

Afranaph Technical Report #8

On the Categories of Clausal Constituents in Lubukusu and Limits to their Selection

Mark Baker[#], Ken Safir[#], and Justine Sikuku⁺
Rutgers University[#] and Moi University⁺

Abstract: The Bantu language Lubukusu has seven distinct kinds of clause-like constituents: finite clauses, subjunctives, actuals, derived nominals, and three types of infinitive, which differ in how their subjects are expressed. We show that two of these clause types are actually headed by a null noun, whereas four are not, and one (the null subject infinitive) is structurally ambiguous in this respect. This structural difference gives a unified account of several empirical differences among the clauses: whether they can be complements of prepositions, whether they can be in the subject position, whether they can extrapose to the right edge of the clause, and whether a phrase can be extracted from them. We then explore the implications of this category difference for a theory of selection, showing that the phrases that are headed by a noun cannot be specially selected for by the matrix verb, whereas those that are not headed by a noun can be. We infer from this that the selection of a verb for its clausal complement must be a syntactic matter as well as a semantic one.

1. Introduction

Since Chomsky (1965), it has been understood that the arguments that satisfy a given predicate are selected both in a semantic and syntactic sense. In other words, it is not enough to know of an argument what sort of semantic type it is; it is also necessary to know what sort of syntactic type it is. For verbs that semantically select for arguments that correspond to events or propositions rather than just animate and inanimate entities, syntactic questions arise about the category of the

argument involved and the influence that a predicate has on the syntactic categories of the arguments that are compatible with it.

Like many languages, the Bantu language Lubukusu (Luyia subgroup, spoken in Western Kenya) has a rather wide variety of clauses and clause-like constituents that correspond to events or propositions. In this paper, we explore the distinctions between those that behave externally like nominals and those that behave externally like non-nominal clauses. At least seven different types of clause-like constituents can be distinguished in Lubukusu, all of which can function as complements to a verb. We discuss the patterns that arise with these various types of clause with respect to several syntactic phenomena: ability to function as the object of a preposition, ability to sit in the subject position, ability to trigger subject agreement, compatibility with clefting the object, ability to extrapose to the right edge of the clause, and status as an island for the extraction of an object. We show that the seven types of clauses fall into three major classes: non-nominal, fully nominal, and one intermediate type, and we propose syntactic structures that account for these patterns. We then finish our study by drawing some conclusions about what aspects of its complement a verb can select for and what aspects it cannot select for, based on the observation that verbs cannot select for particular nominal clause types in the same distinctive way that they can select for particular non-nominal clause types.

2. The Cast

We begin by introducing the seven clause-like constituents in Lubukusu, with a brief outline of the distinguishing properties of each.

The most noun-like of all is what we will call the **li-** nominal shown in (1).¹

(1) **Ba-saani b-eny-a [li-nyw-a lye ka-ma-lwa].**

c2-men SM.c2.TNS-want-fv c5-drink-fv c5.AM c6-c6-beer

‘The men want the drinking of beer.’

Wekesa SM.c1.TNS.want-fv c15-praise-fv Wafula c15-AM Andrea.

‘Wekesa wants ANDREA to praise Wafula.’

Poss subject infinitives are somewhat like **li**-nominals in that a verb stem bears an apparent noun class prefix, here **khu-**, and the subject follows the nominalized verb and is marked with the AM **khu+a/e**. Class 15 **khu-** is a special noun class for infinitival verbs that is not used with simple underived nouns. This construction is also extremely productive—maybe totally productive. Unlike the **li**-nominal, the overt object of the verb stem in a poss subject infinitive may and must be a bare NP, not one marked with an AM, and it needs to be adjacent to the verb, between the verb and the subject. Multiple objects are possible, and objects are as obligatory in poss subject infinitives as they are in finite clauses. Adverbs are freely allowed as well. Unlike the **li**-nominal, the poss subject infinitive can bear an object marker (OM) (**khu-mu-fumy-a khw-a Andrea**, ‘Andrea’s praising him’); passive, reflexive, and reciprocal affixes are all possible as well. The verb stem bears no subject agreement, however. There is no tense marking, but habitual aspect is possible. Overall, this Lubukusu constituent is roughly like so-called Poss **-ing** constructions in English such as **Mary’s praising the child**. Very similar constructions exist in Chishona (Myers, 1987), Kiswahili (Carstens, 1991), Chichewa (Bresnan and Mchombo (1995), Logooli (Mugane 2003a), and Gĩkũyũ (Mugane 2003b) among others. For Kiswahili, Carstens calls this a Poss-*Ing* construction, using the English term. In Kiswahili, Logooli and Gĩkũyũ, the prefix on the nonfinite verb form is the same as the one on the derived nominal. As a result, the way that the direct object is expressed is the only salient difference between these first two constructions in these languages.

A third type of clause-denoting constituent in Lubukusu is the null subject infinitive, shown in (3).

(3) a. **Wekesa eny-a** [khu-fumy-a Andrea].

Wekesa SM.c1.TNS.want-fv c15-praise-fv Andrea.

‘Wekesa wants to praise Andrea.’

b. **A-khalak-a** [khu-ly-a e-ng’eni].

SM.c1.TNS-decide-fv c15-eat-fv c9-c9-fish

‘He decided to eat (a) fish.’

As far as their internal syntax is concerned, null subject infinitives are exactly like poss subject infinitives except for the treatment of the subject, which here is left unpronounced. If the null subject infinitive functions as a complement, its subject behaves like a PRO in English, and it is controlled by a designated argument of the matrix clause (the matrix subject in (3)). If a null subject infinitive functions as a subject, the null subject can be interpreted as a PRO_{arb}, referring roughly to people in general. These Lubukusu constituents are analogous to PRO **-ing** gerunds (**praising Andrea**) or control infinitives (**to praise Andrea**) in English. The analog of this construction also exists in Kiswahili (Carstens 1991), Chichewa (Bresnan and Mchombo 1995), Loloogi (Mugane 2003a), Gikūyū (Mugane 2003b) and presumably many other Bantu languages.²

Next we have the bare subject infinitive, shown in (4).

(4) **Ba-saani b-eny-a** [Wekesakhu-nyw-a ka-ma-lwa].

c2-men SM.c2.TNS.want-fv Wekesa c15-drink-fv c6-c6-beer

‘The men want Wekesa to drink beer.’

Here too the verb form has the apparent noun class prefix **khu-**, but now the subject is a simple overt NP (not marked with an AM) that appears in the normal subject position, before the verb. Except for this difference, the bare subject infinitive behaves just like the poss subject infinitive and the null subject infinitive with respect to its internal syntax. This construction is roughly

comparable to ACC-**ing** constructions (**Wekesa drinking beer**) or **for-to** infinitives (**for Wekesa to drink beer**) in English. Diercks (2012) discusses this Lubukusu construction from the point of Case theory, showing that it also exists in Swahili and Digo. Carstens (1991: ch.5) also discusses the Swahili version in some detail. However, it is not mentioned for Chichewa by Bresnan and Mchombo (1995) or Mchombo (2004), nor by Mugane (2003a) for Logooli. Mugane (2003b:244) explicitly says of Gĩkũyũ that its **ku**+verb constructions “cannot have subjects in the internal syntax”, which we take to mean that there is no bare subject infinitive in that language. This construction may thus have a narrower distribution within the Bantu languages than the other two class-15 constructions, its existence possibly related to the fact that the Case filter does not seem to hold in Lubukusu and certain other Bantu languages (Diercks, 2012); as a result, there is no penalty on having an overt subject with no obvious case marker in this language. However, determining the crosslinguistic distribution of this construction and its causes goes beyond our topic here.

The first four clause-denoting complements just described are non-finite. Lubukusu also has a set of finite clauses, where a finite clause is one that manifests agreement in the form of a “subject marker” (SM) that corresponds to the noun class of the clause’s subject. Indicative finite clauses, illustrated in (5), do not treat the subordinate clause as nominalized in any obvious way.

- (5) **Ba-saani ba-khalak-a [bali/ali/?mbo Wanjala a-nyw-a**
 c2-men SM.c2.TNS-decide-fv COMP Wanjala SM.c1.TNS-drink-fv
ka-ma-lwa].
 c6-c6-beer

‘The men decided that Wanjala drank beer.’

The verb in an indicative clause is inflected for tense-aspect-mood and subject agreement just like a matrix verb in this language, and, when subordinated, the clause is introduced by an overt complementizer: **bali**, **mbo**, or the stem **-li** inflected to agree with the matrix subject in person, number and noun class (see Diercks 2010, to appear).³ These are clearly like finite complementizer phrases (CPs) in English (**that Wekesa drank beer**) and other languages, except that the complementizer cannot normally be dropped, as it can be for many indicative complement clauses in English. Such clauses are also found in Chichewa (Mchombo 2004:27-28), Kiswahili, Kinande and others—with some differences in the syntax of the C head itself, such as whether it agrees or not.

Lubukusu also has a special type of subjunctive complement, shown in (6). This is distinguished from indicative CPs in that there does not need to be an overt complementizer, and the verb bears the suffix **-e**, rather than the otherwise ubiquitous final vowel **-a**.

- (6) **Ba-saani ba-khalak-a** [**Wanjala a-nyw-e** **ka-ma-lwa**].
 c2-men SM.c2.TNS-decide-fv Wanjala SM.c1-drink-SBJV c6-c6-beer
 ‘The men decided Wanjala should drink beer.’

Although subject agreement is present in these clauses, there is no morpheme in the normal tense slot after the SM and before the verb stem. Other than this, the morphosyntax of subjunctive clauses is just like that of a matrix clause. The subject of the subjunctive clause can be null, with a result that looks very similar to a null subject infinitive. However, the null subject can still refer to a prominent antecedent in the discourse not mentioned in the matrix clause, so we take it to be an ordinary null pronoun *pro* licensed by agreement, not PRO, the special subject of control. Semantically, a subjunctive clause is interpreted as expressing an unrealized possibility that is temporally posterior to the eventuality denoted by the matrix clause. Such clauses can have time

adverbs that are different from time adverbs in the matrix, as long as the adverb in the subjunctive clause picks out an interval of time that follows that of the matrix. There is no very close syntactic parallel to this kind of complement in English. A cognate subjunctive form is present in many other Bantu languages, however, including Chichewa (Mchombo 2004:28-29). (A difference is that in Chichewa subjunctives apparently always appear with the overt finite complementizer **kuti**, whereas in Lubukusu an overt complementizer is possible, but often does not appear.)

Finally, Lubukusu has what we call *actual* clauses, like the one illustrated in (7).

- (7) **Wafula a-a-khol-a** [**Wekesa a-a-kul-a** **si-tabu**].
 Wafula SM.c1-TNS-make-fv Wanjala SM.c1-ACT-buy-fv c7-book
 ‘Wafula made Wekesa buy a book (and he actually did).’

These are syntactically like subjunctive clauses in many respects: they also can appear without an overt complementizer, they do have subject agreement, and a null subject behaves like *pro* rather than *PRO*. The striking semantic difference between these clauses and subjunctives is that the eventuality they describe must be actual—true and occurring or occurred—in the utterer’s world. The morphosyntactic difference is that the verb bears the prefix **a-** with nonhigh tone in the T position after the SM and before the verb stem, as well as the normal final vowel **-a**. Hence, (6) does not imply that Wekesa actually drank the beer, only that it was decided that he would, whereas (7) does imply that the buying did in fact happen. Actual clauses cannot be simple matrix clauses, unlike indicatives, and they are entailed to be true, not presupposed.⁴ Again, there is no direct syntactic (or semantic) analog of this in English. The previous generative literature on Bantu languages has rarely if ever recognized these actual clauses as a distinct type, so we have no idea how widespread they might be. For further discussion and analysis of actual clauses, see Safir, Baker, and Sikuku (2012).

