic, social, or cultural.

2.2.2 Incongruity theory

Unlike superiority theory, incongruity theory attributes the source of amusement to the ambiguity or the unexpected ways of looking at a situation. Proponents of this approach believe there cannot be humor without incongruity. Berk (1999) and Carroll (2014), for example, view incongruity as the most basic structure of humor consisting of at least two elements: the "expected content" and the "unexpected content" (see Berk 1999: 7). When the unexpected is what the audience gets, humor becomes funnier. Therefore humor arises from the perception of an incongruity between a set of expectations and what is actually perceived. Attardo & Raskin (1991: 308) argue that there is usually a punchline that "triggers the switch from one script to the other by making the hearer backtrack and realize that a different interpretation of the joke was possible from the very beginning." In Kenyan humor, the unexpected twists in the selected acts are what make the jokes hilarious.

2.2.3 Release theory

Release theory posits that there is a psychic release as a result of humor. As a type of self-deprecating humor, artists may present jokes that target their own follies or defects, as part of a defense or tension-relieving mechanism that is not necessarily intended to be hostile or aggressive (Davies 1990; Gulas 2006). According to Apte (1987: 30), "engaging in mild self-deprecatory humor is generally perceived as clear indication of having a sense of humor." People who are able to laugh at themselves are perceived as having a good sense of humor and being easygoing. Ziv (1988: x) defines self-disparaging humor as those "instances in which we ourselves are 'victims' of the joke." Similarly Berk (1999: 12) observes that "self-effacing or self-deprecating humor in the form of *self-downs* is not only an acceptable form but a highly desirable one to break barriers." Some minority groups may even find strength in self-disparaging humor as that humor helps them rethink the values that bind them together as a community. Gonzales & Wiseman (2005: 172) observe that ethnic humor directed at self is useful among minority groups since these small and marginalized groups will use this humor as "a means of trying to preserve their ethnic identity and tradition." In a similar manner, Labrador (2004: 298) argues that "local comedy and ethnic jokes are important sites for the practice and

performance of local identity and culture." This type of humor helps to construct identity among the affected people (Labrador 2004; Rappoport 2005); and as Apte (1987: 36) observes, comedians who make fun of their ethnic groups are quite popular.

In this article, these approaches are applied in the discussion of Kenyan humor. Each act is analyzed using an approach or a combination of approaches that best captures the intended humor. We also examine how comedians avoid being offensive. This article demonstrates that jokes occur in context and their effectiveness depends largely on shared knowledge and stereotypes. Therefore, the appreciation of ethnic humor depends on how much we know about and identify with the joke and its target and the shared background knowledge.

3 Data and data analysis

Data used in this paper were gathered from Kenyan standup comedy clips from YouTube. These episodes were recorded in Kenya, Rwanda and Zambia. The main show, however, is recorded and produced in Kenya on the Churchill Show at Carnivore Restaurant. The Churchill Show has several videos that have been collected over time and for our analysis we chose over 50 online video clips from some of the most recent episodes. Since we cannot include audio files in this paper, the clips that are relevant to the current analysis are transcribed to help the reader understand the context as well as the analysis.

3.1 Superiority at the local and international level

Ethnicity as defined above can be viewed within the larger Kenyan society as well as at the international level. In the Kenyan context, comedians entertain their audiences using jokes that poke fun at other Kenyan ethnic groups. At the international level, local ethnic differences are overlooked and the focus shifts to the idea of Kenyanness. Comedians' identities continue to evolve depending on the audience as well as the context.

In the joke below, Majimoto, a humorist from the coastal region of Kenya, privileges his coastal people over the Meru people due the coastal people's higher proficiency in Swahili. He imitates the way the Meru people speak Swahili to make his joke successful. What is also interesting is that he delivers this joke to an audience dominated by Meru speakers.

