Non-verbal clauses in Kirundi: Focus and non-verbal predication*

Terrance Gatchalian
2nd PhD Evaluation, McGill University

rev. 29 March 2023

Abstract

Abstract here.

Contents

* Acknowlegements

1 Introduction

One strategy for expressing focus in Kirundi (Bantu, J/D62 spoken in primarily in Burundi) is an $\bar{A}$-fronting construction, obligatorily accompanied by the sentence-initial particle *ni* illustrated in (1b)[1] Anticipating the discussion below, I will refer to these constructions as "clefts" and the derivational process resulting in them as "fronting". I will elaborate on these terms more precisely below.

(1) a. Yohaáni a-a-som-ye igitabu
Yohani 1SM-PST-read-PFV 7.book
'Yohani read a book.' (neutral sentence)

b. Ni **igitabu** Yohaáni a-a-som-yé
NI 7book Yohani 1SM-PST-read-PFV.REL
'It's the book that Yohani read.' (*ni*-accompanied fronting)

Across Bantu and African languages more broadly, there have been various proposals put forward for similar constructions, with analyses falling into two distinct families. One family of analyses analyzes the accompanying particle as a focus marker, typically understood as instantiating a left peripheral Foc head (Rizzi 1997; see Aboh 2016). While this highly influential account has been widely adopted (for example, see Abels and Muriungi 2008, Kĩîtharaka; Hartmann and Zimmermann 2012, Bura (Chadic); Green 2007, Hausa (Chadic), among others), the particular properties of the Bantu constructions, of which Kirundi is a fairly typical exponent, poses a persistent challenge for such an analysis. In particular, the word order of the hypothesized focus head and the fronted constituent have led to non-trivial challenges for adopting the Rizzian proposal (see Schwarz 2003; Yuan 2017a for discussion and a possible solution). This paper defends an alternative view, where constructions like (1b) are analyzed as biclausal clefts (Zentz 2016b).

More specifically, I propose that these structures consist of a syntactically highly reduced and semantically expletive matrix clause. I argue that Kirundi clefts differ sharply from English clefts in that the matrix clause is syntactically non-verbal, headed by a non-verbal predicator *ni*, rather than a verbal copula like English BE. That is, while the Kirundi cleft structure is bi-clausal, they are mono-verbal. I spell out this hypothesis in crucial contrast to the widespread, and often implicit, structural assumption that (cognates of) *ni* is part of the verbal extended projection (for e.g., Zentz 2016b; **?**). The argument rests on taking seriously the formal similarities between the particle in clefts and the copular element in non-verbal predication (see §4 for an explicit definition of the copula as a non-verbal functional item).

As such, I have two interrelated goals in this paper. Firstly, I present an analysis of the *ni* accompanied fronting constructions in Kirundi as an $\bar{A}$-movement construction, supported with novel elicitated data. Second, I spell out the cleft analysis and motivate a non-verbal view of the matrix predicate. In (2), I briefly sketch the particular claims to be established under a cleft analysis, and describe the structural proposal to be defended.

(2) **Claims to establish**

a. The immediately post-*ni* constituent is derived by $\bar{A}$-movement
(For similar analyses in Bantu, see e.g. Schneider-Zioga 2007; Abels and Muriungi 2008; Zentz 2016b: p. 182ff; beyond Bantu, see e.g. Torrence 2013a,b; Klecha and Martinović 2015; Martinović 2021b *inter alia*)

b. The post-*ni* constituent is interpreted as the exhaustive listing satisfying the predicate in the remnant

[1]Glossing abbreviations here. Additional information about the language?

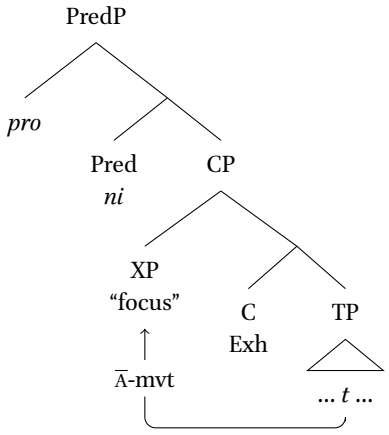
(For exhaustive identification, see Horvath 2005, 2007, 2013; Green 2007; Hartmann and Zimmermann 2012; Klecha and Martinović 2015; Fominyam and Šimík 2017)

c. Kirundi *ni* is a reflex of non-verbal Pred^0
(For similar analyses, see Wasike 2007; Diercks 2010: 193ff)

d. Kirundi *ni*-initial constructions are bi-clausal but mono-verbal, with a defective/expletive *non-verbal* matrix clause
(On clefts generally, see e.g. É. Kiss 1998; Rochemont 1986; Chomsky 1971; Jackendoff 1972; Akmajian 1970; Hedberg 2000; for Bantu-specific analyses, see e.g. Wasike 2007; Diercks 2010; Zentz 2016a,b *inter alia*)

Together, these claims converge on the analysis illustrated in (3) for the sentence given in (1b). The embedded CP is lexically specified to be embedded. The root clause consists of a non-verbal predicator, which I will argue is the same element that appears in nominal and adjectival predication. I argue further using the two contexts in which *ni* occurs, clefts and non-verbal predication, that Pred is distinct from the verbal predicator *v* in that the former obligatorily lacks verbal functional projections such as tense inflection.

(3) $\bar{\text{A}}$-derived cleft construction



In a local sense, this proposal provides a new look at *ni*-constructions in Bantu, and argues for the bi-clausality of these constructions in Kirundi. More broadly, this paper proposes and develops a new structural analysis for clefts, one where the matrix clause is highly deficient. That is, it defines a new way for cleft structures to be bi-clausal: while English clefts consist of two clauses containing verbs (the main predicate and *be*), Kirundi clefts consist of two clauses, only one of which contains a verb. Finally, it defines the properties of a syntactically non-verbal copula (called "particle copulas" in Pustet 2003), which has been often neglected or abstracted over but which I hope to show permits an insightful analysis for a persistently challenging class of constructions in Bantu syntax. Ultimately, the view we come to is one where "cleft" picks out a distinct set of structures across languages. I end in §5.2 by outlining a novel structural typology of clefts where languages vary in the *verbality* of their copulas, and whether the cleft clause is obligatorily embedded or not.

In other words, I expand upon a proposal made by É. Kiss (1998), who proposes that Hungarian pre-verbal focus and English clefts can be given a partially-unified structural analysis. Specifically, the embedded clause in English clefts is an $\bar{\text{A}}$-fronting construction to a clause-internal position, akin to movement to the pre-verbal position in Hungarian. Here I add a third structural possibility, where Kirundi contains the same $\bar{\text{A}}$-fronted clause, but is embedded by a minimal, non-verbal clause that is independently motivated by Kirundi non-verbal predication. The resulting view is that "cleft-like" constructions

are stucturally distinct, but nonetheless share a clause with an Ā-fronted constituent; the status of this clause as a root or embedded clause, and the syntactic nature of the embedder, results in three logically possible structural types, which I show are exhaustively represented by Hungarian, English, and Kirundi. This proposal raises the question of whether clefts in other languages, particularly those with "particle copulas", may best be analyzed along the lines of the present proposal for Kirundi.

This paper reports data from elicitation, undertaken from January 2022 – April 2023 in Montreal, Quebec with three speakers of Kirundi.

2 Kirundi fronting constructions are Ā-derived

Kirundi fronting has been described as one of syntactic configurations which express narrow focus on the fronted constituent (Edenmyr 2001; Lafkioui et al. 2016).[2] Focused elements and *wh*-constituents may optionally appear fronted to the post-*ni* position. This is illustrated by the pair in (1b, repeated in 4). When fronted, they co-occur with an obligatory gap within the remnant.[3]

To anticipate the account to follow, I will call the construction in (4b) a cleft; I will postpone a discussion on this terminology to §3. When relevant, I will make a distinction between the *cleft* as a construction and (Ā-)*fronting* as the derivational process resulting in a cleft. Finally, following much of the literature, I will continue to use the term FOCUS CONSTRUCTION to refer to the constructions such as (4b) in Kirundi and other languages, especially when the source does. There are some concerns that may arise, particularly with respect to the aptness of this term (see §2.2), but I retain this usage below for conciseness and because I am unable to verify the information structural properties of the languages other than Kirundi, which I cite below.