Although we have briefly described the salient properties of the internal syntax of these various kinds of clausal complements for the reader's orientation, they are not our primary concern here. Rather, we focus on the external syntax of these constituents—where they can and cannot appear in larger constructions. So far we have seen that they can all function as the complement of a matrix verb. However, in many other respects, differences in their external syntax appear, suggesting that some are “more nominal” than others. In the sections that follow, we present a range of data relevant to this. Nevertheless, we argue that there is no theoretical need to distinguish a wide range of degrees of nominality, in the manner of Ross's (1973) nouniness squish, or more recent functionalist-typological work that presents nominality as a continuum (Koptjevskaja-Tamm, 1993, Malchukov, 2004). Rather, we show that—with a few interesting qualifications—each clause type in Lubukusu is either fully nominal or not with regard to its external distribution.

More specifically, we claim that **li**-nominals and poss subjects infinitives are fully nominal: they have both phi-features and the power to bind pronouns. In contrast, the C-less clauses (subjunctives and actuals) are totally nonnominal: they do not have phi-features and they cannot bind pronouns. Only one intermediate situation is allowed, that of expressions that are “argumental” but not fully nominal, expressions that do not have a full set of phi-features but can nevertheless bind pronouns/variables. We claim that both finite CPs and bare subject infinitives exemplify this intermediate possibility. Finally, we claim that null subject infinitives are structurally ambiguous: they can have the structure of bare subject infinitives, in which they are less than fully nominal, or that of poss subject infinitives, in which case they are fully nominal. Once that is accomplished, we go on to draw some preliminary consequences of the partition

between nominals and clauses for a theory of selection—the theory of what requirements a matrix verb can put on the structure, form, and meaning of its clausal complement, and why.

3. Clauses as the Object of a Preposition

The first environment in which a distinction among the different clause types can be found in Lubukusu is as the object of a preposition. Like other Bantu languages, Lubukusu has only a small number of prepositions and some of these (e.g., **ne** ‘with’) have meanings that may not lend themselves to taking clausal complements. However, (8) shows that the preposition **khu** when it appears with a verb of speaking can mean ‘about’.

(8) **Wekesa a-kachul-a khu Wafula.**

Wekesa SM.c1.TNS-talk-fv about Wafula

‘Wekesa talked about Wafula.’

This meaning is suitable to having a clausal complement, since one can talk about events, possibilities, and propositions as much as one can about people and things.

The examples in (9) show that **li**-nominals, poss subject infinitives, and bare subject infinitives are possible as objects of **khu**.

(9) a. **Wekesa a-som-a [khu-li-ly-a].**

Wekesa SM.c1.TNS-read-fv about-c5-eat-fv

‘Wekesa read about eating.’

b. **Wekesa a-kachul-a/a-som-a [khu-ch-a khw-a Wafula].**

Wekesa SM.c1.TNS-talk/read-fv about.c15-go-fv c15-AM Wafula.

‘Wekesa talked/read about Wafula’s going.’ (note: *khu+khu+V* truncates to *khu+V*)

- c. **Wekesa a-kachul-a [khu-ch-a].**

Wekesa SM.c1.TNS-talk-fv about.c15-go-fv

‘Wekesa talked about going.’ (note: *khu+khu+V* truncates to *khu+V*)

There is a minor complication in the examples in (9b,c). The preposition **khu** ‘about’ and the class 15 prefix **khu-** are homophonous, and in this case **khu+khu+V** truncates to **khu+V**. We take this to be the result of a common kind of haplology, where a sequence of two identical formatives is simplified at PF—something that has no syntactic significance.⁵

In contrast to (9), (10) shows that bare subject infinitives are very marginal as complements to the preposition **khu**, and the various types of finite clause are completely out.⁶

- (10) a. **??Wekesa a-som-a (khu) [Wafula khu-ch-a].**

Wekesa SM.c1.TNS-read-fv (about) Wafula (about-)c15-go-fv

‘Wekesa read about Wafula going.’

- b. ***Wekesa a-kachul-a khu [bali Wafula a-a-ch-a].**

Wekesa SM.c1.TNS-talk-fv about that Wafula SM.c1-TNS-go-fv

‘Wekesa talked about (it) that Wafula went.’

- c. ***Wekesa a-kachul-a khu [(Wafula) a-ch-e].**

Wekesa SM.c1.TNS-talk-fv about (Wafula) SM.c1-go-SBJV

‘Wekesa talked about (Wafula) going (future, unrealized).’

(NB: with *Wafula* it is OK as ‘Wekesa talked about Wafula, in order for him to go.’)

- d. ***Wekesa a-kachul-a khu [(Wafula) a-a-ch-a].**

Wekesa SM.c1.TNS-talk-fv about (Wafula) SM.c1-ACT-go-fv

‘Wekesa talked about Wafula going (and then he actually did).’

(NB: OK as ‘Wekesa talked about Wafula in order for him to go, and he did.’)

These facts about Lubukusu are quite similar to familiar facts in English, where NPs can be objects of P, including derived nominals and certain gerunds, but finite CPs cannot be objects of P (Stowell, 1981) among others). The relevant constraint is not universal, however; for example, Taraldsen (1984) shows that finite CPs can be objects of Ps in Norwegian, even though they are nonnominal in other respects. With these Germanic cases also in mind, we can state the following, which is presumably the consequence of deeper principles of grammar that we cannot investigate here.

- (11) In every language in which the complement of P cannot be a tensed sentence, the complement of P must be nominal.⁷

Given this generalization plus the fact that Lubukusu does not allow a finite indicative CP as the complement of P ((10b)), we can infer that the various clause types in Lubukusu differ in their nominality: three of them are apparently headed by some kind of noun, whereas the other four are not. And the dividing line does not come exactly where one might have expected it: it splits one type of c15 infinitive, the bare subject infinitive, from the other two.

4. Clauses in subject position

Next we consider the possibility of clause-like constituents in preverbal subject position. Here we see the same distinction in nominality replicated. However, these data are somewhat more nuanced, revealing also a secondary distinction that separates the two C-less finite clauses from finite CPs and bare subject infinitives.

Consider first **li-** nominals. These can appear readily in subject position, as shown in (12). When they do, they trigger the expected class 5 agreement on the verb, matching the class 5 prefix on the nominalized verb. (12b) shows that it is also possible to cleft the object past the

subject in this case. This suggests that the **li**-nominal is in the true subject position, SpecTP, and not in some other position at the left periphery of the clause. (This contrasts with data below.)

- (12) a. [**Li-ch-a ly-a Andrea**] **ly-a-sim-isy-a** **o-mw-aalimu.**
 c5-go-fv c5-AM Andrea SM.c5-TNS-like-CAUS-fv c1-c1-teacher
 ‘Andrea’s going pleased the teacher.’
- b. **O-mw-aalimu niye niye** [**li-ch-a ly-a Wekesa**] **ly-a-siim-isy-a.**
 c1-c1-teacher FOC.c1 FOC.c1 c5-go-fv c5-AM Wekesa SM.c5-TNS-like-CAUS-fv
 ‘It’s the teacher that Wekesa’s going pleased.’

The examples in (13) and (14) show that poss subject infinitives and null subject infinitives are just like **li**-nominals in these respects. They appear readily in subject position ((13a), (14a)), they trigger only the expected agreement on the verb ((13b), (14b))—in this case, c15 **khu-**, not c5, as in the **li**-nominals—and it is possible to cleft an object past the subject in these cases too ((13c) and (14c)).

- (13) a. [**Khu-mu-many-isy-a khw-a Andrea**]
 c15-OM.c1-know-CAUS-fv 15.AM Andrea
khw-a-sim-isy-a **o-mw-alimu.**
 SM.c15-TNS-like-CAUS-fv c1-c1-teacher
 ‘Andrea’s introducing her pleased the teacher.’
- b. ***[Khu-ly-a e-n-yama khw-e Wekesa]li-nyal-a ly-a-belesy-a.**
 c15-eat-fv c9-c9-meat c15-AM Wekesa SM.c5-can-fv SM.c5-TNS-annoy-fv
 ‘Wekesa’s eating meet can be annoying.’ (OK with **khu-nyala khw-a-belesya** (c15))
- c. **Wafula ni-ye ni-ye** [**khu-ch-a khw-a Wekesa**] **khw-a-siim-isy-a.**
 Wafula FOC-c1 FOC-c1 c15-go-fv c15-AM Wekesa SM.c15-T-like-CAUS-fv

‘It’s Wafula that Wekesa’s going pleased.’

- (14) a. **[Khu-nyw-a ka-ma-lwa] khw-a-siim-isy-a Wafula.**

c15-drink-fv c6-c6-beer SM.c15-like-CAUS-fv Wafula.

‘To drink beer pleased Wafula.’

- b. ***[Khu-ly-a e-n-yama] li-nyal-a ly-a-belesy-a.**

c15-eat-fv c9-c9-meat SM.c5-can-fv SM.c5-TNS-annoy-fv

‘Eating meat can be annoying.’ (OK with: **khu-nyal-a khw-a-belesy-a**)

- c. **Wafula ni-ye ni-ye [khu-ch-a] khw-a-siim-isy-a.**

Wafula FOC-c1 FOC-c1 c15-go-fv SM.c15-TNS-like-CAUS-fv

‘It’s Wafula that going pleased.’

To the extent, then, that only nominal phrases can be subjects in SpecTP, these phrases are just as nominal as a **li**-nominal or a simple NP.

Bare subject infinitives are a bit different in these respects. They are sometimes dispreferred in the subject position ((15a)), whereas they are completely comfortable as objects. Moreover, when they do appear in the subject position, they sometimes trigger class 5 agreement **li**- ((15b)), rather than the expected class 15 prefix seen in (13b) and (14b). Finally, clefting an object over this kind of subject is awkward ((15c)), noticeably worse than clefting it over a null subject infinitive or a poss subject infinitive of similar length and complexity.

- (15) a. **(?) [Andrea khu-many-isy-a o-mw-aana] khw-a-siim-isy-a o-mw-aalimu.**

Andrea c15-know-CAUS-fv c1-c1-child SM.c15-TNS-please-fv c1-c-1-teacher

‘Andrea introducing the child pleased the teacher.’

- b. **[Ba-b-aami khu-ly-a e-n-yama] li-nyal-a ly-a-belesy-a.**

c2-c2-leaders c15-eat-fv c9-c9-meat SM.c5-can-fv SM.c5-TNS-annoy-fv

‘The leaders eating meat can be annoying.’

- c. **?Wafulani-ye ni-ye [Wekesakhu-ch-a] ly/khw-a-siim-isy-a.**

Wafula FOC-c1 FOC-c1 Wekesa c15-go-fv SM.c5/c15-TNS-like-CAUS-fv.

‘It’s Wafula that Wekesa going pleased.’

The examples in (16) provide a head-to-head comparison between a bare subject infinitive and a poss subject infinitive, showing that the latter is better than the former in subject position. (In contrast, the opposite preference is very often expressed in complement position; see below.)

- (16) a. **?[Wafula khu-ch-a] khw-a-kan-ibw-a.**

Wafula c15-go-fv SM.c15-TNS-forbid-PASS-fv

‘Wafula going was forbidden.’ (worse than (16b))

- b. **[Khu-ch-akhw-a Wafula] khw-a-kan-ibw-a.**

c15-go-fv c15-AM Wafula SM.c15-TNS-forbid-PASS-fv

‘Wafula’s going was forbidden.’