(1) Swahili spoken with Meru accent

Wajua mi natatizika sana nyie watu wa bara mwatatizika sana na Kiswahili. Halafu si ndio tumebarikiwa twajua Kiswahili sana. Ndio maana wale mapresenta wa ripoti za saa moja wengi ni wa Mombasa. Sasa wajua nini, kuna ripota mmoja ni rafiki wangu sana. Sasa tulikuja naye hapa town moja sijui inaitwa nini – akaripoti hapo. Anaanza tu kuripoti: "Tuko katika mji huu wa Chuka, kaunti ya Meru ambapo asubuhi ya leo katika bucha ya Bwana Karimi kumetokea wizi wa mabavu." Wameru wakamkatisha: "Mbro tangaza ukweli. Hao watu hawajaiba mbavu, wameiba steak." [Laughter]

> 'You know I am always troubled by the way you people from upcountry struggle with Swahili. And then for us we are blessed to know Swahili very well. That is why most 7:00 pm news reporters are from Mombasa. Now you know, there is one reporter who is a great friend of mine. Now I accompanied him to a town nearby whose name I can't recall– and he was reporting. He begins to report: "We are in Chuka town, of Meru County where this morning at Mr. Karimi's Butchery, there has been robbery with violence." Meru people cut him short: "Brother, report the truth. Those people did not steal ribs, they stole steak."' (Churchill Show 2015a)

Here, Majimoto pokes fun at the pronunciation and limited Swahili proficiency of Meru people. The phrase *wizi wa mabavu* which means 'violent robbery' is misunderstood by the Meru people to mean *mbavu* 'ribs' and hence the rest of the joke. To the Meru people, who are known for their nasalized speech, *Mabavu* and *mbavu* are homophonous. It may be argued that the humorist is bragging about the coastal people's high proficiency level in Swahili. Most of the coastal people speak Swahili natively while other Kenyans acquire Swahili as a second language. In this context, the superiority humor is due to the failure by the Meru speakers to decode a Swahili idiom, as native speakers would do. Majimoto's joke therefore expresses the idea of one ethnicity being superior to another linguistically. Although Kenyans are multilingual, their levels of proficiency vary and they do not necessary seek to acquire the coastal Swahili variety.

Another humorist, Eric Omondi of Luo ethnicity capitalizes on shared local and international knowledge and stereotypes. In some cases his jokes concern competing ethnic communities such as the Luo and the Kikuyu, the most dominant ethnic groups in Kenya politically and, to some extent, economically. In the following joke, the humorist portrays Kikuyus as entrepreneurs whereas the Luos are portrayed as proud, snobbish, and arrogant people. Viewed through the lens of the superiority framework, this humorist privileges one group over another.

(2) English

The Luos and Kikuyus [...] (are) very different, very unique, almost the same but very different. Like the Kikuyus, Kikuyu is the national language, and Luo is the international. [laughter] For Luos, money is not a problem and for Kikuyus, money is not everything; it is the only thing... [laughter]. (Churchill Live 2011)

As Apte (1985) points out, there is nothing new that the comedians bring to the audience but rather it is how they repackage what is already known in form of jokes to entertain their audience that triggers excitement and laugher. Omondi portrays the Luo as a superior ethnic community. He introduces the two ethnic groups as being "almost the same but very different," signaling the groups' political dominance. But he also succeeds at privileging his Luo group over Kikuyus. If Kikuyus dominate at the national level, then the Luos are superior by dominating at the international level. Omondi further jokes that money is not a problem to the Luos; however to Kikuyus it is the only thing, thus satirizing the Kikuyus' obsession with money. The wordplay and the unexpected twists further make the joke more hilarious.

In yet another joke Omondi connects with his audience by portraying Luos as superior and arrogantly so, living lavishly and driving the latest models of cars; while the Kikuyus are misers, saving every penny and driving old outdated cars. The humor comes not only from shared stereotypes about the lifestyles and boastful nature of the Luo people, but also from the humorist's stylization of this conversation.