In the remainder of this brief introductory section, I will outline the basic facts concerning Kirundi fronting, with an eye towards demonstrating some of the typically discussed differences displayed across analogous constructions in other Bantu languages.

(4) a. Yohaáni a-a-som-ye igitabu
Yohani 1SM-PST-read-PFV 7.book
'Yohani read a book.' (neutral sentence)

b. Ni **igitabu**$_1$ [Yohaáni a-a-som-yé ___$_1$]
NI 7book Yohani 1SM-PST-read-PFV.REL
'It's THE BOOK that Yohani read.' (fronting with *ni*)

As in many Bantu languages, Kirundi exhibits a subject/object asymmetry with respect to fronting (see, e.g.. Zentz 2016b on Shona). Wh-objects may remain in situ and objects may receive focus interpretation post-verbally, as seen in (5).[4] Subjects, however, are obligatorily fronted when they are wh-words or are focused, as seen by the obligatory clefting in (6).

[2]The other means of expressing narrow (or constituent) focus as opposed to broad (or sentence) focus is the conjoint/disjoint alternation (Meeussen 1959; Ndayiragije 1999; van der Wal 2017; Nshemezimana and Bostoen 2017), where the disjoint is taken to correspond to narrow focus on the sentence-final constituent. Beyond the contrast with exhaustivity effects used to argue for a dissociation of fronting and focus in §2.2, I will have little to say about this sentence-final constituent focus.

[3]There are some exceptions where locative adjuncts are optionally resumed, but I put these aside. These exceptions are common across similar constructions across languages, such as Hungarian (É. Kiss 1987).

[4]Post-verbal objects however, may not necessarily be in-situ; see Ndayiragije (1999).

(5) a. Yohaáni a-a-som-ye ikí?
Yohani 1SM-PST-read-PFV what
'Yohani read what?'

b. N'ikí Yohaáni a-a-som-yé
NI-what Yohani 1SM-PST-read-PFV
'What did Yohani read?'

(6) a. * Ndé a-a-som-ye igitabu?
who 1SM-PST-read-PFV 7.book
Intended: 'Who read the book?'

b. Ni-ndé a-a-som-yé igitabu?
NI-who 1SM-PST-read-PFV 7.book
'Who read the book?'

Finally, in some Bantu languages, the same particle used in focus constructions is also used pre-predicatively to signal predicate focus. This is exemplified in with data from Kîîtharaka (7; Abels and Muriungi 2008) and Kikuyu (8; Schwarz 2003, 2007)

(7) **Kîîtharaka pre-predicate focus marker (Abels and Muriungi 2008)**

a. N-Aana a-gûr-ir-e î-buku
FOC-1.Ana 1.SM-buy-PERF-FV 5-book
'ANA bought a book.'

b. Maria n-a-gûr-ir-e î-buku
1.Maria FOC-1.SM-buy-PERF-FV 5-book
'Maria BOUGHT A BOOK.'

(8) **Kikuyu pre-predicate focus marker (Schwarz 2003: p. 140, 142)**

a. **ne**-kee Abdul a-ra-nyu-ir-ɛ
FM-what A. SM-T-drink-ASP-FV
'WHAT did Abdul drink?

b. Abdul (ne) a-ra-nyu-ir-ɛ mae
Abdul FOC SM-T-drink-ASP-FV 6.water
'Abdul DRANK WATER.'

The *ni* particle in Kirundi, however, does not have the pre-predicative distribution (similar to, e.g., Shona; Zentz 2016a,b). Instead, predicate focus is marked with a distinct verbal prefix, *-ra-*, called the disjoint marker (Nshemezimana and Bostoen 2017) or the anti-focus marker (Ndayiragije 1999).

(9) a. * Yohaáni ni a-a-som-yé igitabu.
Yohani NI 1SM-PST-read-PFV 7.book
'What did Yohani read?'

b. Yohaáni a-ra-som-ye igitabu.
Yohani 1SM-DJ.PST-read-PFV 7.book
'Yohani [read the book]$_F$'

Having looked at the basic properties of Kirundi *ni* constructions in the context of similar structures in the Bantu language family, I will now to discussing their derivation via $\bar{A}$-movement.

2.1 The fronted constituent is derived by Ā-movement

In this section, I present data to establish the claim that the fronted constituent arrives in the post-*ni* position through Ā-movement. I will show that fronting constructions have three properties which are typically assumed to diagnose Ā-movement (see, e.g., Safir 2019 for a discussion of these diagnostics).

Firstly, Ā-movement can establish a long-distance dependency with its extraction site, by-passing multiple intervening subjects, as seen in (10).

(10) **Long-distance dependencies**

a. Kagabo yavúze kó Yohaáni yībaza kó Petero akūnda Kēza.
Kagabo a-a-vúg-ye kó Yohaáni a-ī-baz-a kó Petero a-kūnd-a Kēza
1.Kagabo 1SM-PST-say-PFV C Yohani 1SM-RFLX-think-IPFV C Petero 1SM-love-IPFV Keza
'Kagabo said that Yohani believes that Petero loves Keza.'

b. Ni Kēza Kagabo yavúze kó Yohaáni yībaza kó Petero
Ni Kēza$_i$ Kagabo a-a-vúg-ye [kó Yohani a-ī-baz-a [kó Petero
NI 1.Keza 1.Kagabo 1SM-PST-say-PFV C 1.Yohani 1SM-RFLX-think-IPFV C Petero
akūnda.
a-kūnd-a ___$_i$]]
1SM-love-IPFV
'It's Keza that Kagabo said that Yohani believes that Petero loves.'

Secondly, this dependency is island sensitive, showing that it is indeed a movement dependency. This is illustrated for adjunct islands (11), relative clause islands (12–13), and a language-specific island formed with a quotative complementizer *ngo* (14). For discussion on this last island, see Ndayiragije (1999).

(11) **Adjunct Islands**

a. n-a-gīye kw' isoko [kubēra n-kenér-ye umukâté].
1SG.SM-PST-walk.PFV to store because 1SM-need-PFV bread
'I went to the store because I needed bread.'

b. * Ni umukâté n-a-gīye kw' isoko [kubēra n-kenér-ye ___].
NI bread 1SG.SM-PST-walk.PFV to store because 1SG.SM-need-PFV
'It's bread that I went to the store because I need.'

(12) **Relative clause island (object)**

a. Ni igitabu n-a-gúr-ye ___ [umugēnzi wā-nje a-a-som-yé].
NI 4.book 1SG.SM-PST-buy-PFV 1.friend 1-1SG.POSS 1SG.SM-PST-read-PFV
'It's the book that I bought ___ [that my friend read].'

b. * Ni umugēnzi wā-nje nagúze igitabu [___ yasómye].
NI 1.friend 1-1SG.POSS 1SG.SM-PST-buy-PFV 4.book 1SG.SM-PST-read-PFV
Intended: 'It's my friend that I bought the book [that ___ read].'

(13) **Relative clause island (subject)**

a. Ni umūntu n-riko n-ra-ronder-a ___ [a-a-na-ib-ye
NI 1.person 1SG.SM-PROG? 1SG.SM-DJ-search.for-IPFV 1SM-PST-1SG.OBJ-steal-PFV
terefone].
5.phone
'It's the person I'm looking for ___ [who stole my phone].'

b. # Ni terefone n-riko n-ra-ronder-a umūntu [a-a-na-ib-ye
NI 5.phone 1SG.SM-PROG? 1SG.SM-DJ-search.for-IPFV 1.person 1SM-PST-1SG.OBJ-steal-PFV
___].