Carstens (1991) makes some similar observations for Kiswahili. She mentions (p. 194, n. 11) that bare subject infinitives (her Acc-**ing** gerunds) are not acceptable in subject position for all Swahili speakers (implying by contrast that poss subject constructions are uniformly possible). She also shows that poss subject constructions can be the subjects of relative clauses in Kiswahili—which may involve both the head of the relative and the finite verb moving past the subject—whereas bare subject constructions cannot be subjects in this construction. This contrast is similar to ours between (13c) and (15c).⁸ Similarly, Mugane (2003b:252-253) shows that poss subject infinitives (and other infinitives with nominal modifiers) can appear in subject position in Gĩkũyũ, but infinitives without nominal modifiers and with more verbal properties (such as aspect markers) cannot be.

Indicative CPs are very much like bare subject infinitives in these respects. They too are not impossible but sometimes dispreferred or a bit uncomfortable in subject position. When they are accepted, they can also trigger class 5 agreement **li-** (not class 15 agreement **khu-**) even though they do not themselves bear a class 5 gender prefix ((17a)).⁹ And they too do not like to have an object clefted past them to sentence initial position; indeed (17b) is sharply ungrammatical. Clefting the object is possible in Lubukusu only if the finite sentential subject is extraposed to the right, as shown in (17c) (see section 4 below).

- (17) a. **[Bali Wafula a-a-ch-a] ly-a-siim-isy-a Wekesa.**
 COMP Wafula SM.c1-TNS-go-fv SM.c5-TNS-like-CAUS-fv Wekesa
 ‘That Wafula went pleased Wekesa.’
- b. ***Wekesa ni-ye ni-ye [bali Wafula a-a-ch-a]**
 Wekesa FOC-c1 FOC-c1 COMP Wafula SM.c1-TNS-go-fv
ly-a-siim-isy-a.
 SM.c5-TNS-like.CAUS-fv
 ‘It is Wekesa that that Wafula left pleased.’
- c. **Wekesa ni-ye ni-ye ly-a-siim-isy-a [bali Wafula a-a-ch-a].**
 Wekesa FOC-c1 FOC-c1 SM.c5-TNS-please-fv COMP Wafula SM.c1-TNS-go-fv
 ‘It is Wekesa that it pleased that Wafula left.’

The behavior of finite CPs in subject position in Lubukusu is similar to that in English. CPs in English also seem to appear in the subject position ((18a)). However, they do not agree with T as robustly as NPs do, since the conjunction of two CPs is (under some conditions) treated as singular, whereas the conjunction of NPs is always plural ((18b), cf. McCloskey 1991).

- (18) a. **That John dropped his plate is unfortunate.**

- b. **[That John dropped his plate and that Mary saw him do it] is/?? are unfortunate.**

Moreover, other elements cannot move past clausal subjects in English either; in particular, **wh**-phrases and Tense heads cannot move past a CP “subject” into SpecCP or the highest C position in an English question. Finite CP “subjects” are strikingly worse than equally complex Poss-**ing** gerunds in this respect.

- (19) a. ***Who did that John dropped his plate upset?**

- b. **Who did John’s dropping his plate upset?**

The contrast between (19a) and (19b) in English is very reminiscent of the one between (13c) and (17b) in Lubukusu.

A classic generative analysis of (19) and related facts in English is that CP in English cannot truly occupy SpecTP. Rather it is a kind of topic in the left periphery of the clause, which binds an empty category of some sort in the true subject position. That is why a filled C position cannot come before it in (19a) (Koster, 1978, Stowell, 1981). As Carstens (1991) does for Kiswahili, we propose essentially the same analysis for Lubukusu. More specifically, we assume that CP in Lubukusu has no number-gender features—which is very plausible, given that it has no discernible class prefix of its own, the way that true nominals do. However, there is a firm requirement that T must agree with the phrase in SpecTP in Lubukusu, as in other Bantu languages (Baker, 2003, Baker, 2008, Carstens, 2005). Therefore, it stands to reason that a phrase without phi-features cannot occupy SpecTP directly in this language. What it can do is occupy a discourse-related position in the left periphery of the clause and bind a null pronoun in the true SpecTP position. Since the CP is in a discourse sensitive position, it is not surprising that judgments as to its acceptability are somewhat variable in simple elicitation situations, where the discourse context is not controlled for. Since it is not in SpecTP but rather binds a pronoun there,

T can agree with the pronoun, rather than directly with CP itself. Indeed, the class 5 **li-** is one form of expletive subject agreement in other sentence types in Lubukusu as well (see section 4). Finally, since CP is in the left periphery of the clause, not in SpecTP, it is not unexpected that there would be interference with other movement into the left periphery, as in clefting, yielding the badness of (17b). Indeed, we propose more precisely that the clause is in SpecTopP and focusing puts an NP in SpecFocP, where FocP is below TopP, not above it, in the style of Rizzi (1997) and much related work. On this view, we would expect that (17b) would become much better with the finite CP before the clefted/focused object, and indeed this is true:

- (20) (?)**[bali Wafula a-a-ch-a], Wekesa ni-ye ni-ye ly-a-siim-isy-a.**
 that Wafula SM.c1-TNS-go-fv Wekesa FOC-c1 FOC-c1 SM.c5-TNS-like-CAUS-fv
 ‘That Wafula left, it is Wekesa that it pleased.’

The schematic structures we have discussed here are summarized in (21).

- (21) a. [_{VP} verb [_{XP}]] condition on XP: Must be compatible with selecting verb,
 but otherwise OK for NP and CP alike
- b. [_{TP} [_{XP}] Agr+T VP] condition on XP: must have features for T to agree with;
 OK for NP, not for CP
 Allows [_{FocP} NP Foc [_{TP} [_{XP}] Agr+T VP]
- a. [_{TopP} [_{XP_i]} [_{TP} *pro_i* Agr+T VP] condition on XP: must be able to bind pronoun in TP.
 OK for CP, but T agrees with *pro* not CP itself.
 Not allowed: [_{FocP} NP Foc [_{TopP} [_{XP_i]} [_{TP} *pro_i* Agr+T VP]

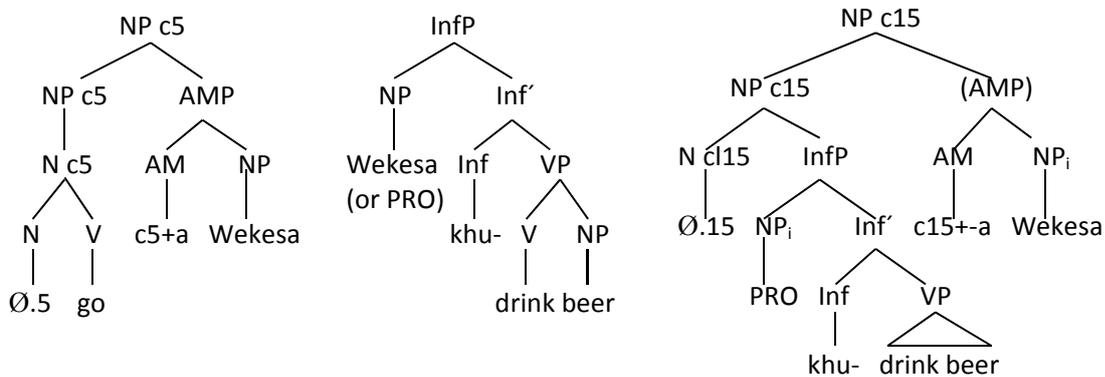
The pattern of facts for bare subject infinitives in Lubukusu is very much like the pattern of facts for indicative CPs. Therefore we extend the same analysis to them, concluding—more surprisingly—that these constituents also do not have gender and number features. This bars

them from true SpecTP position and goes much of the way toward explaining the variability in subject agreement with preverbal bare subject infinitives: T is not really agreeing with the infinitive at all, but rather with the null subject that it binds. The relationship between T and the bare subject infinitive is thus indirect, mediated by a *pro*, and this makes possible a degree of mismatch between the apparent gender of the infinitive and the subject agreement on the verb. We also expect the degradation when clefting happens over a bare subject infinitive, as seen in (15c)—although we have no formal explanation for why the degradation is sharper with indicative CPs than with bare subject infinitives.¹⁰

By way of contrast, the same line of reasoning implies that poss subject infinitives do have true phi-features—specifically class 15 features, just as **li-** nominals have c5 features. We can conclude from this that the **khu-** prefix on the verb is not itself truly a nominalizing prefix, nor does it bear class 15 features, since the same prefix is present both in constituents that have phi-features (poss subject infinitives, null subject infinitives) and in ones that do not (bare subject infinitives). On this point we disagree with Myers (1987) (for Chishona), but agree with Carstens (1991) for Swahili and Bresnan and Mchombo (1995:233-235) for Chichewa, who claim that **ku**+verb in those languages can be either nominal or verbal in nature (see also Mugane (2003b) on Gikũyũ.). More specifically, in our version we take **khu-** on verbs to be of category T, not of category N; it is a nonfinite tense marker, not a nominalizer. As such, it can head a TP, which does not have phi-features.¹¹ This is the structure of a bare subject infinitive. But this TP can also function as the complement of a null noun head which does have phi-features, specifically class 15 features. When the nonfinite TP is embedded in an NP with a null head in this way, it can have a phrase headed by the associative marker adjoined to it. This is an option for any ordinary NP in Lubukusu, used for expressing the possessor of the NP (e.g., **si-tabu sy-a Wekesa** ‘c7-book c7-

AM Wekesa’, ‘Wekesa’s book’) and other kinds of nominal modification. For these event-denoting NPs the normal—perhaps the only— interpretation of the possessive-like NP is as controlling a PRO subject inside the nonfinite NP. This results in a poss subject infinitive. The two structures are compared in (22). Also included is a **li**-nominal, where we assume that a null noun of class 5 combines directly with the verb—not with a whole vP, as the null noun of class 15 does. This accounts for the fact that the internal structure of **li**-nominals is essentially the same as that of NPs with simple noun heads in Lubukusu.¹²

(22) a. **li**-nominal b. bare subj infinitive c. poss and null subj infinitive



Note also that we allow the subject in SpecInfP (=TP) to be either overt, as in (22b), or PRO, as in (22c). Following Diercks (2012), we tentatively assume that both options are possible because case theory simply does not hold in Lubukusu, so there is no need for the subject in (22b) to receive case. This is particularly plausible for Lubukusu, in which no NP ever varies in its form depending on its syntactic position—not even pronouns as in English.¹³ The structure in (22c) can also exist without an AMP adjoined to NP, just as it is possible to have an ordinary NP with a simple N and no AMP (i.e., with no possessor or modifier). Thus (22c) without the AMP is a structure for null subject infinitives.

Given the structure in (22c), we would expect that other kinds of NP modifiers, in addition to the AMP, could adjoin to the NP node in poss subject infinitives and null subject infinitives. Indeed, Bresnan and Mchombo (1995:220) and Mugane (2003a:209, 2003b:247) show that a class 15 demonstrative can modify such infinitives in Chichewa, Logooli, and Gikūyū. The same is true in Lubukusu, as shown in (23b,c). (The **li**-nominal in (23a) is also included for comparison.) However, we predict that the demonstrative cannot appear in construction with a bare subject infinitive, since that lacks an NP node to adjoin to. That is also correct, as shown by (23d).