(3)   Announcer speaks in English; Caller speaks in English with an unmistakably Luo accent

You know I used to hear people say Luo is a lifestyle I never understood what it meant until two weeks ago a certain Luo man made a phone call to a certain radio station.
Caller: Hello, we are on our way to Thika. We are going to support Gor Mahia.
Announcer: OK sir, how many are you?
Caller: Do not ask me how many we are; ask how many cars have we. [laughter]
Announcer: OK sir, I am sorry. How many cars do you have?
Caller: Do not ask me how many cars have we, ask me which models? [laughter]
Announcer: OK I am really sorry, which model sir?
Caller: We are having 30 Mercedes and 4 NZs. That means we have 30 Luos and 4 Kikuyus. (Churchill Live 2011)

At the international level, the emerging African film industry, such as Nollywood in Nigeria and Riverwood in Kenya, gives Omondi a subject of entertainment. Nollywood films saturate the Kenyan market and although there is rivalry, artists from both West and East Africa have been collaborating in different entertainment ventures. When Omondi compares Kenyans and Nigerians he rises above the ethnic boundaries and views himself as a Kenyan, privileging Kenyanness over Nigerianess. At this level the projection is elevated to that of a Kenyan irrespective of the local differences that exist among certain Kenyan ethnic communities. He pokes fun at Nigerian etiquette and mannerisms. As Labrador (2004) notes, we poke fun at other people's follies, thankful that we are not in their position. Omondi's jokes privilege the groups that he belongs to and therefore it becomes "we" versus "them." In example (4), Omondi constructs complex ethnic boundaries that portray a positive Kenyan identity compared to the Nigerian one that is vain.

(4)   English

You know Nigerians, they talk big but when it comes to actions, there is a problem. Like I saw a Nigerian who went to a hotel and he was asking for everything, "Waiter do you have fried chicken? Ok do you have pizza, pizza, big large pizza? Ok wait a minute, do you have milkshake? OK, do you have the big hamburger [hombaga]? Ok if you have all those, in that case, give me a cup of tea." (Omogi 2012)

Additionally, in example (5) Omondi praises Kenyan women for their politeness and good manners as contrasted to Nigerian women. This performance can be viewed from both superiority and incongruity frameworks.

(5)   English, with Nigerian portrayed as speaking in a heavy Nigerian-English accent

Kenyan women are so kind and polite and patient […] Even Safaricom is one of our sponsors. They chose a Kenyan lady to be talking to us on the phone. When you call Safaricom, there is a lady who is very polite and very patient. She tells you the mobile subscriber cannot be reached. And you can call again and she tells you the mobile subscriber cannot be […] [The audience fills in.] And you can call a million times and she will be so patient, she will keep telling you, the mobile subscriber cannot be reached. She will even go into Swahili and say, "Samahani, mteja wa nambari uliyopiga hapatikani." [Audience echoes.] Sorry, the mobile subscriber cannot be reached.
Go to Nigeria; the correspondence in Nigeria, the lady from Nigeria, she tells you once! "The mobile subscriber cannot be found!" [In a heavy Nigerian English accent]. [Laughter].
Try calling again. "Hello, didn't I just tell you the mobile subscriber cannot be found?"
[Laughter] "What is your problem, stupid man?" [In a heavy Nigerian English accent]. (Omogi 2012)

Another humorist who compares ethnicities across national borders is Pablo from Uganda. He entertains his audience by portraying Kenyans' inability to give directions (6). However, because of his desire not to antagonize the Kenyan audience, he resorts to inclusivity as seen in clip (7), and softens the joke showing that Ugandans are not any different.