Intended: *'It's the phone I'm looking for the person [who stole ___].'
(OK under interpretation: 'It's the phone I'm looking for ___ [that the person stole].)'

(14) ***ngo*-islands**

a. Petero a-a-vúg-ye [ko/ngo Yohàáni a-a-nyō-ye amâzi].
1.Petero 1SM-PST-say-PFV C/C.QU 1.Yohani 1SM-PST-drink-PFV 5.water
'Peter said that Yohani drank milk.'
verb is *kunywa* '*drink*'

b. Ni amâzi Petero a-a-vúg-ye [ko/*ngo Yohàáni a-a-nyō-ye ___.]
NI 5.water 1.Petero 1SM-PST-say-PFV C/C.QU 1.Yohani 1SM-PST-drink-PFV
'It's (only) water that Peter said Yohani drank.'

Together, these properties establish that fronting is related to its extraction site via Ā-movement. However, as noted by both Torrence (2013a,b) and Hartmann and Zimmermann (2012), this data is compatible with two hypotheses: a null-operator analysis and a promotion analysis. The null-operator analysis posits Ā-movement of a phonologically null operator Op, which is bound by the ostensibly fronted constituent. Under a promotional analysis, wherein the fronted constituent is directly extracted. Following argumentation by Torrence for Wolof and Hartmann and Zimmermann for Bura, we expect to see reconstruction of the fronted constituent in the latter, but not the former. The following data shows that fronted constituents do indeed reconstruct into the extraction site, supporting a promotional analysis.

Examples are given for Condition A and Condition C. In (15), we see that a pronominal interpreted as a reflexive anaphor must be interpreted as bound by the embedded subject. That is, the reflexively interpreted anaphor must reconstruct into its extraction site for Condition A.

(15) **Condition A reconstruction**

a. Yonaáni$_1$ a-a-vúg-ye [kó Petero a-a-bōn-ye [ubwīwé$_{*1/2}$ bwambure]]
Yohani 1SM-PST-say-PFV C Petero 1SM-PST-see-PFV [his.own nakedness]
'Yohani$_1$ said that Peter$_2$ saw his own$_{*1/2}$ nakedness' (Condition A)

b. N' [ubwīwé$_{*1/2}$ bwambure] Yonaáni$_1$ a-a-vúg-ye [ko Petero$_2$ a-a-bōn-ye
FOC his.own$_{*1/2}$ nakedness Yohani 1SM-PST-say-PFV C 1.Petero 1SM-PST-see-PFV
___$_1$]

'It's his own$_{*1/2}$ nakedness who Yohani$_1$ said Peter$_2$ saw.' (Condition A reconstruction)

In (16a), we see that proper name *Yohaáni* cannot be interpreted as co-referential with the matrix *pro*-dropped subject. We see in (16b) that the R-expression is still ungrammatical when interpreted as

coreferential with the subject in the remnant clause. In other words, it obligatorily reconstructs for Condition C.

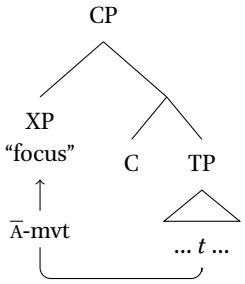
(16) **Condition C reconstruction**

a. *pro*$_{*1/3}$ a-a-vúg-ye [kó Petero a-a-bōn-ye Yohaáni$_1$]
pro 1SM-PST-say-PFV C Petero 1SM-PST-see-PFV 1.Yohani
'He$_{*1/3}$ said that Peter saw Yohani$_1$' (Condition C violation)

b. Ni Yohaáni$_1$ [*pro*$_{*1/2}$ a-a-vúg-ye kó Petero a-a-bōn-ye ___$_1$]
FOC 1.Yohani *pro* 1SM-PST-say-PFV C 1.Petero 1SM-PST-see-PFV
'It's Yohani$_1$ who he$_{*1/2}$ said Peter saw.' (Condition C reconstruction)

Together, the data in this section support the hypothesis that the fronted constituent is directly Ā-moved into the post-*ni* position, as illustrated in (17).

(17) **Claim (2a): fronting is Ā-derived**



2.2 The fronted constituent is exhaustively identified

In many languages which have been described as having a dedicated structural focus position, distinct from the base position where prosodic focus can be assigned, the non-canonical word order has been shown to have an additional, truth-conditional interpretive effect.[5]

For at least some speakers, Kirundi shows similar interpretive effects with *ni*-clefts but not with other positions also compatible with focus (such as sentence-final position, Ndayiragije 1999). This can be seen in, for example, mention-some contexts which do not pragmatically support an exhaustive answer (Cable 2008).

(18) Ni ivya-he biharūro bi-gabúr-w-a na kabiri?
NI 8-which 8.number 8SM-divide.REL-PASS-IPFV by two
'What are the even numbers?' ('Which numbers can be divided by two?')

a. Ibi harūro bi-gabur-w-a na kabiri ni (nka) kabiri, kane, na gatandatu.
8.DEM 8.number 8SM-divide-PASS-IPFV by two NI (about) two four and six
'The numbers divisible by two are (for example) two, four, and six.' (not exhaustive)

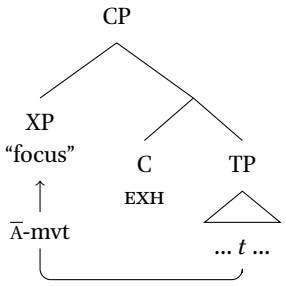
b. # ni kabiri, kane, na gatandatu bi-gabúr-w-a na kabiri
NI two four and six 8SM-divide-PASS-IPFV by two

[5]There is equally work suggesting that this exhaustivity interpretation is a presupposition, and demonstrating experimentally that the effect is either not present in certain contexts or is weaker than typically assumed (e.g. Büring and Križ 2013).

'It is two, four, and six that are divisible by two.'
(ok if these exhaustively pick out options from a list)

The above pair illustrates that the exhaustiveness is not tied to *ni*, which appears in both examples, but rather in the movement into post-*ni* position in the cleft example (18b).

(19) **Claim (2b): Ā-movement is for exhaustivity**



For now, I will assume that this position is the specifier of an embedding exhaustifying operator in C, as in (19) which attracts [+exhaustive] constituents to its specifier (Horvath 2007, 2013). While this will likely turn out to be too powerful, it is attractive as an initial hypothesis since it divorces exhaustivity (and focus) from the particle *ni*.

3 Structurally asymmetrical clefts: embedding by a non-verbal head

Having established that the fronted constituent is Ā-moved to its surface position, I will develop the central structural claim in this section, namely that these constructions are biclausal clefts, crucially with a syntactically deficient matrix clause. Firstly, in §3.1, I will show that the remnant clause is an embedded, but non-relative, clause (contrary to previous descriptions in Edenmyr 2001; Lafkioui et al. 2016) and that the fronted constituent lands clause-internally. Then, I turn briefly to the proposal on the syntactic identity of *ni*, where I propose that it is a non-verbal copula (and crucially distinct from the verbal copula). The upshot of this analysis is spelled out, where the lexical availability of a non-verbal copula results in fully expletive and defective clauses in clefts. This is unlike English, where the matrix clause of clefts centred around the copula necessarily encodes verbal categories such as tense and agreement, on par with lexical verbs.

3.1 Cleft clauses are embedded clauses

In this section, I will outline three morphosyntactic properties that show cleft clauses are embedded. I show that prior analyses which claim that clefts include a relative clause cannot be straightforwardly maintained, firstly on the basis that the properties used to establish this view are in fact properties of non-matrix predicates more generally, and that the fronted constituent does not behave like the head of a relative clause.

In §3.1.1 section, I will first exemplify these so-called "relative" properties exemplified by clefts and show that, while these properties are indeed shared with relative clauses in the language, they are not themselves sufficient diagnostics for a formal identification of the cleft clause with relative clauses in general (see É. Kiss 1998 for a similar claim on English clefts). I then go on in §3.1.2 to present two further arguments to show that the remnant clause cannot be straightforwardly analyzed as a relative clause:

the "head" of the ostensible relative clause is significantly more free than in relative clauses elsewhere, and that relative clauses lack the reconstruction effects seen in the previous section.