- (23) a. **Wekesa a-a-siim-a li-nyw-a ly-e ka-ma-lwa li-no.**
 Wekesa SM.c1-TNS-like-fv c5-drink-fv c5-AM c6-c6-beer c5-this
 ‘Wekesa likes this drinking of beer.’
- b. **Wekesa a-a-siim-a khu-nyw-a ka-ma-lwa khu-no.**
 Wekesa SM.c1-TNS-like-fv c15-drink-fv c6-c6-beer c15-this
 ‘Wekesa likes this drinking beer.’
- c. **Wekesa a-a-siim-a khu-nyw-a ka-ma-lwa khw-a Wafula khu-no.**
 Wekesa SM.c1-TNS-like-fv c15-drink-fv c6-c6-beer c15-AM Wafula c15-this
 ‘Wekesa likes c1-T-like-fv this drinking beer by Wafula.’
- d. ***Wekesa a-a-siim-a Wafula khu-nyw-a ka-ma-lwa khu-no.**
 Wekesa SM.c1-TNS-like-fv Wafula c15-drink-fv c6-c6-beer c15-this
 ‘Wekesa likes this Wafula drinking beer.’

Parallel facts are found with attributive adjectives, where (24a-c) are possible, but (24d) is not.¹⁴

- (24) a. **Wekesa a-a-siim-a li-nyw-a ly-e ka-ma-lwa li-layi.**
 Wekesa SM.c1-TNS-like-fv c5-drink-fv c5-AM c6-c6-beer c5-good

‘Wekesa likes the good drinking of beer.’

- b. **Wekesa a-a-siim-a khu-nyw-a ka-ma-lwa khu-layi.**

Wekesa SM.c1-TNS-like-fv c15-drink-fv c6-c6-beer c15-good

‘Wekesa likes good drinking of beer.’

- c. **Wekesa a-a-siim-a khu-nyw-a ka-ma-lwa khw-a Wafula khu-layi.**

Wekesa SM.c1-TNS-like-fv c15-drink-fv c6-c6-beer c15-AM Wafula c15-good

‘Wekesa likes Wafula’s good drinking of beer.’

- d. ***Wekesa a-a-siim-a Wafula khu-nyw-a ka-ma-lwa khu-layi.**

Wekesa SM.c1-TNS-like-fv Wafula c15-drink-fv c6-c6-beer c15-good

‘Wekesa likes Wafula good drinking beer.’

The structural analysis in (22) also accounts immediately for the data in the previous section given (11): poss subject infinitives and (some) null subject infinitives have the noun head (null, as it happens) that they need in order to be the object of a preposition, as does the **li**-nominal. The bare subject infinitive, in contrast, does not (nor does a finite CP). So we have converging evidence that some of these constructions count as true NPs whereas others do not. And indeed, these structures are broadly speaking similar to what other generative researchers have proposed for parallel distinctions in other Bantu languages, apart from some differences of detail, many of them theoretically driven.¹⁵

The subject position reveals a second distinction as well. Although finite CPs and bare subject infinitives cannot truly occupy SpecTP, they can create the illusion of doing so by binding the subject position from the left periphery. But even this is impossible for subjunctive and actual clauses. These are completely out in preverbal position, even with default/expletive **li**-agreement on the verb and no constituent clefted passed them, as shown in (25).

(25) a. *[Wekesa a-ch-e] li-kha-siim-isy-e Wafula.

Wekesa SM.c1-go-SBJV SM.c5-FUT-like-CAUS-SBJV Wafula

‘That Wekesa would go would please Wafula.’ (also bad with other SMs)

b. *[A-a-ch-a engo] ly-a-nyal-ikh-a.

SM.c1-ACT-go-fv home SM.c5-TNS-able-STAT-fv

‘That he went home (successfully) was manageable.’

(These examples remain bad even if an overt complementizer like **bali** is added at the beginning.)

We describe this fact by saying that the C-less clauses not only lack phi-features, but even lack the ability to bind a null pronoun in the subject position; they cannot participate at all in anaphora or the formation of a binding chain. As such, they cannot be properly linked to the subject theta role of the verb in (25), and cannot be interpreted at LF.

In summary, then, we have a three-way distinction in “nominality”: there are fully nominal expressions, headed by a noun with phi-features and binding power, intermediate expressions without phi-features but with binding power, and completely non-nominal categories with neither phi-features nor binding power. Each degree of nominality is exemplified by at least two similar-behaving constructions. The overall pattern is summarized in Table One.

Table One

Construction type	Complement of verb	object of P	Preverbal “subject”	True Subject in Spec, TP	Modified by Dem or Adj
li- nominal	Yes	Yes	Yes	Yes	Yes
Poss subject infinitive	Yes	Yes	Yes	Yes	Yes
Null subject infinitive	Yes	Yes	Yes	Yes	Yes
Bare subject infinitive	Yes	No	Yes	No (?)	No

Indicative CP	Yes	No	Yes	No (*)	(No)
Subjunctive clause	Yes	No	No	No	(No)
Actual clause	Yes	No	No	No	(No)

5. Extraposition of subjects

Next we consider a syntactic context that favors non-nominal constructions over nominal ones, namely extraposition of subjects. A priori, we would expect this to give the same pattern as PPs or sentential subjects, but in reverse: the clause types that were bad in PPs and in true SpecTP position should be grammatical when extraposed, whereas the types that were good before should be bad here. This expectation turns out to be substantially correct, but with one interesting wrinkle: now null subject infinitives pattern with bare subject infinitives rather than with poss subject infinitives.

In fact, Lubukusu distinguishes at least two sorts of so-called extraposition. The first is a kind of impersonal passive, in which the verb bears a passive suffix (or a detransitivizing stative suffix) and its thematic complement remains after the verb, not moving to a preverbal position as either a subject or a topic. In this construction, the matrix verb bears the expletive agreement **ka-** (class 6) (or sometimes also **e-**). This sort of impersonal passive is not allowed when the verb takes a simple NP complement, as in (26).¹⁶

(26) ***Ka-bon-w-a** **e-twika** **ano** **likoloba**.

SM.c6.TNS-see-PASS-fv c9-giraffe here yesterday

‘A giraffe was seen here yesterday.’

It is also bad if the verb takes a **li-**nominal or a poss subject infinitive as its complement:

(27) a. ***K-eny-ebw-a** [**li-nyw-a** **ly-e** **ka-ma-lwa**].

SM.c6.TNS-want-PASS-fv c5-drink-fv c5-AM c6-c6-beer

‘It is wanted/needed the drinking of beer.’

- b. ?***Ka-khingilil-w-a** [khu-ly-a ka-ma-toore khw-a mayi].

SM.c6.TNS-prevent-PASS-fv c15-eat c6-c6-bananas c15-AM mother

‘It was prevented Mother’s eating bananas.’

However, the impersonal passive is possible with all the other clause types: with null subject infinitives and bare subject infinitives, as well as with all of the finite clauses.

- (28) a. **K-eny-ebw-a** [khu-nyw-a ka-ma-lwa].

SM.c6.TNS-want-PASS-fv c15-drink-fv c6-c6-beer

‘It is wanted to drink beer.’

- b. **K-eny-ebw-a** [Wekesa khu-som-a si-tabu].

SM.c6.TNS-want-PASS-fv Wekesa c15-read-fv c7-book

‘It is wanted for Wekesa to read the book.’

- c. **Ka-khalak-w-a** [ka-li/ba-li Wekesa a-nyw-e ka-ma-lwa].

SM.c6.TNS-decide-PASS-fv COMP Wekesa SM.c1-drink-SBJV c6-c6-beer

‘It was decided that Wekesa should drink beer.’

- d. **K-eny-ebw-a** [(Wekesa) a-nyw-e ka-ma-lwa].

SM.c6.TNS-want-PASS-fv Wekesa SM.c1-drink-SBJV c6-c6-beer

‘It’s wanted/needed (for Wekesa) to drink beer.’

- e. **Ka-khalak-w-a** Wafula a-a-b-a ne kumulukha.

SM.c6.TNS-decide-PASS-fv Wafula SM.c1-ACT-be-fv with party

‘It was decided that Wafula would have a party (and he did).’

The second type of extraposition found in Lubukusu involves psych verbs like ‘please’ and ‘bother’, which naturally take clause-like subjects—the sort of verbs that served as the matrix verb in the previous section. Here too the clause-like subject can appear after the verb, at the end of the clause, with no overt phrase in SpecTP. An interesting difference is that with this construction the default/expletive agreement on the matrix verb is class 5 **li-**, not class 6 **ka-**. Why Lubukusu has more than one expletive subject available, and why they distribute themselves in this way, is an intriguing question that we (unfortunately) have nothing to say about at this time. But what is significant for our current topic is that some clause-like constituents can extrapose in this way and some cannot. In particular, **li-**nominals and poss subject infinitives resist this extraposition, whereas they are fine in preverbal subject position.

(29) a. **??Ly-a-siim-isy-a Wekesa [li-ch-a ly-a Wafula].**

SM.c5-PST-please-fv Wekesa c5-go-fv c5-AM Wafula

‘It pleased Wekesa Wafula’s leaving.’

b. **?Ly-a-siim-isy-a Wekesa [khu-ch-a khw-a Wafula].**

SM.c5-PST-please-fv Wekesa c15-go-fv c15-AM Wafula

‘It pleased Wekesa Wafula’s leaving.’ (not as good as (30b))

In contrast, the other types of clause can undergo this sort of extraposition, including (slightly marginally) the null subject infinitive.¹⁷

(30) a. **(?)Ly-a-siim-isy-a Wekesa [khu-ch-a].**

SM.c5-PST-please-fv Wekesa c15-go-fv

‘It pleased Wekesa to go.’

b. **Ly-a-siim-isy-a Wekesa [Wafula khu-ch-a].**

SM.c5-PST-please-fv Wekesa Wafula c15-go-fv

‘It pleased Wekesa for Wafula to leave.’

c. **Ly-a-siim-isy-a Wekesa [ba-li/*a-li/*li-li Wafula a-a-ch-a].**

SM.c5-PST-please-fv Wekesa c2/*c1/*c5-COMP Wafula SM.c1-PST-go-fv

‘It pleased Wekesa that Wafula left.’

Lubukusu is not unique in this respect. English also allows impersonal passives and extraposition with finite CPs but not with simple nominals. Furthermore, Poss-**ing** gerunds in English are more like simple nominals than like infinitival clauses in this respect.

- (31) a. ***It is believed Mary.**
b. **It is widely believed that John will win the election.**
c. ***It is expected John’s winning the election.**
- (32) a. ***It bothers me Mary.**
b. **It bothers me that John will win the election.**
c. ***It bothers me John’s winning the election.**

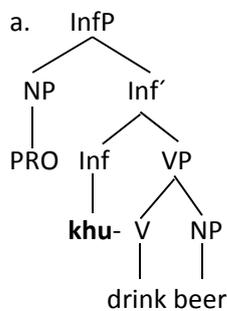
It appears, then, that the Lubukusu facts should be explained in terms of general principles.

Our rough-and-ready interpretation of this paradigm is as follows. We may assume that expletive pronouns in natural languages are possible, but they are used only as a kind of last resort, to fill the obligatory subject position (“satisfy the EPP”) when nothing else will do. So we might say that if the highest argument inside vP is a candidate for satisfying the EPP property of T, it blocks the insertion of an expletive pronoun in Spec, TP position¹⁸ (compare Chomsky (1991) on **do**-support as a last resort relative to verb movement, and Chomsky (1995) on Move blocking Merge). This makes expletive insertion impossible, not only with simple NPs, but also with clause-like constituents that have a null noun head, namely **li**-nominals and poss subject infinitives. If however the highest XP in the clause is not a perfect candidate for satisfying the EPP property of T, then merging an expletive pronoun in SpecTP is possible. Now we already

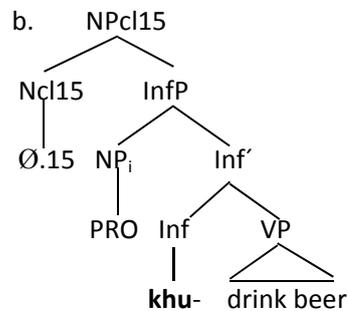
know that bare subject clauses and finite CPs in Lubukusu (not to mention C-less clauses) do not have phi-features, and thus cannot sit in SpecTP; hence they cannot satisfy the EPP property of T by themselves. Therefore, expletive insertion is possible in these configurations. This explains why the constituents that can sit in SpecTP disallow extraposition and the constituents that cannot sit in SpecTP allow it, accounting for most of the data above.