(6)   English

We all don't know how to give directions. I met a young boy. He was around 9 years old and I asked him, "Where is the museum?" And the boy looked at me and told me "I don't know." And so as I was walking back the boy called me in the most local way (whistles) and I knew that was meant for me because I am a VIP (Very Improved Peasant). So I come back thinking that maybe the boy has remembered where the museum is, instead he showed me his friend and told me "This my friend, also he don't know. (Omogi 2012)

(7)   English

But it is not only in Kenya. It also happens in Uganda. You ask for directions at your own risk. (Omogi 2012)

So while the first part of the joke was humorous due to its incongruity, there is a possibility of it being interpreted as offensive for portraying a sense of superiority by an outsider. However, the artist quickly returns to self-deprecation and places Kenyans and Ugandans on the same level. Unlike Omondi who does not make such apologies in his talk about Nigeria, maybe because Kenya and Nigeria are not neighbors and therefore offending Nigerians will not have any serious consequences, Kenya and Uganda are neighbors and offending one might have serious repercussions. The context favors Omondi, unlike Pablo whose act is in front of Kenyan audience.

3.2 Incongruity

Incongruity theory attributes humor to an unexpected twist. In this section we show examples of how comedians employ incongruity to create humor and entertain their audience. In example (8), MC Jessy ridicules televangelists who lie about receiving special messages from God.

(8) Swahili spoken in a Kikuyu accent

Ningependa uangalie kwenye television, Mimi nimekuwa mchungaji kwa miaka twenty […] na Mungu amenionyesha pande hii ya reft kuna Mzee mmoja ako na miaka ninety five. Amekuwa akililia Mwenyezi Mungu kwamba haoni. Macho yake anafungua lakini anaona giza. Mungu ameniambia nikuulize. Mzee miaka ninety five, umeishi ninety five years, ni nini ingine unataka kuona.

'I would like you to look at your TV screen [addressing viewers], I have been a pastor for twenty years […] the Lord has shown me in this left side, there is a ninety five year old man who has been crying to the Lord that he can't see. He opens his eyes but all he sees is darkness. The Lord has told me to ask you [now directing his message to the old man], "Old man, you have lived for ninety five years, what else do you need to see?"' (Churchill Raw 2014)

Usually the audience will expect the preacher to bring some good news to the worshippers and so this joke takes a humorous twist when the preacher tells the old man that he has lived long enough and has seen everything there is to see and therefore he should not be bothering God anymore by asking for better eyesight. If the comedian didn't add this unexpected turn or if he told the old man that God had promised sight, nothing could have been funny.

In the joke below, Smart Joker and his friend parody robbers; and as robbers, they have a "constitution" to regulate their robbery. Their first rule states

(9) Swahili in a Luhya accent

Rule namba one kwa constitution ilisemanga hivi: Hakuna kuipia mama mja mzito […] tumuachie nafasi azae mtoto apate laptop.

'Rule number one on the constitution states that: There should be no robbing of a pregnant woman […] we should give her time to deliver so that the child can get a laptop.' (Churchill Show 2013)

Smart Joker entertains his audience by his strong western Kenya Luhya regional accent (Luhya) that clearly portrays where he comes from. However, what makes the above joke funny is the twist in especially the second part. Smart Joker sounds very human when he states that they should not steal from a pregnant woman and the audience might just conclude that even robbers are understanding enough not to bother pregnant women. In the video, the audience does not laugh at the joke until Smart Joker adds the twist or the unexpected ending. It is then that the audience realizes they were wrong in thinking the robbers were being kind as most people generally are to pregnant women.

Apparently the most valued item for the robbers is a laptop and they are so desperate for laptops that they are willing to wait until a child is born and grows to be of an age that he/she can be given a laptop at school which the robbers can then steal. It is shared knowledge among Kenyans that when the current Kenyan government was campaigning, they promised to give school-going children laptops. (Unfortunately the project did not take off.)

3.3 Release or self-deprecation

As mentioned earlier, release or self-deprecating humor involves jokes that target one's own follies. The examples in this section involve artists who expose some of the follies of their own ethnic groups and, because they belong to those groups, the jokes are considered self-deprecating.

In the following example, another comedian Owago-Onyiro, laughs at his own "Luoness."