The upshot of this argument and analysis will turn out to be that, while clefts are often described as having a relative clause (Akmajian 1970; Chomsky 1971; Hedberg 2000; Lafkioui et al. 2016 for Kirundi specifically), the Kirundi data presented here provides reason to believe that the state of affairs is not so straightforward. This data sets the stage for a "promotional analysis" of clefts developed further in §3.2, whereby the clefted constituent is directly $\bar{A}$-moved to the embedded-clause initial position and embedded by a matrix predicate (see Torrence 2013b on Wolof and É. Kiss 1998 on English).

3.1.1 "Relative properties" are embedded properties

The remnant clause in cleft constructions differs from matrix clauses in three ways: (i) the presence of a high tone ("relative" tone; Zorc and Nibagwire 2007; Lafkioui et al. 2016), (ii) the choice of negation (Chaperon to appear) and (iii) the obligatory absence of disjoint marking (Ndayiragije 1999). Below, I address each in turn.

Firstly, consider the the tone on neutral sentences in (20a, 21a), in contrast with high tone that appears on the verb in remnant clauses shown in (20b, 21b).[6]

(20) **Clefts take embedded tone**

a. Yohaáni a-a-som-ye igitabu
Yohani 1SM-PST-read-PFV 7.book
'Yohani read a book.'

b. Ni igitabu$_1$ [Yohaáni **a-a-som-yé** ___$_1$]
NI 7book Yohani 1SM-PST-read-PFV.EMB
'It's THE BOOK that Yohani read.'

(21) a. Yohaáni a-a-gī-ye ku kazi mugatôndo
Yohani 1SM-PST-go-PFV to work in.morning
'Yohani went to work this morning.'

b. Ni Yohaáni$_1$ [___$_1$ **a-a-gī-yé** ku kazi mugatôndo]
NI Yohani 1SM-PST-go-PFV.REL to work in.morning
'It's YOHANI who went to work this morning.'

This tone pattern is shared with relative clauses (as seen in (22a)), suggesting to previous researchers that the cleft clause is a relative clause (Ndayiragije 1999; Lafkioui et al. 2016). However, non-$\bar{A}$ contexts such as embedded clauses with the complementizer *kó* also require the embedded tone pattern, illustrated in (22b).[7] As such, the use of the so-called "relative tone" is not a sufficient diagnostic to identify the remnant of clefts as a relative clause.

(22) **Embedded tone across contexts**

[6]The actual pattern is slightly more complicated, interacting with the lexical high tone and the high tone associated with the tense prefix. The complication is that the recent past has the high tone, which neutralizes the tonal contrast between the high tone which appears in embedded contexts.

[7]See Zorc and Nibagwire 2007: p. 325 for a list of complementizers with the same property.

a. N-a-bōn-ye igitabu Yohaáni **a-a-som-yé**
1SG.SM-PST-see-PFV 7.book 1.Yohani 1SM-PST-read-PFV.EMB
'I saw the book that Yohani read.' (Relative clause)

b. N-a-vug-ye kó Yohaáni a-a-som-yé igitabu
1SG.SM-PST-say-PFV that Yohani 1SM-PST-read-PFV.EMB 7.book
'I said that Yohani read a book.' (Complement clause)

A similar argument can be made from other properties shared by the remnant clause and relative clauses generally being shared by all non-matrix clauses. Consider the negation data in (23). Kirundi predicate negation is expressed by one of two syntactically conditioned markers: a pre-subject-marker *nti-* which occurs in matrix clauses (23a), and a post-subject-marker *-ta* which occurs in remnant clauses of clefts, relative clauses, and embedded clauses (23b); see Chaperon to appear for an account. In short, the choice of negation marker correlates with the matrix/non-matrix distinction rather than an A/$\bar{\text{A}}$ distinction (Ndayiragije 1999).

(23) **Clefts take secondary negation**

a. Yohaáni nti-a-kor-á imikâté
Yohani NEG-1SM-make-IPFV.REL 4.bread
'Yohani didn't make bread.'

b. Ni Yohaáni a-da-kor-á imikâté
NI Yohani 1SM-NEG-make-IPFV.REL 4.bread
'It's YOHANI who didn't make bread.'

Finally, the availability of conjoint/disjoint alternation (or antifocus marker) has been taken to diagnose $\bar{\text{A}}$-movement (Ndayiragije 1999: p. xx; Nshemezimana and Bostoen 2017). The compatibility of the disjoint marker can therefore plausibly be used as an argument to unify the remnant of clefts and relative clauses. Once more, the split is in fact between matrix/non-matrix clauses. While the disjoint/anti-focus marker *-ra-* is available in matrix clauses, it is ruled out in clefts.

(24) **No *-ra-* in clefts (Ndayiragije 1999: p.407)**

a. Ni abâna ba-á-(*ra)-nyôye amatá
NI 2.children 2SM-DIST.PST-RA-drink.PERF 6.milk
'It was children who drank milk.'

b. Ni amatá abâna ba-á-(*ra)-nyôye
NI 6.milk children 2SM-DIST.PST-RA-drink.PERF
'It was milk that children drank.'

The data above has shown that a relative clause analysis of the remnant clause of clefts rests on tenuous arguments, where the properties motivating such an analysis are too widespread across non-relative clause contexts to be seen as a sufficiently diagnosing a relative clause structure. Nonetheless, the cleft clause falls into a natural class with other non-matrix clauses in the language, suggesting that they are indeed embedded. I will return to this point in §3.2.1, where I discuss various proposals regarding the number of clauses in similar cleft-like structures across Bantu.

3.1.2 Against fronted constituents as the head of relative clause

Above, I established that cleft clauses are embedded, and suggested that the embedded clause is not in fact a relative clause on the basis of the wider distribution of so-called relative properties. Here, I present two arguments which further demonstrate the non-identity between Kirundi relative clauses and cleft clauses.

The first argument is a familiar argument from É. Kiss (1998), brought up in the context of the extraposition analysis of English clefts. The observation is that clefted constituents are significantly freer than the head of a relative clause with respect to what constituent can occupy this position. Consider the extraposition analysis of Akmajian (1970), whereby the cleft clause is derivationally related to a free-relative. While unproblematic for cases where the head of the relative clause is nominal, as in (25), challenges arise when the clefted constituent cannot be the head of a relative clause such as in (26)

(25) **Extraposition analysis (É. Kiss 1998: p. 257)**

a. $[_{CP}$ who is sick$]$ is me $\rightarrow$

b. it$_i$ is me [who is sick]$_i$

(26) a. It was to John [that I spoke]

b. * $[_{CP}$ that I spoke$]$ was to John

Kirundi shows similar freedom in what can be clefted. The example in (27a) shows an adverb can be fronted; in (27b), a full clause can occupy this position. It is unclear what structure a clausally-headed relative clause would have.

(27) a. Ariko ni keénshi tu-ya-reéng-a
but NI often 1PL.SBJ-PRES-6OBJ-violate-FV.REL
'But it is often that we violate them (the laws)'
(Lafkioui et al. 2016: p. 82)

b. Ni [kubêra n-kenéy-e u-mu-kâté] n-a-gīy-e kw' ī-sokó
NI because 1SG.SBJ-need-PFV AUG-3-bread 1SG.SBJ-go.REL-PFV to AUG-5.store
'It's because I needed bread that I went to the store.'

I take these data as arguing against analyzing the clefted constituent as the head of a relative clause. However, this does not strictly exclude an analysis similar to the extraposition analysis, whereby the clefted constituent and the cleft clause (as relative clause) are in a looser relationship. Below, I present data that excludes the relative clause analysis of Kirundi clefts by showing that the structures underlying the two constructions distinct. Specifically, I show that they differ in whether the $\bar{A}$-moved constituent reconstructs into the clause: clefted constituents do, relativized constituents do not.