The one apparent anomaly is the null subject infinitive. This is the one construction type that is found both in true subject position and in extraposed position. However, there is a natural analysis of these as being structurally ambiguous. In the previous section, we said that null subject infinitives have an analysis in which an InfP (a nonfinite TP) headed by **khu-** is embedded under a null noun head of class 15 ((22c)), but no associative phrase modifier is adjoined to the NP to express the subject. That analysis emerges naturally as a variant of our analysis of poss subject infinitives, and it explains why null subject infinitives can be complements of P and specifiers of TP. Let us assume that the null subject when the AMP is missing is just PRO. But there is another, equally natural analysis of null subject infinitives in which they are variants of bare subject infinitives, in which the structure in (22b) happens to have a null subject. We therefore get a second analysis of null subject infinitives in which they are InfPs (TPs) without a nominal head or phi-features. The two structures are compared in (33).

(33) Structure 1 (=Bare-Inf)



Structure 2 (=Poss-INF)



The (33a) version of the null subject infinitive should not be possible in PP or in SpecTP, but it can be extraposed, as we see in (28a) and (30a). We conclude that the especially wide distribution of null subject infinitives is not a counterexample to the usually discrete distinction between structures that are headed by a noun and structures that are not. Rather it is a result of this particular structure happening to have two natural analyses, one as a nominal structure headed by a noun head and one that is a purely clausal structure headed by infinitival T.

6. Extraction from a clausal complement

The last major syntactic distinction that we discuss here before turning to issues of selection concerns the possibility of extracting the direct object out of a clause-like constituent in Lubukusu by way of clefting/focus movement. Simple extraction of the object out of the unambiguously nominal constituents is bad. More specifically, extraction out of a **li**-nominal is marginally possible if and only if an AM-marked resumptive pronoun is included, not otherwise, whereas extraction of the object out of a poss subject infinitive is sharply bad.

(34) a. ?**Ka-ma-lwa ni-ko ni-ko Wekesa eny-a [li-nyw-a**
 c6-c6-beer FOC-c6 FOC-c6 Wekesa SM.c1.TNS-want-fv c5-drink-fv
***(ly-a-ko)].**

c5-AM-pronoun.c6

‘It is beer that Wekesa wants the drinking of (it).’

b. ***Sii-na ni-syo Wekesa eny-a [khu-kul-a**
 c7-what FOC-c7 Wekesa SM.c1.TNS.want-fv c15-buy-fv
 ___ **khw-a Wafula]?**
 ___ c15-AM Wafula

‘What does Wekesa want Wafula’s buying?’

In contrast, extracting the object out of any of the unambiguously nonnominal clauses is perfectly good, including the indicative CP in (35b), the C-less subjunctive in (35c), and the actual in (35d). The successful extraction from the bare subject infinitive in (35a) forms a striking minimal pair with the poss Subject infinitive in (34b).

- (35) a. **Sii-na ni-syo Wekesa eny-a [Wafula khu-kul-a --]?**
c7-what FOC-c7 Wekesa SM.c1.TNS-want-fv Wafula c15-buy-fv
‘What does Wekesa want Wafula to buy?’
- b. **Sii-na ni-syo Wekesaa-khalak-a [a-li/ba-li a-kha-kul-e --]?**
c7-what FOC-c7 Wekesa SM.c1.TNS-decide-fv c1/c2-COMP SM.c1-FUT-buy-SBJV
‘What did Wekesa decide that he will buy?’
- c. **Sii-na ni-syo Wekesa a-khalak-a (bali) [Wafula a-kul-e --]?**
c7-what FOC-c7 Wekesa SM.c1.TNS-decide-fv COMP Wafula SM.c1-buy-SBJV
‘What did Wekesa decide for Wafula to buy?’
- d. **Sii-na ni-syo Wekesa eny-a [Wafula a-a-kul-a --]?**
c7-what FOC-c7 Wekesa SM.c1.TNS-want-fv Wafula SM.c1-ACT-buy-fv
‘What did Wekesa want Wafula to buy such that he did buy it?’

This clearly matches the distinction between nominal and nonnominal constituents that we have seen in previous sections. Here we simply observe that the following restriction holds quite strictly in Lubukusu (and many other Bantu languages).

(36) Generalization: There is no **wh**-extraction from NP/DP in Lubukusu.

Presumably this is ultimately a special case of the Complex NP Constraint (Ross, 1967), and should be deduced from a more general theory of island phenomena (with allowance for some

escape hatches from NP in some languages but not others), but we do not pursue that here. See also Carstens (1991:202), who discovers the same difference between poss subject construction, which block extraction, and bare subject constructions, which permit it, in Kiswahili.

What should we expect then for extraction from null subject infinitives? We expect that it should be possible when they are in complement position, because they can have a parse like (33a) in which they are no different from bare subject infinitives. Hence, extraction from them under this analysis should be possible, just as it is in (35a). This is correct, as shown by (37).

(37) **Sii-na ni-syo Wekesa eny-a [khu-kul-a --]?**

c7-what FOC-c7 Wekesa SM.c1.TNS.want-fv c15-buy-fv

‘It is beer that Wekesa wants to drink.’

Similarly, Bresnan and Mchombo (1995:223) show that it is possible to extract the object (by relativization) out of a null subject infinitive in Chichewa. (They do not, however, discuss whether the extraction becomes bad if there is a possessive subject with the infinitive.)

However, we might expect things to turn out differently if some other factor forces the null subject infinitive to have the structure in (33b), with a null noun head. Then extraction from the null subject infinitive should run afoul of (36), just as extraction from poss subject infinitives does. The crucial case would be when the null subject infinitive is the complement of a preposition like **khu**. As we saw in section 2, this forces a nominal head to be present to satisfy (11), the condition that the complement of a P in some languages must be nominal. But then extraction of the object from a null subject infinitive functioning as the object of a P should be degraded relative to (37), the null head that helps to satisfy (11) causing a violation of (36). This is correct:¹⁹

(38) **??Sii-na ni-syo Wekesa a-kachul-a khu-kul-a?**

c7-what FOC-c7 Wekesa SM.c1.TNS-talk-fv (about)-c15-buy-fv

‘What did Wekesa talk about buying?’ (improves with OM –**si-** as resumptive on INF)

This confirms our hypothesis that null subject infinitives are structurally ambiguous between an NP structure and an InfP structure. If they were simply intermediate cases on a continuum of nouniness, then we would not necessarily expect this interaction.

The data we have considered is now summarized in Table Two, which expands on Table One. There are still three major types of constituent, with the same distinctions being made by several different phenomena. The notable change to the overall layout is that null subject infinitives now have their own distinctive pattern. However, this is simply the union of the possibilities that exist for bare subject infinitives and those that exist for poss subject infinitives, reflecting the fact that null subject infinitives are structurally ambiguous. This completes our analysis of the structure and categorial status of the various clause types found in Lubukusu.

Table Two

	PP	Preverbal	True Subj	Nominal modifiers	Extraposed	Extraction from	Vs select for it (see below)
li-nominal	Yes	Yes	Yes	Yes	No	No	No
Poss-INF	Yes	Yes	Yes	Yes	No	No	No
Null-INF	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bare-INF	No	Yes	No (?)	No	Yes	Yes	Yes
Finite CP	No	Yes	No (*)	no	Yes	Yes	Yes
Subjunctive	No	No	No	no	Yes (pass)	Yes	Yes
Actual	No	No	No	no	Yes (pass)	Yes	Yes

7. Implications for a theory of selection

With this descriptive and analytical material in hand, we would like to explore in brief an implication that it has for the theory of selection.

In general, we say that a particular predicate selects for a given clause type when it puts requirements on what sort of clause it can take as its complement. Those requirements can be stated in formal-grammatical terms, or in semantic terms, or in both. For example, in English some verbs select only for an infinitival complement, some for an **-ing** complement, some for a finite CP complement, and some allow more than one of these possibilities (perhaps with shifts in the meaning of the selecting verb, perhaps not) (Rosenbaum, 1967). Although the phenomenon is familiar and ubiquitous, a principled theory of which predicates select for which kinds of complements and why is an underdeveloped area of linguistic theory. The theoretical gap is even more pronounced when one seeks a crosslinguistic theory, which is able to account for differences in selection across languages in terms of (i) differences in the meanings of the selecting

predicates, and (ii) differences in the stock of clause types that are allowed in a given language (for example, whether the language has a distinct subjunctive clause, whether it has a distinction between an infinitive and a gerund, whether it—like Mohawk—lacks any sort of nonfinite clause whatsoever, etc.). A principled and general theory of this type has rarely been attempted, much less achieved, beyond a few restricted subdomains.

Like other languages, Lubukusu has verbs that select for clauses of different types. So far we have tended to use verbs that take a relatively wide range of complements, so as to provide near minimal pairs. But other verbs are more choosy. In a preliminary survey of clausal complementation in Lubukusu that investigated some 75 verbs, we have identified the following verb classes, sorted as to what range of clausal complements they select:

- (39) a. Verbs that take null subject infinitive only: **khilwa** ‘fail’, **lekha** ‘avoid’, **loba** ‘resist’, **mala** ‘finish’, **chililila** ‘continue’...
- b. Verbs that take null subject infinitive or bare subject infinitive: **biyila** ‘hate’, ...
- c. Verbs that take null subject infinitive or actual clause only: **rakikha** ‘start’
- d. Verbs that take subjunctive or actual only: **khola** ‘make’, **lekha** ‘let’...
- e. Verbs that take null subj infinitive, bare subject infinitive, subjunctive, or actual: **enya** ‘want’, **khaka** ‘try’, **yingilila** ‘force’, ...
- f. Verbs that take indicative CP only: **amoonya** ‘mumble’, **subila** ‘believe’, **reeba** ‘ask’...
- g. Verbs that take indicative CP or null subject infinitive: **manya** ‘know’, **okesya** ‘show’, **iyikiicha** ‘regret’, **rya** ‘fear’, **khebulila** ‘remember’....
- h. Verbs that take all of the above: **boolela** ‘tell’, **nyala** ‘be able’, **siima** ‘like’, **khalaka** ‘decide’, **loma** ‘say’, **iyama** ‘accept, permit’, **panga** ‘plan’,

The details of this list are tentative, both because of the preliminary nature of our research and because we need more precise and principled answers to certain analytical questions about whether one should count as the same or different verbs that have the same root but different-seeming meanings (from the English perspective, anyway) and verb stems that are related to each other by adding causative and applicative affixes.²⁰ But already one can discern what seem to be intriguing patterns in this list, including many similarities with selection in English, a few systematic-looking differences, and perhaps a few idiosyncratic differences.

But in the context of our study of the structure and distribution of the various clause-like constituents, one very clear and non-subtle generalization stands out, as stated in (40).

(40) No verb in Lubukusu selects specifically for a “clause” that is headed by N.