(10) Swahili and English, with some Kikuyu

   a. Ujaluo utakuua.
       '"Luoness" will kill you.'

   b. Lazima upende mahali unatoka.
       'You have to appreciate your origins.'

   c. People from Nyanza like prefixes. A prefix is a title that comes before the name: Dr Geofrey Otieno, Engineer Obado, Lecturer Omondi. People from Central also like something called suffixes. Suffix is a title that comes after the name:

   d. Mwangi wa Equity, Wanjiku wa makara.
       'Mwangi of Equity' [equity bank], 'Wanjiku of Makara' ['charcoal' in Kikuyu, or 'Wanjiku who runs the charcoal business'] (Churchill Show 2013)

This comedian addresses several issues and, since he is a Luo, he can poke fun at the Luo lifestyle without worrying about the consequences. He is clearly making fun of the Luos for their arrogance and their love of lengthy titles. However, since he is Luo, he has the right to say whatever he wants about himself and his people. This can be likened to the situation in the USA where it is OK for an African American to use the N-word but it would be considered inappropriate if someone from a different race used the same term. It is also clear that through Onyiro's jokes, the Kikuyu ethnic group is associated more with business ventures, unlike the Luo's lifestyle.

In this next joke, Omwami is supposedly Luhya as indicated by his name and he jokes freely about the Luhya people.

(11) Swahili with Luhya features

Unajua Mluhya ni mtu anakutusi lakini vile amepackage hiyo matusi uwezi jua amekutusi. Mluhya anaweza sema, ah, "Bosi yaani umefika, asante sana kwa kufika show [...]" [stretching out his hand to give a handshake] "kamatako hii."

'You know a Luhya can curse you but the way the curse is "packaged" you won't realize you have been cursed. A Luhya can say to you, "Boss, so you made it here, thank you for coming to the show [...]" [high five]' (Churchill Raw 2014)

The wordplay here is the twist that delivers the humor to an audience that shares in the language stereotypes. There is a tendency for Luhya dialects to add a particle or suffix like *ko* and *nga* to a conjugated verb. When *ko* is added to the verb *kamata* 'grasp', the new word created could be a dimunitive for 'butt'. Consequently in Omwami's joke, the high five could also mean 'you little butt [...]'

Kenyan audiences laugh at the Luhya speech, but at the same time they realize that such jokes empower the Luhya speakers. If the speaker chooses to, s/he can use the language to attack his boss; but since the listeners know that Luhyas often add that suffix to their words, nobody can punish the speaker. The artist being Luhya himself has the right to make fun of his own people without offending them.

In the next example, Majimoto this time pokes fun at his community that is often considered less educated and limited in its knowledge of English. As a coastal resident, he pokes fun at his own Mombasa people because he is one of them. This will not be an easy joke for an outsider to make considering that the people from this area are limited in their English proficiency due to marginalization and unfair allocation of public resources. Moreover, this is supposedly a joke about his own mother. What the audience finds funny is the misinterpretation of the phrase *high class* by the mother. *High class* is used here to mean 'sophisticated /classy,' but the mother understands it to mean an upper grade in school.

(12) Swahili with English

Ikabidi nimchukue bwana mtoto nikampeleke nyumbani. Sasa mtoto wa Kikamba ikabidi nikamintroduce kwa mother, unaona. "Mama waona bwana nakuletea mkaza mwana – mkaza mwana sasa ni bibi ya mtoto wako. Sasa nakuletea mkaza mwana mtoto wa kikamba, mtoto high class. Mama nakuletea mtoto wa high class." Mama aliposikia mtoto wa "high class," mama akaanza kunisomea: "utafungwa wewe, utafungwa wewe yaani mtoto bado yuasoma yuko high class wewe wamchukua wamleta nyumbani wataka kumfanya bibi".