Consider the data on binding below. As seen in (28), the pronominal phrase *ukuboko$_i$ kw-iwe* 'his/her/its arm', when the pronominal is interpreted as a bound-variable pronoun, must be interpreted with the pronominal being bound by the most local c-commanding DP. In other words, (28) is only good when the bound-variable pronoun is the embedded subject; it behaves like an anaphor.

(28) Yohaáni$_1$ a-a-vug-ye kó Petero$_2$ a-a-komerek-tse ukuboko$_i$ kw-iwe$_{*1/2}$
Yohani 1SM-PST-say-PFV that Peter 1SM-PST-wound-PFV 5.arm LK-his.own

'Yohani said that Peter hurt his own arm.'

Now, consider the data with respect to two Ā-fronting processes: relativization in (29a) and clefting in (29b). The result is that the bound-variable pronoun must be interpreted as anaphoric to the matrix subject when relativized, and as anaphoric to the embedded subject when clefted. Clefting, but not relativization, reconstructs for Condition A.

(29) **Relative clauses do not reconstruct for Condition A**

a. Yohaáni$_1$ a-a-hamb-ir-iye [ukuboko$_i$ kw-iwe$_{1/*2}$]$_i$ [Petero$_2$ a-a-komerek-tse
Yohani 1SM-PST-bury-APPL-PFV 5.arm LK-his.own Peter 1SM-PST-wound-PFV
___$_i$]

'Yohani bandaged his own arm that Peter hurt.'

b. Ni [ukuboko$_i$ kw-iwe$_{*1/2}$]$_i$ Yohaáni$_1$ a-a-vug-ye kó Petero$_2$ a-a-komerek-tse
NI 5.arm LK-his.own Yohani 1SM-PST-say-PFV that Peter 1SM-PST-wound-PFV
___$_i$

'It's his arm that Yohani said Peter bandaged.

A similar observation can be made from Condition C reconstruction. In (30), we see the baseline case where the possessor *Kēza* must be interpreted as disjoint from a pronominal subject, in accordance with Condition C. When the object is relativized as in (31a), however, the pronominal subject of the relative clause may be interpreted as co-referential as the possessor of the relativized object. This is not the case for the cleft in (31b), where the possessor is once more obligatorily referentially disjoint from the subject. In other words, clefts, but not relative clauses, reconstruct for Condition C.

(30) *pro*$_{*1/2}$ a-a-shír-ye igitabu cā Kêza$_1$ ku mêzá
pro 1SM-PST-put-PFV 7.book 7.LK 1.Keza on 5.table
'She$_{*1/2}$ put Keza$_1$'s book on the table.'

(31) **Relative clauses do not reconstruct for Condition C**

a. N-a-som-ye [igitabu cā Kêza$_1$]$_i$ *pro*$_{1/2}$ a-a-shír-ye ___$_i$ ku mêzá
1SG.SM-PST-read-PFV 7.book 7.LK 1.Keza 1SM-PST-put-PFV on 5.table
'I read Keza$_1$'s book that she$_1$/he$_2$ put on the table.'

b. Ni [igitabu cā Kêza$_1$]$_i$ *pro*$_{*1/2}$ a-a-shír-ye ___$_i$ ku mêzá
NI 7.book 7.LK 1.Keza 1SM-PST-put-PFV on 5.table
'I read Keza$_1$'s book that *she$_1$/he$_2$ put on the table.'

The above data all show that the remnant of fronting shares some surface properties with relatives clauses, but cannot ultimately be identified as a relative clause. One proposal to derive this distinction is to analyze clefts as involving a directly Ā-moved (promoted) constituent, and the relative clause as involving a null operator co-referential with the nominal it is adjoined to. This analysis, which I adopt here, is spelled out in (32)

(32) **Relative clauses and clefts are structurally different**

a. $[_{DP}$ DP$_1$ $[_{CP}$ Op$_{1,i}$ C $[_{TP}$ ... t_i ... $]$ $]$ $]$ (Relative clause)

b. $[_{\text{CP}}\ \text{DP}_i\ \text{C}_{\textbf{EXH}}\ [_{\text{TP}}\ \ldots\ t_i\ \ldots\ ]\ ]$ (Cleft)

In this section, I have established that the cleft clause patterns like other embedded clauses in Kirundi, and that there are reasons to argue against identifying it as a relative clause, despite surface similarities. I showed that reconstruction in Kirundi distinguishes between the clefts and relative clauses, and that clefts are more flexible in what can occur in their clause-initial position, noting similar observations made for English. In the following section, I will consider the second half of my proposal: the nature of the embedding material in clefts.

3.2 The structure of the matrix clause: Kirundi *ni* is non-verbal Pred

We turn now to the question of what *ni* is doing in these structures. In this section, I will lay out the assumptions I make and discuss the initial motivation for the hypothesis that *ni* is a non-verbal element, which I take structurally to be an instance of non-verbal Pred (see Adger and Ramchand 2003). The discussion, for now, will be limited to outlining the central consequence of this hypothesis for the structural possibilities available for cleft constructions across languages. In §3.2.1, I overview a debate about the number of clauses present in similar constructions across Bantu languages, and call into question one diagnostic presented as conclusive by Zentz (2016a,b). Then, in §3.2.2, I claim that Kirundi C heads are differentiated into a single matrix C head and multiple obligatorily embedded ones, where all CPs with $\bar{\text{A}}$ positions fall into the latter. When one of these heads is present, an expletive PredP is used to satisfy this structural requirement without adding additional semantically substantive content; in other words, matrix clauses in clefts are devoid of semantic content (in contrast to the view that matrix material in clefts are semantically substantive, as argued for English by Hedberg 2000).

As as starting point, I will consider proposals forwarded for clefts in other Bantu languages. There are two relevant analytical choice points in the literature for cleft constructions: firstly, whether the construction is mono-clausal or bi-clausal (Zentz 2016a); and secondly, the syntactic/semantic nature of the accompanying particle (analogues of Kirundi *ni*). With respect to the first choice point, I will argue that the mono-/bi-clausal distinction is too blunt, and that Bantu clefts (and Kirundi in particular) show us that a more fine-grained notion of mono-/bi-verbality provides a more insightful way of understanding cleft structures and the often-made observation that clefts and non-verbal predication share significant formal properties (Green 2007). The second choice point regarding the nature of the *ni* is related to the question of verbality in the upper clause of clefts. However, much work on Bantu (which adopt a cleft analysis) assumes that analogues of *ni* is the copula and that it is therefore a verb (Zentz 2016a,b).

I argue, following the typological distinction made by Pustet (2003) between verbal copulas (such as English *be*) and non-verbal copulas, that the matrix clause of clefts is a non-verbal expletive embedding structure. Crucially, *ni* is not syntactically verbal, and does not support the projection of verbal functional material (the extended projection of the verb).

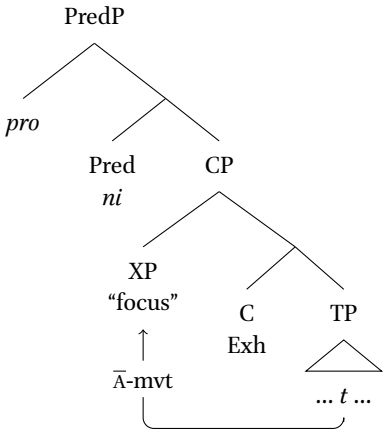
(33) **Claim (2c): *ni* is Pred**
Pred introduces the subject for non-verbal predicates, subject to the restrictions discussed in §4.

(34) **Claim (2d): *ni* fronting constructions are mono-verbal**
The matrix clause in clefts is semantically expletive, and is syntactically non-verbal. As such, it cannot project verbal functional material such as tense.