This generalization covers two clear subcases. First, there is no verb in Lubukusu that selects a **li-**nominal as complement, to the exclusion of other types of nominals or clauses. **Li-**nominals are strikingly unlike the superficially not-that-different null subject infinitives in this respect. The two are similar in that the verb is “nominalized” by attaching an apparent noun class prefix (**li-** or **khu-**), the object is generally expressed, and the subject need not be. But many verbs select for null subject infinitives, and no verb selects for a **li-**nominal. (40) also covers the fact that many verbs select for bare subject infinitives, but none select for poss subject infinitives, even though the two are superficially similar in every way except for the expression of the subject. Our analysis reveals what these two unselected clause types have in common: both of them are headed by a null N, whereas bare subject infinitives are not and null subject infinitives need not be.

The observation in (40) also makes a more subtle prediction about null subject infinitives. We have learned that these are structurally ambiguous: they can have a structure in which they are only an InfP (TP), like bare subject infinitives, and they can have a structure in which they are an

InfP embedded under a null noun of class 15, like poss subject infinitives. If (40) is correct, then verbs that select for bare subject infinitives must be selecting for them in their InfP guise, never in their NP guise. This makes a testable prediction. If there were in fact verbs in Lubukusu that selected for bare subject infinitives in their NP guise only, these could be identified by the fact that extraction of an object from their complement would be degraded, given that NPs are islands to extraction in Lubukusu (see (36); compare extraction from the complement of P in (38)). Generalization (40) implies that there should be no such verb in Lubukusu; rather, all verbs that take null subject infinitives should permit the object of that infinitive to be extracted, as in (37). A first-pass investigation supports this prediction. We checked a sample of seventeen verbs that select null subject infinitives as their complements, chosen from various semantic classes (**iyama** ‘accept’, **manya** ‘know how to’, **boolela** ‘tell X’, **siima** ‘like’, **nyala** ‘to be able to’, **enya** ‘to want’, **ima** ‘to stop’, **loba** ‘to avoid’, **khaka** ‘to try’, **chililila** ‘to continue’, **rakikha** ‘to start’, **panga** ‘to plan’, **khilwa** ‘to fail’, **lekha** ‘to neglect’, **khalaka** ‘to decide’, **kumbuka** ‘remember’, **biyila** ‘to hate’). We tested to see if there was any notable variation in the acceptability of extracting the object **siima** ‘what’ of the nonfinite clause in such constructions. We found that the results were uniformly grammatical. This suggests that verbs always select the version of the null subject infinitive that is not embedded in an NP projection, just as (40) predicts.

In contrast, there are verbs that select for essentially every non-nominal clause type we have identified: for null subject infinitives only, for actuals and subjunctives only, for indicative CPs only, and there are verbs that select for more than one non-nominal clause type. So the split is a fairly clean one: anything that turns out to be a nominal complement in our investigation cannot be differentially selected (for nominal type), and almost anything that did not turn out to be an NP can be a differentially selected clause type.²¹

The claim that no verb selects differentially for a truly nominal clause in Lubukusu needs some clarification. It is not that **li**-nominals and poss subject infinitives are impossible as the complements of verbs in Lubukusu. On the contrary, we have seen that they are possible in this position. However, the verb in question always takes some other kind of clausal complement as well, and that other version is always considered clearly more “normal”, less marked semantically. In addition, the verb in question also allows an ordinary NP argument, not derived from a verb. (Indeed, most Lubukusu verbs seem to allow this.) The following examples give a sense of this. While (41b) is the “normal” way to describe this situation, (41a) is regarded as stilted and artificial –less natural than (41b), though “grammatically possible.”

(41) a. **?A-rakikh-a li-nyw-a ly-e ka-ma-lwa.**

SM.c1.TNS-start-fv SM.c5-drink-fv c5-AM c6-c6-beer

‘He started the drinking of beer.’

b. **A-rakikh-a khu-nyw-a ka-ma-lwa.**

SM.c1.TNS-start-fv SM.c15-drink-fv c6-c6-beer

‘He started drinking beer.’

c. **Wafula a-rakikh-a si-tabu.**

Wafula SM.c1.TNS-start-fv c7-book

‘Wafula started (to read) the book.’

The contrast in (42) is similar.

(42) a. **??A-khak-a li-nyw-a ly-e ka-ma-lwa.**

SM.c1.TNS-try-fv c5-drink-fv 5-AM c6-c6-beer

‘He tried the drinking of beer.’

b. **A-khak-a khu-nyw-a ka-ma-lwa.**

SM.c1.TNS-try-fv c15-drink-fv c6-c6-beer

‘He tried to drink beer.’

- c. **Wekesa a-khak-a ka-ma-lwa.**

Wekesa SM.c1.TNS-try-fv c6-c6-beer

‘Wafula tried (tasted) the beer.’

The examples in (43) and (44) show a similar contrast between a bare subject infinitive and a poss subject infinitive. The former is a perfectly normal, vanilla-flavored sentence; the latter is not impossible, but it is more marked, more unusual, used only for a special pragmatic effect.

- (43) a. **?/#Wekesa eny-a khu-mu-fumy-a khw-a Andrea.**

Wekesa SM.c1.TNS.want-fv c15-OM.c1-praise-fv c15-AM Andrea

‘Wekesa wants Andrea’s praising him.’

- b. **Wekesa eny-a Andrea khu-mu-fumy-a.**

Wekesa SM.c1.TNS.want-fv Andrea c15-OM.c1-praise-fv

‘Wekesa wants Andrea to praise him.’

- c. **Basaani b-eny-a ka-ma-lwa.**

c2-men SM.c2.TNS-want-fv c6-c6-beer

‘The men want beer.’

- (44) a. **#Wekesa a-khak-a khu-ly-a ka-ma-toore khw-e o-mw-aana w-ewe.**

Wekesa SM.c1.TNS-try -fv c15-eat-fv c6-c6-bananas c15-AM c1-c1-child c1-his

‘Wekesa tried his child’s eating bananas’ (“The structure is OK but has no meaning”)

- b. **Wekesa a-khak-a o-mw-aana w-ewe khu-ly-a ka-ma-toore.**

Wekesa SM.c1.TNS-try-fv c1-c1-child c1-his c15-eat-fv c6-c6-bananas

‘Wekesa tried for his child to eat bananas’ (normal)

As far as we can tell, the preference never goes the other way around: neither the **li**-nominal nor the poss subject infinitive is ever the most preferred, most neutral complement of the verb that takes a complement describing an event, but it can serve to describe an event if a nominal complement is natural with that verb at all. We interpret this as meaning that verbs looking to combine with a clausal constituent of some type can never explicitly select for a particular type of constituent headed by N (or D) to satisfy this need, but if the verb can also combine with a simple NP/DP, then event-denoting NP/DPs such as **li**-nominals and poss subject infinitives can pass as alternatives to non-nominal clausal complements.²²

Indeed, English seems to be essentially the same as Lubukusu in this respect. English has so-called action nominals and gerunds, both of which use the suffix **-ing**. But when the two are used as complements they differ both in their degree of acceptability and in subtle semantic ways.

- (45) a. ?#Alice saw Bill's drinking of beer on the wharf. (action nominal)
 b. Alice saw Bill drinking beer on the wharf. (gerund)
 c. Alice saw the event.

While (45a) can be taken to mean what roughly what (45b) does, (45a) is very stilted and sounds like the viewing of a ceremonial event rather than an ordinary episodic one. Moreover, the complement in (45a) seems to be possible whenever an NP like **the event** is, whereas whether the complement in (45b) is possible varies with the matrix verb, as shown in (46).

- (46) a. ?#Alice attended Bill's drinking of beer on the wharf
 b. *Alice attended Bill drinking beer on the wharf
 c. Alice attended the event.

Although (46a) is still stilted, it can be understood in the same ceremonial way that (45a) is, whereas the Acc-**ing** gerund that is possible as the complement of **see** in English ((45b)) is impossible as the complement of **attend** ((46b)). The Lubukusu facts have the same quality.

Is (40) then a hint as to what a more general theory of selection should be like? We suspect the answer is “yes”, that it points to a distinct role for a syntactic mechanism that conditions the result of semantic selection. While we take it that a semantic role in selection is clearly necessary, a syntactic role in selection does not seem conceptually necessary, though some have argued that it is empirically necessary (e.g., Grimshaw 1979, Pesetsky 1982, and more recently Landau 2004). But the difference in gross semantics between (say) a poss subject infinitive and a null subject infinitive appears to be very slight, if any; yet one of them can be specially selected for, and the other cannot be. Hence, the syntax must be playing a significant role in selection. Even if one argues that there are type-theoretic differences between one kind of event and another, the mapping between semantic type and syntactic form remains non-trivial, as the examples discussed here show.

It appears, then, that many non-nominal clausal complements are more susceptible to outside selectional influence on their internal modal/temporal structure than nominals are. But it is also notable that indicative CPs differ somewhat in this respect from the other kinds of clausal complements. If a matrix verb in English selects an indicative CP, then it is compatible with any tense value at all within that clause. For example, one might think that **remember** would be semantically biased toward events in the past, since one cannot remember things that will happen in the future. This bias does show up clearly when the complement of **remember** is a nonfinite verb with **-ing**, as in (47a). Nevertheless, when **remember** appears with an indicative CP, that CP can be in future tense as well as in past tense, as shown in (47b,c).

- (47) a. John remembers his team winning the league championship last year/#next year.
 b. John remembers that Mary went to London last year.
 c. John remembers that Mary will go to London next year.

Conversely, one might expect a verb like **anticipate** to be semantically biased toward events in the future given its lexical meaning. Again this bias shows up when it appears with a nonfinite – **ing** complement, but not when it appears with a finite CP.

- (48) a. John anticipates Mary going to London next year/#last year.
 b. John anticipates that Mary will go to London next year.
 c. John anticipates that Mary went to London last year.

Although slightly unusual, (48c) seems fine with a sort of epistemic meaning, in which John does not know for sure whether Mary went to London on company business in the previous year, but he knows that the company records have recorded this information, and he predicts that when he looks into the matter he will confirm that she did in fact go. In contrast, (48a) with the past adverb strongly resists this meaning. In short, then, it appears that when a verb selects for a finite indicative C then V cannot put any strong requirements on the TP complement of that C.

Essentially the same thing is true in Lubukusu. Recall that this language has verbs that select for infinitival, subjunctive, or actual clauses. Here selection, whether syntactic or semantic, seems to be for different values of a T node (infinitive **khu-**, actual **a-** or subjunctive **-e**) in a TP that is not embedded in a CP.²³ Lubukusu also has some predicates select for indicative CPs exclusively, but for those predicates any finite value of TP is possible inside the CP, as in English: the T could be present, any of several degrees of past, or any of the several futures—any kind of T that is allowed in an unselected matrix clause. Perhaps the most striking near-minimal pair in this respect is subjunctive and future. Both of these morphemes express future time relative to some

reference point as part of their lexical meaning.²⁴ But there are two clear differences. First of all, subjunctive clauses can appear without a C, as seen in several examples above, whereas future clauses cannot:

- (49) **Ba-booli ba-bool-a *(ba-li) o-mwaami a-kh-eny-e**
 c2-reporters SM.c2.TNS-say-fv c2-COMP c1-leader SM.c1-FUT-want-SBJV
bu-ng’ali bw-ewe.
 c14-right c14-his

‘The reporters say that the president will claim his rights.’