'It was necessary I take her home. Now, the Kamba girl, I had to introduce her to my mother, you see. "Mother you see this is your daughter-in-law – a daughter-in-law is a wife of your son/child. Now I am bringing to you a daughter-in-law, a Kamba girl, a high-class girl. Mother, I bring you a high-class girl." When mom heard "a high class girl" she started lecturing me, "You will be jailed, you will be jailed; the girl is still in school in 'high class' and you are taking her and bringing her home to be your wife." (Churchill Show 2015b)

3.4 How do humorists avoid being offensive?

In our analysis we notice a number of tactics the comedians employ to avoid being offensive especially to other ethnic groups. Some of the common tactics they use include self-effacing, inclusivity, outright apology, and careful explanation of their joke.

3.4.1 Self effacing

As we have seen in the examples above, most artists choose to poke fun at themselves or at their own ethnic groups instead of laughing at other groups. Luo comedians such as Omondi and Owago make fun of Luo practices, while Luhya comedians like Omwami freely joke about their Luhya follies. Instead of outsiders poking fun at the Luhya or Luo accents in spoken English, the comedians from those groups take it upon themselves to bring this out in the jokes. Owago, for example, discusses the issue of "shrubbing" among Luos. *Shrubbing* (or *srub*) is a Sheng word which refers to the unintentional [ʃ]/[s] substitution characteristic of many Luo speakers of English or Swahili as second languages.

(13)   English with Sheng code-mixing (in bold)

You say haa you Luos you **srub**. **Mimi ninasrub**? Be careful if you have a friend from Nyanza. Luos only **srub** what they don't like. **Kama sida** [laughter] [...]**mandasi** [...] **shitikashel** [more laughter] [...] **aishishi**. **We maisani mwako ushawai ona mjaluo akisema Mashidishi ama Range Rover Shport**?
'You say that haa you Luos you shrub. Do I shrub? Be careful if you have a friend from Nyanza. Luos only shrub what they don't like. Things like problems (standard pronunciation [ʃida]) [...] difficulties [...] City Council (standard [sitikansel]) [...] ICC (standard [aisisi]). In your entire life, have you ever heard a Luo say "Mashdishi" (< Mercedes) or "Range Rover Shport"? (< Sport)"'
(Churchill Show 2013)

Here Owago, a Luo comedian, pokes fun at himself and his fellow Luos by performing and stylizing an embarrassing language aspect that many Luos struggle with. However, he makes it light and even empowers his people by arguing that their pronunciation mistakes are deliberate and intentional. To Owago, it is not necessarily a pronunciation problem that they struggle with, but rather something they have power over and can choose to avoid whenever they want.

3.4.2 Inclusivity

Some comedians avoid being offensive by finding a way to include themselves in the group they are poking fun at. They do this either by showing that they are good friends and are therefore in agreement with what that particular group does, or by showing that that particular folly is widespread and not just limited to one specific group. Pablo from Uganda in example (7) above employs inclusivity. In example (14) below, Omondi begins his joke by introducing Kikuyus as friends.

(14)   English with Swahili code mixing (in bold)

I love Kikuyus, I love Kikuyus because they are very, they are very entrepreneurial, they are very business minded yeh. **Kama** last Sunday I was with a Kikuyu friend in church; **nimekaa hapa, beste yangu amekaa hapa.** My tight friend. **Pastor akakuja akasema,** "and Jesus went to heaven to make each one of us a mansion. Everyone will have his own mansion." **Huyo Mkikuyu akaniangalia akaniambia, "Eriko si tukifika heaven tuishi kwa moja turent hiyo ingine?"** That is a business mind.

'I love Kikuyus, I love Kikuyus because they are very, they are very entrepreneurial, they are very business minded, yeh. **Like** last Sunday I was with a Kikuyu friend in church; **I am seated right next to my best friend.** My tight friend. **The pastor said**, "and Jesus went to heaven to make each one of us a mansion. Everyone will have his own mansion." **That Kikuyu turned to me and said, "Eric, when we get to heaven, can't we just stay in one of the mansions and rent out the other one?"** That is a business mind.' (Churchill Live 2011)

In this case Eric Omondi, who is from a different ethnic group, pokes fun at the extreme love of money expressed by Kikuyus; but he makes it less offensive by introducing his joke with the fact that he loves Kikuyus, and the person he is talking about is a very close friend of his – "a tight friend"— so close that his joke cannot be intended to hurt or ridicule him.