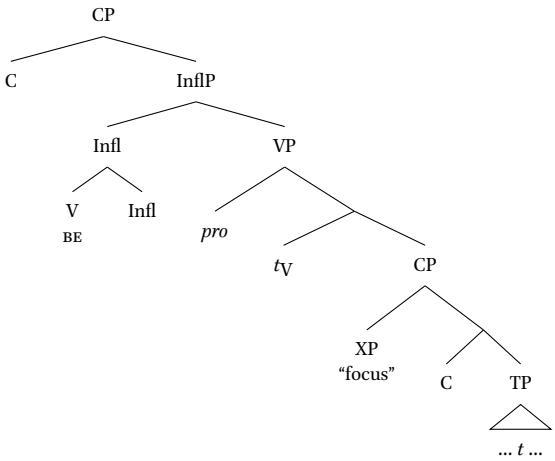
The resulting construction is a cleft with a defective, non-verbal root clause, distinct from English type "symmetric" clefts where both clauses contain a verbal extended projection. The proposal for Kirundi

clefts is illustrated in (35a), where the relevant property is the lack of verb in the upper clause. This contrasts with an English-type cleft as illustrated in (35b), where the upper clause contains a verbal element BE, and is surmounted by additional functional material typically associated with verbs. I will return to develop this idea in §5.2, after going through the proposal in more detail in this section.

(35) a. Kirundi asymmetric cleft



b. English symmetric cleft (after É. Kiss 1998)



The difference between the two structures is tied entirely to the syntactic properties of the embedding material; in other words, the availability of the structure (35a) in Kirundi is due to the availability of a non-verbal copula in the lexicon. The remaining task will then be to motivate independently the existence of these non-verbal root clauses I will postpone motivating this structural possibility to §4, where I discuss the distribution of *ni* and the verbal copula in non-verbal predication. In the remainder of this section, I will discuss the clausality question in more detail, defending the bi-clausal status of clefts.

3.2.1 The clause-hood question: mono-clausal vs. bi-clausal clefts

In this section, I will discuss one previously proposed diagnostic for the number of clauses in the cleft construction. While the question regarding the number of clauses in clefts is closely-related to the question of whether *ni* is a copula or a focus-marker, I show that they are not equivalent. For instance, Wasike (2007) analyzes the Lubukusu equivalent to *ni* as a head within the left-periphery of the cleft clause, but

nonetheless maintains the bi-clausality of the overall construction.[8] In the end, I will argue here this question of whether clefts are bi-clausal or mono-clausal, in the case of languages like Kirundi, is too coarse of an opposition, leading to the ambiguity of the previously proposed diagnostics.

Instead, I propose that clefts can be bi-clausal, but need not consist of two *symmetrically sized* clauses. Specifically, I will show that Kirundi clefts consist of a full CP cleft-clause which is embedded by a syntactically reduced PredP. I argue that this PredP lacks the TP functional projection; I will provide independent evidence that this PredP is a licit matrix clause in the language, despite being syntactically quite bare, in the following section where I discuss copular clauses with *ni*.

As a starting point, I will follow closely the overview of clause-hood in clefts across different Bantu languages presented by Zentz (2016b: p. 1598ff.), who notes that while many diagnostics are ambiguous, there is sufficient evidence to differentiate bi-clausal constructions from mono-clausal ones. The sole diagnostic that he argues is able to differentiate the two is a diagnostic due to Schwarz (2003); Abels and Muriungi (2008), namely whether topicalization of temporal adverbials in cleft construction patterns with bi-clausal or mono-clausal contexts. I will show in this section that topicalization of temporal adverbials does not unambiguously determine the number of clauses in clefts in Kirundi.

The crucial generalization for Schwarz (2003), Abels and Muriungi (2008), and Zentz (2016b) is the clause-boundedness of left-dislocation. In Kîîtharaka, but not in Shona, left-dislocation of temporal modifiers is permitted from clefts but not from relative clauses. Consider firstly the Kîîtharaka data in (36). In (36a), we see that focus constructions permit the left-dislocated temporal adverb *îgoro* 'yesterday' to be interpreted as modifying the main predicate (the seeing). However, in the unambiguously bi-clausal relative clause shown in In (36b), the same adverb cannot be interpreted as modifying the main predicate.

(36) **Kîîtharaka relative clauses and focus construction differ for modifier left-dislocation (Abels and Muriungi 2008: p. 725)**

a. **î-goro**$_2$ i-mw-amba$_1$ Peter a-ra-on-ir-e t_1 t_2.
5-yesterday FOC-1-thief 1.Peter 1.SM-REC.PST-see-PFV-FV
'Yesterday, THE THIEF Peter saw.' (Left-dislocation ok for focus construction)

b. * **î-goro**$_2$ boriisi ba-ka-thaik-a [$_{RC}$ mw-amba$_1$ û-ra Peter
5-yesterday 2.police 2.SM-FUT-arrest-FV 1-thief 1-that 1.Peter
a-ra-on-ir-e t_1 t_2].
1.SM-REC.PST-see-PFV-FV
'Yesterday, the police will arrest the thief that Peter saw.'
(No left-dislocation for relative clause)

This contrast is taken by both Abels and Muriungi (2008) and Zentz (2016b) to reveal the presence of a clause-boundary in the relative clause case (36b), and the lack of a corresponding boundary in the focus construction (36a). The conclusion is therefore that the focus constructions are mono-clausal.

Consider now the corresponding data in (37) for Shona. We see that adverbial left-dislocation in focus construction shown in (37a) is ungrammatical, as is the relative clause case shown in (37b).

(37) **Shona relative clauses and clefts both disallow modifier left-dislocation (Zentz 2016b: p.167)**

[8] Related proposals consider the copular element to be within the left-periphery of a mono-clausal construction. See O'Neill (2019) on English amalgam specificational clauses, and Martinović (2021a) on Wolof left-peripheral non-verbal predication.

a. * **Nezuro**$_2$ i-m-bavha$_1$ ya-aka-on-a t_1 t_2.
yesterday NI-9-thief 9.NSE-1.SM.TA-see-FV
'Yesterday, it's A THIEF that s/he saw.' (No left-dislocation for cleft)

b. * **Nezuro**$_2$ ma-purisa a-cha-sung-a [$_{RC}$ m-bavha$_1$ ya-aka-on-a t_1 t_2].
yesterday 6-police 6.SM-FUT-arrest-FV 9-thief 9.NSE-1.SM.TA-see-FV
'Yesterday, the police will arrest the thief that s/he saw.'
(No left-dislocation for relative clause)

From this contrast, Zentz (2016b: p.168f.) concludes that the cleft is bi-clausal. I agree with the general line of reasoning here, but demonstrate that the condition operative in Shona cannot be transfered to Kirundi. With respect to the temporal modification, the Kirundi data patterns like Kîîtharaka. While this would suggest, following Zentz's argumentation, that Kirundi clefts are mono-clausal, I will argue that there is another reason why temporal adverbs are unable to be fronted. Crucially, the possibility of left-dislocating from clefts is instead due to the absence of a TP in the matrix clause of Kirundi clefts.[9]

Turning to Kirundi, we see firstly that, like Shona and Kîîtharaka, temporal modifiers cannot be left-dislocated from relative clauses, as seen in (38). In contrast, the left-dislocation of temporal modifiers from clefts is grammatical (39).

(38) **Kirundi relative clauses disallow modifier left-dislocation**

a. N-zō-vug-an-a umugabo [$_{RC}$ a-a-tsîn-ze ihiganwa ryo kwiíruká
1SG.SM-FUT-speak-COM-IPFV 1.man 1SM-PST-win-PFV 5.competition 5.LK to.run
mūndwi ihezé].
last.week
'I will speak to the man who won the race last week.'

b. * **Mūndwi ihezé** n-zō-vug-an-a umugabo [$_{RC}$ a-a-tsîn-ze
last.week 1SG.SM-FUT-speak-COM-IPFV 1.man 1SM-PST-win-PFV
ihiganwa ryo kwiíruká].
5.competition 5.LK to.run
Intended: 'Last week, I will speak to the man who won the race.'