Second, there are verbs that require their complement to be subjunctive (or actual), but there are no verbs in Lubukusu that require their complement to be future. Thus, the sorts of verbs that like subjunctive clauses cannot be satisfied with a future clause instead, nor is there another class of verbs that requires future the way that (say) **khola** ‘make’ requires subjunctives. For example, the verb **ikomba** ‘to wish’ is possible with either future or subjunctive, with similar meanings, whereas the verb **enyā** ‘to want’ is possible only with subjunctive, not with future. But there is no similar verb that is acceptable with a future complement but not a subjunctive one.

- (50) a. **Wekesa e-komb-a o-mw-aana a-b-e o-mu-lamu.**
 Wekesa SM.c1.TNS-wish-fv c1-c1-child SM.c1-be-SBJV c1-c1-healthy.

Wekesa wished for the child to be healthy.’

- b. **Wekesa e-komb-a bali o-mw-aana a-kha-b-e**
 Wekesa SM.c1.TNS-wish-fv COMP c1-c1-child SM.c1-FUT-be-SBJV
o-mu-lamu.

c1-c1-healthy.

Wekesa wished that the child will be healthy.’

being unable to see past a C node either. These cases thus strongly suggest that selection is strictly local, in the syntactic sense suggested in (52).

Although we consider this range of evidence to be strong, and (52) is relatively uncontroversial, we must acknowledge that it is not so trivial and innocuous as to be without potentially embarrassing problems. For example, in English there are some verbs that allow **for-to** complements and others that only allow infinitival complements without **for**.

(53) a. **John hoped (for us) to win.**

b. ***John tried (*for us) to win.**

If English **for** counts as the head of CP, as has been the standard analysis, then it ought to free its complement from selection, making possible any sort of TP complement inside CP. This is obviously false (***John hoped for us/we will win**). An obvious suggestion to make here is that the verb selects for its complement CP (headed by **for** or by \emptyset) and then the complementizer selects for the TP type, here nonfinite, headed by **to**. Although this sort of chained selection is technically compatible with (52), it threatens to empty it of empirical content, since any apparently long distance selection relationship could potentially be reanalyzed as a sequence of local selection relationships in this way, emptying (49) of its force. Similar cases may arise in Lubukusu. While it is true that subjunctive clauses are possible without an overt C, unlike indicative clauses, subjunctive clauses can appear with a C as well, as in (54).

(54) **Wekesa a-pang-a (bali) Wafula a-b-e ne kumulukha.**

Wekesa SM.c1.TNS-plan -fv COMP Wafula SM.c1-be-SBJV with party.

‘Wekesa planned for Wekesa to have a party.’

Other verbs that allow both options include **khaka** ‘to try’, **ikomba** ‘to wish’, **yingililia** ‘to force’.

It is not clear (to us) whether there is a semantic difference here or not, whether these are more

like C-less subjunctives that happen to have a C, or whether they are like finite CPs, where one of the choices freely available for unselected TP is subjunctive. If the former is the case, then we have the problem that the presence of C in (54) apparently does not prevent the matrix V from selecting for the subjunctive T/Mood.

We have no solution to offer here to the issue of chained selection raised by (53) and (54). One may certainly hope that a fuller understanding of the semantic aspects of selection will help to explain why (53) and (54) are possible but other apparently long distance selection is not. For now, we simply want to emphasize that we have shown that a solution needs to be found, because we have shown clearly that, apart from these special cases, there *are* strong and substantive structural limits to what can select what. It is a clear result of our study that having a nominal head prevents a clause-like constituent from being selected in a distinctively clause-like way. Such a constituent might be possible as an NP object, but it cannot show any of the special selectional properties associated with infinitives, subjunctives, and other special forms of clausal complementation in languages of the world. This was already included as a property distinguishing the clause types back in Table Two, closely correlated with the other differences we studied in earlier sections. This in turn strongly suggests that the syntactic contribution to selection must be strictly local in terms of tree geometry, as expressed in (52). We hope then that this study will function as a sort of prelude to a more general and principled theory of selection, capable of explaining in principle when selection is and is not possible, and accounting for the similarities and differences between (say) Lubukusu and English, or other pairs of languages.

8. General Conclusion

In this paper, we have explored the borderline between clausal and nominal syntax in Lubukusu. The result is a sharp distinction, not a graded difference, which is evident in the external distribution of the constituent types we have investigated. Constituents that are headed by a null noun—including **li**-nominals and poss subject infinitives—can be the complements of Ps and can be in the true subject position (SpecTP), but they cannot extrapose to the right edge of the clause and NPs cannot be extracted from them. In contrast, constituents that are not headed by a noun—bare subject infinitives, indicative CPs, subjunctive clauses, and actual clauses—cannot be in PP or in true subject position (SpecTP) but they can be extraposed and can have an NP extracted from them. We thus have a clear distinction between categories that are nominal for their external syntax and ones that are not—one that separates some class 15 infinitives from others.

An apparent exception to this binary distinction is the null subject infinitive, but that turns out simply to be structurally ambiguous: it has one parse in which it occurs in a null headed NP and behaves just like a poss subject infinitive, and another parse in which it is not embedded in NP and behaves just like a bare subject infinitive. In this way, we sort out the various types of clause-like constituents into discrete types in Lubukusu, each with consistent and theoretically tractable properties.

Finally, we use these results as an opportunity to reflect on the nature of selection between a verb and its complement. In particular, we observe that verbs in Lubukusu can exercise special selection for any of the non-nominal clause types, but none of them show special selection for any of the nominal clause types. We conclude from this that selection is strictly local in a structural syntactic sense, such that even a phonologically null noun head with little obvious semantic value disrupts it. This also seems to be true for indicative complementizers, which prevent the matrix verb from selecting particular values of Tense in its complement, whereas such selection is

possible in reduced clauses that have no obligatory complementizer. We offer this as a starting point for a deeper and more principled theory of selection that is crosslinguistically valid.

References

- Alexiadou, Artemis, Haegeman, Liliane, and Stavrou, Melita. 2007. *Noun phrase in the generative perspective*. Berlin: Mouton de Gruyter.
- Baker, Mark. 2003. Agreement, dislocation, and partial configurationality. In *Formal approaches to function in grammar*, eds. Andrew Carnie, Heidi Harley and Mary Ann Willie, 107-134. Amsterdam: John Benjamins.
- Baker, Mark. 2005. On gerunds and the theory of categories. Ms.
- Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Bresnan, Joan, and Mchombo, Sam. 1995. The lexical integrity principle: evidence from Bantu. *Natural Language and Linguistic Theory* 13:181-254.
- Bresnan, Joan. 1997. Mixed categories as head sharing constructions. Paper presented at *Proceedings of the LFG97 Conference*, Stanford University.
- Carstens, Vicki. 1991. The Morphology and Syntax of Determiner Phrases in Kiswahili, UCLA.
- Carstens, Vicki. 2005. Agree and EPP in Bantu. *Natural Language and Linguistic Theory* 23:219-279.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1991. *Some notes on economy of derivation and representation: Principles and parameters in comparative grammar*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Collins, Chris. 2002. Eliminating labels. In *Derivation and explanation in the minimalist program*, eds. Samuel Epstein and T. Daniel Seely, 42-64. Oxford: Blackwell.
- Diercks, Michael. 2010. Agreement with subjects in Lubukusu, Georgetown University: Ph.D. dissertation.
- Diercks, Michael. 2012. Parameterizing Case: evidence from Bantu. *Syntax* 15.

- Diercks, Michael. to appear. Indirect Agree in Lubukusu complementizer agreement. *Natural Language and Linguistic Theory*.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10:279-326.
- Johnson, Kyle. 1988. Clausal gerunds, the ECP, and government. *Linguistic Inquiry* 1988:583-610.
- Julien, Marit. 2002. *Syntactic heads and word formation*. New York: Oxford University Press.
- Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.
- Koster, Jan. 1978. Why sentential subjects don't exist. In *Recent Transformational Studies in European Languages*, ed. S.J. Keyser. Cambridge, MA: MIT Press.
- Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22:811-877.
- Lefebvre, Claire, and Muysken, Pieter. 1988. *Mixed categories: nominalizations in Quechua*. Dordrecht: Kluwer.
- Malchukov, Andrej. 2004. *Nominalization / verbalization: constraining a typology of transcategorial operations*. Munich: LINCOM Europa.
- Malouf, Robert. 2000. *Mixed categories in the hierarchical lexicon*. Stanford, Calif.: CSLI publications.
- McCloskey, James. 1991. There, it, and agreement. *Linguistic Inquiry* 22:563-567.
- McFadden, Thomas. 2004. The position of morphological case in the derivation, University of Pennsylvania: Ph.D. dissertation.
- Mchombo, Sam. 2004. *The syntax of Chichewa*. Cambridge: Cambridge University Press.
- Mugane, John. 2003a. Infinitive-gerunds as extrasequential hybrid constructions in Logooli.
- Mugane, John. 2003b. Hybrid constructions in Gikuyu: Agentive nominalizations and infinitive-gerund constructions. In *Nominals: inside and out*, eds. Miriam Butt and Tracy Holloway King, 235-266. Stanford, Calif.: Center for the Study of Language and Information.
- Myers, Scott. 1987. Tone and the Structure of Words in Shona, The University of Massachusetts-Amherst.
- Pesetsky, David. 1982. Paths and categories, MIT: Ph.D dissertation.
- Reuland, Eric. 1983. Governing -ing. *Linguistic Inquiry* 14:101-136.

- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281-338. Dordrecht: Kluwer.
- Rosenbaum, Peter. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.
- Ross, John. 1967. Constraints on variables in syntax, MIT.
- Ross, John. 1973. Nouniness. In *Three dimensions of linguistic theory*, ed. O. Fujimura, 137-258. Tokyo: TEC Company.
- Safir, Ken, Baker, Mark, and Sikuku, Justine. 2012. Actual clauses in Lubukusu and the theory of clausal complementation. Paper presented at *World Congress of African Linguistics 7*, Buea University, Cameroon.
- Stowell, Timothy. 1981. Origins of phrase structure, MIT: Ph.D. dissertation.
- Wechsler, Stephen, and Zlatić, Larisa. 2003. *The Many Faces of Agreement*. Stanford, Calif.: CLSI.

Notes:

¹ Abbreviations used in the glosses include: ACT, actual mood; AM, associative marker; COMP, complementizer; CAUS, causative; cX, noun class X (X ranging from 1 to 20); FOC, focus; FUT, future; fv, final vowel; OM, object marker; PASS, passive; PST, past; SBJV, subjunctive; SM, subject marker; STAT, stative; TNS, tense.

² Carstens (1991) calls the Kiswahili version the infinitive. Her reason for calling this construction an infinitive whereas she calls the other **ku**+verb constructions gerunds has to do with a putative correlation with tense interpretation. She claims that **ku**+verb constructions in Kiswahili that have a null subject also have an unrealized future interpretation with respect to the matrix verb, whereas **ku**+verb constructions with a bare subject are truly tenseless, like gerunds in English. However, her primary examples of this correlation concern complements of the

predicate ‘remember’. We believe that when one looks at a wider range of matrix predicates (in Lubukusu, anyway), the temporal properties of the nonfinite clause are more complex, and a correlation with having an overt subject breaks down. For example, the future complement of ‘want’ in Lubukusu can have an overt subject (unlike what Carstens reports for Kiswahili), whereas verbs like ‘fail’ and ‘avoid’ occur with null subject infinitives but there is nothing future or semantically tensed about their complements.

³ There are some interesting semantic differences between these three forms of C having to do with factivity and evidentiality—whether the speaker is committed to the truth of the complement or not. In addition, the three do not have the same distribution in all respects; for example, only **bali** can be used in sentential subjects. See Diercks 2010 for some discussion. We do not investigate these semantic distinctions or the finer-grained distributional differences here, but hope to return to them in future work.