3.4.3 Outright apology and careful explanation of their joke

Sometimes the jokes are too satirical and the artists resort to outright apology or justification for their choice of words.

In this next joke Mammito talks about human beings having been created from various types of soil: loam, clay, and sand. She then goes on to say:

(15)   Swahili with English code-mixing

Mafans wa Gor Mahia hao hawakutengenezwa na mchanga, hao walikaviwa kutoka kwa mawe. [Loud laughter.] Hao ndio unasikianga […] Siyo kwa ubaya. Mawe ni kitu ya […] tunasikianga Yesu ni mwamba. Mawe ni kitu ya muhimu sana. Wafans wa Gor Mahia ndio mnasikianga Stone Age people, Stone Age people.

'Gor Mahia fans were not created out of soil but they were carved out of rocks. Those are the people […] No ill intended. Rocks are very important […] we talk of Jesus being a Rock. Rocks are very important. Gor Mahia fans are what you often hear referred to as Stone Age people, Stone Age people.' (Churchill Raw 2014)

In this clip the artist wants to poke fun at the Gor Mahia fans who are mainly of the Luo ethnic group. Gor Mahia is one of the major Kenyan soccer leagues. Many Kenyans know that whenever this team loses a game, there is usually a huge fracas and its fans end up throwing stones at the opposing fans. Mammito, however, tries to defend her use

of the term *mawe* 'rock' here by even calling attention to the fact that Jesus is referred to as the rock of salvation. She is, however, very intentional about discussing the disgusting practice of throwing rocks at opponents. It is a backward practice and people who do that belong to the Stone Age.

While Mammito tries to explain this joke and sounds apologetic to show her sensitivity to local ethnicity, she doesn't do the same for groups that will not pose danger.

(16) Swahili with English code-mixing

Wazungu, hao walitengenezwa kutoka kwa sand, sand ni ngumu kumold, ndio hawananga shape.

'Europeans, they were created from sand, sand it is difficult to mold, and that is why they don't have a shape.' (Churchill Raw 2014)

Apparently since there are hardly any Europeans in her audience, Mammito does not feel obligated to offer any apologies. However, the analogies she creates are very humorous and depend on shared knowledge.

4 Conclusion

While the Kenyan humor industry is still in its developmental stages, there is every evidence that it is a growing field and Kenyans are rapidly learning to appreciate their own humor and to reward their artists. The humor captures the ethnolinguistic diversity inherent in the Kenyan society and shows that the humorists are fully aware of this diversity and are very cautions not to intentionally offend specific ethnicities. Kenyans are also learning to laugh at their follies, not necessarily because they are proud of them but because these follies define who they are as Kenyans. This humor seems to be a way forward for Kenyans to reflect on various issues and maybe serves as a channel for correcting socially unacceptable behavior without causing too much pain and repercussions.

The jokes show a clear interface between language ideologies that shape people's language attitudes. It is evident that speakers use language, dialects, and accents as identity markers.

References

Apte, L. Mahadev. 1985. *Humor and laughter: An anthropological approach.* Cornell, NY: Cornell University Press.

Apte, L. Mahadev. 1987. Ethnic humor versus "sense of humor": An American sociocultural dilemma. *American Behavioral Scientist* 30. 27–41. DOI:http://dx.doi.org/10.1177/000276487030003004

Attardo, Salvatore & Victor Raskin. 1991. Script theory revis (it) ed: Joke similarity and joke representation model. *HUMOR: International Journal of Humor Research* 4(3-4). 293–347.

Berk, Ronald. 1999. *Humor as an instructional defibrillator: Evidence based techniques in teaching and assessment.* Sterling: Stylus publishing.

Carroll, Noël. 2014. *Humour: A very short introduction.* Oxford: Oxford University Press.