(39) **Kirundi clefts allow modifier left-dislocation**

a. Ni [$_{CP}$ Kagabo [$_{Rem}$ a-a-tsîn-ze ihigawa ryo kwiíruka **mūndwi ihezé**]]
NI Kagabo 1SM-PST-win-PFV 5.competition 5.LK to.run last.week
'It's Kagabo that won the race last week.'

b. **Mūndwi ihezé** ni Kagabo a-a-tsîn-ze ihigawa ryo kwiíruka
last.week NI Kagabo 1SM-PST-win-PFV 5.competition 5.LK to.run
'Last week, it's Kagabo that won the race.'

The observed variation between the Shona, Kîîtharaka, and Kirundi data is summarized in the table in (40). This picture, taking the argumentation of Abels and Muriungi (2008) and Zentz (2016b) at face value, directly contradicts the claim made in my account that Kirundi clefts are bi-clausal.

[9] I do not have an account for the differences in grammaticality between the Shona and Kirundi cases, however. To the extent that Shona clefts are indeed bi-clausal, the account presented below suggests that the matrix clause in Shona clefts may be syntactically richer than that of Kirundi clefts. Whether this can be substantiated empirically cannot be confirmed at present.

(40) **Summary of cleft variation with respect to left-dislocated modifiers**

	Left-dislocation ok?		Analysis of cleft
	Rel. clause	Cleft	
Shona	*	*	bi-clausal (Zentz 2016a,b)
Kîîtharaka	*	✓	mono-clausal (Abels and Muriungi 2008)
Kirundi	*	✓	bi-clausal (present proposal)

However, an alternative account for the left-dislocation facts is available. Note that what is crucial for Zentz is that left-dislocation is clause-bound in the sense that left-dislocated modifiers may not cross a clause-boundary. However, under the hypothesis that clefts are indeed bi-clausal. Kirundi must crucially permit clause-boundary-crossing for left-dislocated elements. The proper generalization for Kirundi instead appears to be the presence of an intervening tense projection.

(41) **Generalization for temporal modification**
(Left-discloated) temporal modifiers cannot be interpreted as modifying past any temporal domain, or TP node.

In other words, the matrix clause in clefts (headed by *ni*) lacks a possible adjunction site for temporal modifiers. One piece of evidence that *ni*-clauses lack the TP projection entirely comes from coordination data. Note firstly that the coordinator *kāndi* is able to coordinate two verbal predicates (42a), and is also able to coordinate two nominal predicates accompanied by *ni* (42b).

(42) **Coordination with *kāndi***

a. Yohaáni a-ra-som-a kaāndi a-ra-andik-a
Yohani 1SM-DJ-read-IPFV COORD 1SM-DJ-write-IPFV
'Yohani reads and writes.'

b. Yohaáni ni umusomyi kaāndi ni umwanditsi
Yohani NI reader COORD NI writer
'Yohani is a reader and a writer.'

However, non-verbal predicates with *ni* are unable to coordinate with verbal predicates.[10]

(43) **VP-predicates and *ni*-predicates cannot be coordinated**

a. ?* Yohaáni a-ra-som-a kaāndi ni umuwanditsi
Yohani 1SM-DJ-read-IPFV COORD NI writer
Intended: 'Yohani reads and is a writer.'

Given the data presented in this section, we have seen that one argument for the presence/absence of a clause-boundary in cleft constructions taken to demonstrate the bi-/mono-clausal status of clefts in two Bantu languages. I showed that the argument is not so unambiguous, and that Kirundi seems to present an intermediate state of affairs – that is, temporal adverbs may be left-dislocated in clefts (and not in relative clauses), but there is nonetheless reason to believe that the cleft is bi-clausal. I will discuss this latter point in more detail in the following subsection. Here, I showed that the restriction operative

[10]There is some variation with respect to the strength of this ungrammaticality. In any case, the example in (**??**) is degraded with respect to the examples in (42a) and (42b).

on left-dislocated temporal modifiers in Kirundi is tied specifically to the presence of TP, which clefts are argued here to lack, rather than the weaker condition on clause-boundedness proposed by Schwarz (2003: p. 80) for Kikuyu and adopted by Abels and Muriungi (2008) for Kîîtharaka and Zentz (2016b) for Shona.

3.2.2 Matrix and non-matrix C

In this section, I discuss further arguments for the bi-clausality of cleft clauses, showing that they show properties of embedded clauses more generally. PredP is the minimal material permitted in the language to satisfy this embedded clause's requirement that it be selected.

Given that I have argued immediately above that previously forwarded arguments for mono-/bi-clausality of clefts is not as conclusive as it initially appeared, this section will present one additional argument. In doing so, I address an analytical question that has remained in the background up to this point: namely, if fronting in Kirundi is to an embedded clause-internal position, why is the defective matrix clause headed by *ni* required? To answer this question, consider again the data presented in (45), showing that the remnant of clefts has properties shared by other embedded clauses. In each of the examples in (45), the verb is marked with a high tone on the second mora of the stem (in these examples, falling on the final vowel).

(44) Yohaáni a-a-som-ye igitabu mu gatôndo
Yohani 1SM-REC.PST-read-PFV 7.book in 12.morning
'Yohani read a book this morning.'

(45) **Embedded clauses in Kirundi**

a. N-a-vug-ye kó Yohaáni a-a-som-yé igitabu mu gatôndo
1SG.SM-REC.PST-say-PFV COMP Yohani 1SM-REC.PST-read.EMB-PFV 7.book in 12.morning
'I said that Yohani read a book this morning.'

b. N-a-bon-ye igitabu Yohaáni a-a-som-yé mu gatôndo
1SG.SM-REC.PST-se-PFV 7.book Yohani 1SM-REC.PST-read.EMB-PFV in 12.morning
'I saw the book that Yohani read this morning.'

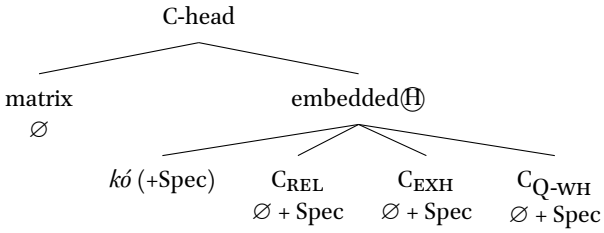
c. Ni igitabu Yohaáni a-a-som-yé mu gatôndo
NI 7.book Yohani 1SM-REC.PST-read.EMB-PFV in 12.morning
'It's the book that Yohani read this morning.'

d. Ni iki Yohaáni a-a-som-yé mu gatôndo?
NI what Yohani 1SM-REC.PST-read.EMB-PFV in 12.morning
'What did Yohani read this morning?'

The shared properties seen in these data lead to a generalization that clauses in Kirundi are headed by C marked as either matrix or embedding. The exhaustivity operator, as well as the operators driving *wh*-movement and relativization, are taken here to be covert members of class of embedded complementizers in Kirundi.[11] This is summarized in (46). I take the high tone appearing on the verb to correspond to embedding complementizer.

[11] At least a subset of these are overt in some Bantu languages. See, for example, Schneider-Zioga 2007 on Kinande.

(46) **Inventory of the C system in Kirundi**



One consequence of the observed split between matrix C and embedded C heads is that all heads which have an Ā-position are not licit root clauses. In other words, movement to Spec,CP appears to be obligatorily selected. This requirement finds some support in the subject/non-subject asymmetry with *wh*-in-situ (and focus-in-situ, to the extent that the IAV/sentence final position is a dedicated focus position). Consider the data in (47), which shows that subject *wh*-questions are obligatorily clefted, unlike non-subject *wh*-questions in (48).