⁴ In this respect, actual clauses are different from factive clauses, the presupposed truth of which survives under negation. In contrast, negation of the matrix verb is incompatible with the complement being an actual, since the event is negated, and therefore did not take place.

⁵ The P **khu** in Lubukusu also has the property that, if a particular noun has a double class marker (e.g. **ka-ma-lwa** ‘c6-c6-beer’), the first of those markers is absent after **khu** (e.g. **khu ma-lwa** ‘about c6-beer’). Hence, an alternative analysis of (9b,c) would be that the single ch15 prefix **khu-** counts as an outer prefix, rather than an inner one, and thus does not appear after the P **khu** for this reason. Note that its presence in (9a) shows that the **li-** prefix in a **li-**nominal counts as an inner affix in these terms.

⁶ Note that (10a) is not very good regardless of whether an overt **khu** precedes the subject or not. Which form one expects here would depend on the details of the haplology rule: whether the

overt subject counts as destroying the contiguity between the preposition and the homophonous verb prefix or not. The details do not matter much, since any plausible version is bad.

⁷ We also leave partly open just what factors go into making a clause “nominal.” For example, it may help if there is something in C⁰ or SpecCP that has nominal features that can be inherited by the CP as a whole; see, for example, Ross 1973 on the presence of a nominal **wh**-phrase in SpecCP allowing that CP to be the complement of a P in English. In this respect, the fact that Lubukusu’s complementizers **bali** and AGR+**li** derive from a verb (see note 9), not from anything nominal in nature, may play a role in explaining why CP cannot be a complement of P in this language.

⁸ Carstens (1991:194-195) however claims that when an bare subject construction does appear in subject position in Kiswahili, it always triggers class 15 agreement on the main verb (**ku-**), unlike what we see in Lubukusu. The difference may not be unexpected in that class 5 can be used as an expletive form of agreement independently in Lubukusu (see section 4), but not in Swahili.

Rather, Carstens points out that class 17 **ku-** is a form of expletive agreement in Swahili, and this is homophonous with class 15 agreement in that language. This homophony might conceal the possibility that true subject-verb agreement can fail with a bare subject construction in Swahili too (compare Carstens 1991:206).

Carstens (1991:195) does show that poss subject constructions and true nominalizations can trigger c15 *object* agreement in Kiswahili, whereas bare subject constructions cannot, a similar contrast. We cannot replicate this result for Lubukusu for two reasons. First, OMs in Lubukusu do not have as many agreement properties as OMs in Swahili do, and may be incorporated pronouns instead (Diercks and Sikuku in progress); second (and perhaps related),

there is no class 15 OM in Lubukusu that one can try using in this way. (Gĩkũyũ is like Lubukusu in this respect, according to Mugane 2003b:250)

⁹ Note that the complementizer **bali** might be decomposed into **ba+li**, in which case it does contain a morpheme **-li**, but (if anything) this is presumably related to the present tense copula **-li** found in Lubukusu. The homophony with class 5 prefixes is purely accidental.

¹⁰ This may be an incompleteness in our account, to be addressed by a finer analysis of the left periphery than we have attempted. But it may also be that the surface similarity of bare subject infinitives to poss subject infinitives clouds the judgments for processing reasons. Something similar happens with some speakers of American English, who allow Poss **Ing** gerunds in Spec,TP (see (19b)) but offer variable and uncertain judgments for Acc-**Ing** gerunds in the same position. Reuland (1983), for instance, gives an example analogous to (ii) as being bad, but the first two authors of this article find it only awkward.

(i) (?)**John dropping his plate was unfortunate**

(ii) ?**Was John dropping his plate unfortunate?**

¹¹Bresnan and Mchombo (1995:235) present one argument that even the non-nominal version of a **ku+verb** infinitive has class features in Chichewa: it can serve as an antecedent for a class 15 pronominal object marker (OM) on a subsequent verb. This argument is not replicable for Lubukusu, because it does not have a class 15 OM in its inflectional inventory (see note 8). Bresnan and Mchombo's argument is also not iron clad, in that pronouns can show semantic agreement rather than grammatical-gender agreement with their antecedents in discourse in many languages. For example, in a language in which the diminutive of 'girl' is grammatically neuter, its referent can still be referred to in the next clause by a feminine pronoun (Wechsler and Zlatic

2003:13). Therefore, one cannot always infer the gender of a phrase from that of a pronoun that seems to refer back to it, as Bresnan and Mchombo propose to do.

¹² In Kiswahili, Logooli and Gĩkũyũ, where class 15 is used for derived nominals as well as for the various types of infinitive/gerund, we would still posit the structure in (21a), but the gender borne by the nominalizer that attaches to verbs directly would be class 15, just as is the nominalizer that selects a VP/InfP complement in (22c). As a result, the prefix that manifests on V is *ku-* in both (21a) and (21c) in those languages.

¹³ But see McFadden 2004 and others for a similar claim about English, based in part on the existence of Acc **Ing** gerunds in seemingly free variation with PRO **Ing** gerunds. The fact that bare subject infinitives are apparently bad in Gĩkũyũ (Mugane 2003b) might suggest that the Case filter does apply in that, fairly closely related language. One would of course want to see other effects of such a significant grammatical difference as that, however.

¹⁴ It is not, however, possible for adjectives to adjoin to gerund-like constructions in all languages. For example, Carstens (1991) shows that an adjective bearing class 15 morphology can adjoin to a derived nominal, but not to the poss subject construction or the bare subject construction in Swahili. This restriction is also found in English, where demonstrative plus gerund is (marginally) possible (**All this criticizing Mary must stop!**) but having an adjective modify a gerund is not (***Constant criticizing Mary must stop!**; compare **The constant criticizing of Mary must stop!**). It is not clear (to us) why adjectives can modify NPs of this sort in some languages but not others. See Alexiadou et al. (2007) for some proposals that address this by saying that gerunds do not have an actual N head, although they do have the higher functional superstructure of a DP.

¹⁵ Here are some more specific comparative notes for the detail-minded. Our (22c) is similar to Myers' (1987) structure, except that he does not distinguish the infinitival head **ku-** from the

nominalizing head, but has a single head that both hosts **ku-** and heads an NP. Carstens' (1991) versions of (22a) and (22b) are almost identical to ours, except that what we call Inf (T) she calls M. Carstens' version of (22c) has two functional heads, D and #, where we have only the lexical head N. Carstens also posits additional T and C nodes above InfP (her MP) in null subject infinitives, because of her view that infinitives with null subjects have a different tense value from gerunds with bare or possessive subjects. Since we do not think her view on this holds up in general (see note 2), we do not enrich the structure with these covert categories.

(22c) is also very much like Bresnan and Mchombo's (1995:234-238) proposal, except that they have no null N head of the NP, following a theoretical/conceptual preference to avoid null heads at the cost of violating the endocentricity of phrase structure; see also Mugane (2003b) on Gikūyū, who follows them in this. Another theoretical option for capturing approximately the same idea is proposed by Bresnan (1997), who says that a single complex word like **khu**+drink counts simultaneously as the (semantic, f-structure) head of both the VP and the c15 NP in a structure like (22c)—a view adopted by Mugane (2003a). We take the choice between exocentricity, null heads, and shared heads to be primarily a theoretical one, and do not emphasize the difference here.

More seriously different from our analysis—and less adequate in our view—would be a claim that a structure like (22c) has only a single head, but it is categorically ambiguous, having features of both V and N simultaneously, as in Lefebvre and Muysken (1988) (for Quechua) and Malouf (2000) (for English). See Bresnan (1997) and Baker (2005) for some criticism of this.

¹⁶ Examples like this can be much improved, perhaps to full acceptability, if T agrees with the postverbal NP, rather than bearing expletive *ka-* agreement; so (26) is ? with a class 9 prefix **y-a-bon-w-a** (and the stative version **y-a-bon-ekh-a** is even better). This raises questions for Baker's

(2008) and Dierck's (2010) hypothesis that agreement is always upward with SpecTP in Bantu/Lubukusu, but is not directly relevant here. For some reason, examples like (27b) (and also (28a,b)) are ruled out even with c15 agreement **khu-**, which one would think should be possible for poss subject and bare subject infinitives, given that they are simply NPs. We don't know why.

¹⁷ We have not found any cases of subjunctive or actual clauses following this pattern.

Presumably this is related to the fact that those clause types are also bad as preverbal "subjects", as shown in (25). This fact might help refine our understanding of why these clause types cannot be sentential subjects: it may imply that, in addition to their not being able to occupy (or bind) the SpecTP position, they cannot even receive an external theta role in (say) SpecvP.

¹⁸ Locative expletives like **there** in English used in presentational sentences are presumably a bit different in these respects, and similarly for Lubukusu. We do not pursue the differences here.

¹⁹ Note however that (38) is not as bad as the extraction from the poss subject infinitive in (34b).

We are not entirely sure why there is this difference. One possibility is that the poss subject Infinitive violates another condition in addition to (36)—for example the ban on extracting out of a specific nominal, if the definite NP in the AM phrase renders the NP as a whole specific, as possessors in English do. Another possibility is the surface similarity of the various kinds of infinitive clouds the judgments somewhat, for performance reasons (compare note 10). We extracted a **wh**-phrase in these examples rather than a referential noun phrase to make it less desirable to have an OM on the embedded verb functioning as resumptive pronoun. This would otherwise be preferred in (38) and it washes out the effect.

Johnson (1988) observes that extraction out of gerunds that are complements to argumental PPs is perfectly grammatical in English (**Who did you count on seeing at the**

party?). This difference is not unexpected given that (36) is true of Lubukusu (and other Bantu languages), but not of English, which generally does allow extraction out of nonspecific NPs.

²⁰ Another need is to learn better how to distinguish arguments from adjuncts, given that infinitives (both null subject and bare subject types), subjunctive clauses, and actual clauses can all be used as what seem to be purposive adjunct clauses, as well as true complement clauses.

²¹ There even seem to be a few verbs that select a bare subject infinitive but not a null subject infinitive, such as **buula** ‘reveal’ and **bukula** ‘consider’. To what extent this effect is syntactic or semantic in nature is an interesting question that we must leave for future research.

²² Interestingly, Mugane (2003b:250-251) says that infinitives with nominal modifiers are not merely marked but fully ungrammatical as complements of verbs like ‘want’ and ‘like’ in Gĩkũyũ. It would be interesting to know if these verbs can take simple NP objects in this language. If not, then this result is predictable.

²³ The assumption that **khu-** and **a-**, are T heads does not seem particularly controversial, although other views could be considered. For subjunctive **-e**, the matter might be more controversial. Some might want to see this as a head of a different functional category (MoodP, perhaps), lower than T in clause structure. This might enable one to explain in these terms why subjunctive is a suffix and the others are prefixes, by saying that the verb in Lubukusu moves up to Mood but not as high as T, along the lines of Julien 2002. Now we do not think that a correlation between affix position (prefix versus suffix) and the height/scope of the affix holds up in general in Lubukusu (Lubukusu does have other T-like suffixes, for example, including recent past *-ile*), and the strong distributional similarities between subjunctive clauses and actual clauses might be captured more easily if one says that actual and subjunctive are two instances of the same category. But we do not need to insist on this: if subjunctive clauses are really MoodPs that are not embedded in a

larger TP or CP, then the value of Mood can still be selected syntactically by the matrix verb, in accordance with the view we are building to in (52).

²⁴ We do not claim that the meanings are identical, only that they are similar. For instance, it could be that subjunctives are future with respect to the time of the matrix verb (relative future), whereas futures are future with respect to the time of the utterance (absolute future).