Chiaro, Delia. 1992. *The language of jokes: Analyzing verbal play.* London: Routledge.

Churchill Live. 2011. Episode 15 part 3. [video file]. Retrieved from: https://youtube.com/QTU85uhtLOc.

Churchill Raw. 2014. Churchill Raw, Season 2, Episode 38.

Churchill Show. 2013. Oct 7. Owago Onyiro— prefixes and suffixes [video file]. Retrieved from: www.youtube.com/watch?v=6evFwvAEJvQ.

Churchill Show. 2015a. May 14. Majimoto in Meru [video file]. Retrieved from: https://www.youtube.com/watch?v=40S6z929yQo.

Churchill Show. 2015b. Season 04, Episode 51 [video file]. https://www.youtube.com/watch?v=G7fGTQ3DDLw, accessed 2015-03-15.

Davies, Christie. 1990. *Ethnic humor around the world: A comparative analysis.* Bloomington: Indiana University Press.

Gonzales, M. Estrella & Richard Wiseman. 2005. Ethnic identification and the perceived humor and rudeness of ethnic jokes. *Intercultural Communication Studies* 14(2). 170–183.

Gulas, Charles S. 2006. *Humor in advertising: A comprehensive analysis.* Armonk, N.Y.: M.E. Sharpe, Inc.

Kimani, Sheila. 2014. Daniel Ndambuki a. K. A Churchill. Standard Digital 3 November 2014.

Labrador, Roderick. 2004. We can laugh at ourselves: Hawaii ethnic humor, local identity and the myth of multiculturalism. *International Pragmatics Association* 14(2/3). 291–316.

Norrick, Neal & Delia Chiaro (eds.). 2009. *Humor in interaction.* Amsterdam/Philadelphia: John Benjamins.

Omogi, Reuben. 2012. Sep 23. Kenyan comedy [video file]. Retrieved from: https://youtu.be/vtGMJsZVm3Q?t=3m.

Rappoport, Leon. 2005. *Punchline: The case for racial, ethnic, and gender humor.* Westport, CN: Praeger Publishers.

Raskin, Victor. 1985. *Semantic mechanisms of humor.* Boston: D. Reidel Publishing Company.

Rhoades, John. 1977. *Linguistic diversity and language belief in Kenya: The special position of Swahili.* Syracuse, NY: Maxwell School of Citizenship & Public Affairs, Syracuse University.

Tumusiime, James. 2013. *What makes Africans laugh? Reflections of an entrepreneur in humour, media and culture.* Kampala, Uganda: Fountain Publishers.

Ziv, Avner (ed.). 1988. *National styles of humor.* Westport, CT: Greenwood Press.

Name index

Language index

Did you like this book?

This book was brought to you for free
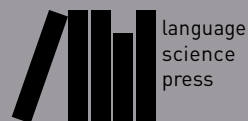
Diversity in African languages

Diversity in African languages contains a selection of revised papers from the 46th Annual Conference on African Linguistics, held at the University of Oregon. Most chapters focus on single languages, addressing diverse aspects of their phonology, morphology, semantics, syntax, information structure, or historical development. These chapters represent nine different genera: Mande, Gur, Kwa, Edoid, Bantu, Nilotic, Gumuzic, Cushitic, and Omotic. Other chapters investigate a mix of languages and families, moving from typological issues to sociolinguistic and inter-ethnic factors that affect language and accent switching. Some chapters are primarily descriptive, while others push forward the theoretical understanding of tone, semantic problems, discourse related structures, and other linguistic systems. The papers on Bantu languages reflect something of the internal richness and continued fascination of the family for linguists, as well as maturation of research on the family. The distribution of other papers highlights the need for intensified research into all the language families of Africa, including basic documentation, in order to comprehend linguistic diversities and convergences across the continent. In this regard, the chapter on Daats'íin (Gumuzic) stands out as the first-ever published article on this hitherto unknown and endangered language found in the Ethiopian-Sudanese border lands.