(47) **Subject *wh*-question is obligatorily ex-situ**

a. * Ndé a-a-som-yé igitabu mu gatôndo?
who 1SM-REC.PST-read.EMB-PFV 7.book in 12.morning
Intended: 'Who read the book this morning?'

b. Ni ndé a-a-som-yé igitabu mu gatôndo?
NI who 1SM-REC.PST-read.EMB-PFV 7.book in 12.morning
'Who read the book this morning?'

(48) **Non-subject *wh*-question is optionally ex-situ**

a. Yohaáni a-a-som-yé iki mu gatôndo?
Yohani 1SM-REC.PST-read.EMB-PFV what in 12.morning
'What did Yohani read this morning?'

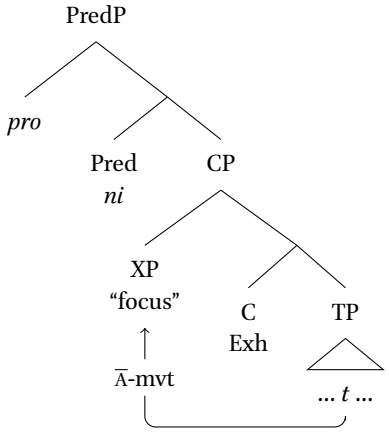
b. Ni iki Yohaáni a-a-som-yé mu gatôndo?
NI what Yohani 1SM-REC.PST-read.EMB-PFV in 12.morning
'What did Yohani read this morning?'

On the assumption that all *wh*-questions have an operator in C that attracts wh-constituents, the above contrast is unexpected. However, Ndayiragije (1999) among others proposes a low, VP-periphery to which VP-internal constituents may move. The in-situ strategy, under this view, is movement into a distinct low Ā-position. Adopting this proposal, the observed asymmetry, then, arises because subjects are too high to be licensed in this VP-periphery, and must instead be licensed in CP of the root clause. By hypothesis, the presence of this Ā-position hosting projection cannot be a licit root clause and therefore the cleft strategy is employed.

The result is that constituent questions of low, internal arguments may be formed by two means: movement into the low VP-periphery or movement into the high CP-periphery. The former is the "in-situ" strategy, whereas the latter is the "ex-situ" strategy and as a result of the particular complementizer system of Kirundi, requires embedding under further material, the most minimal of which is the non-verbal Pred *ni*. Constituent questions of external arguments, however, are introduced too high for movement into the low VP-periphery, and is obligatorily moved to the CP-periphery.

While the above view is ultimately stipulated as a lexical property of the elements in the inventory of C-heads in Kirundi for the moment, it adequately unifies the obligatory presence of a selecting element in relative clauses, clefts, and complement clauses. Ultimately, we would like to substantiate this claim but I will leave this for future work. For now, consider the final structural hypothesis for Kirundi clefts repeated once more in (49).

(49) **Final analysis of Kirundi clefts**



This analysis captures the Ā-properties of the fronted constituent, takes into account the exhaustiveness of the fronted constituent, and the obligatory embedding predicate which in Kirundi is a syntactically minimal clause. This view suggests that, in considering similar constructions across Bantu, we must disentangle our notion of clause-hood from the presence of a syntactically verbal element. Naturally, this analysis ultimately rests on the claim that non-verbal Pred is a grammatical matrix construction. In the following section, I show that the restricted distribution of *ni* in non-verbal predication supports the generalization that inflectional structure is fully absent in contexts with *ni*, substantiating this claim further. After this, I discuss alternative analyses, and then spell out precisely what I mean by claiming *ni* to be a Pred head in §5.2, in light of a the typological predictions made by this account.

4 Non-verbal predicates: Kirundi *[Infl Pred]

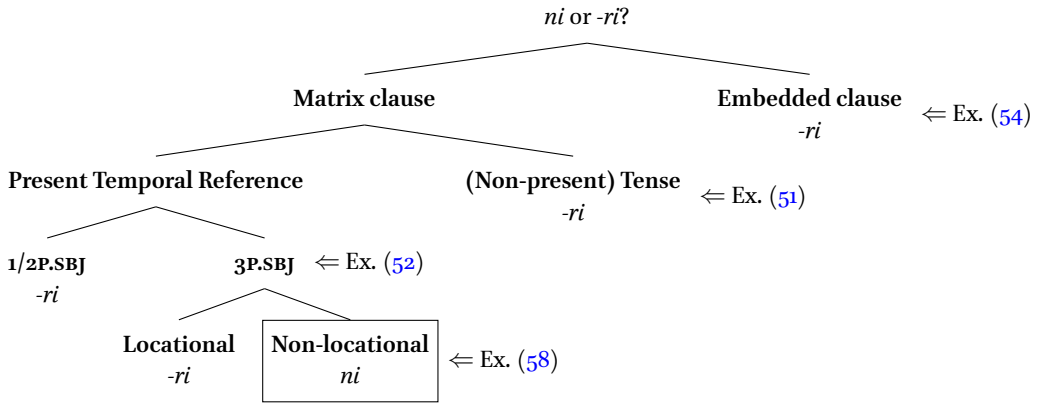
The analysis presented above rests on the assumption that *ni* is a non-verbal predicative element, Pred, and presupposes that PredP is a permitted matrix clause in Kirundi. In this penultimate section, I motivate this view from data on non-verbal predication, which I show empirically justifies this presupposition: that *ni* is restricted to syntactically non-verbal contexts and can be used as a grammatical matrix clause, and that the verbal *-ri* derives a syntactically (defective) *verbal* clause and thereby requires verbal inflectional material. In sum, the data in this section provides additional support for a matrix PredP in Kirundi, which obligatorily lacks verbal functional structure.

The main generalization is that *ni* is, as in many Bantu languages, restricted to non-locational predication of third-person subjects in present tense. I account for this distribution entirely on the syntactic requirements of predication other than this restricted context. In other contexts, a verbal copula *-ri* is used instead. While Jerro (2013) provides an account based on a proposed semantic difference between copulas, I will argue that this reflects the syntactic/semantic properties of the predicate rather than those of the copula.

In other words, Kirundi has two "copularizations" strategies for non-verbal predicates (Pustet 2003), which is determined by both the predicate and downstream requirements that can only be fulfilled by verbal INFL. When predication requires no INFL structure, the non-verbal Pred *ni* is used; when INFL structure is required independently, the verbal copula *-ri* is used to support it.

More generally, the properties which determine the choice of copula across Bantu, as well as the number of copulas, varies considerably (see Gibson et al. 2019; Gluckman 2022). For Kirundi, the basic claim is that the copula choice reflects a *structural* difference rather than simply a lexical one and that this structural difference reflects independent syntactic requirements to bind an eventuality argument (Adger and Ramchand 2003; Welch 2012) and to licence nominal person features (Béjar and Rezac 2003). When neither of these requirements are present, the use of Pred *ni* is licensed. The distribution of *ni* is summarized below in the decision tree in (50).

(50) **Contexts of use**



The remainder of the section presents the data and develops the analysis for non-verbal predication, with a view to demonstrating that PredP is a licit root clause.

4.1 Distribution of *ni* in non-verbal predication

Non-verbal predication in Kirundi with *ni*, as in many other Bantu languages (Gibson et al. 2019), is restricted to present-tense non-locational predicates with third-person subjects. The basic character of the generalizations to be drawn in this section is that *ni* is banned from contexts where there are either interpretive and syntactic requirements needing to be met by INFL further on in the derivation. Consider firstly, the tense restriction illustrated in (51).[12]

(51) **Overt tense requires *-ri***

a. Umwígīsha **ni** Yohaani
1-teacher *ni* John
'The teacher is John'

b. Keerá, Yohaáni yá**ri** umwígīsha
keerá Yohaani a-á-**ri** umwígīsha
before John 3SG.SM-PST-*ri* 1.teacher

[12]More properly, and anticipating the discussion below, the *ni* clauses might be better considered tense-less. The present temporal interpretation of these instead comes from the spatio-temporally undifferentiated character of predication without eventuality arguments.

'John was a teacher, a while ago.'

Further, all non-third-person subjects are banned from predication with *ni*, as can be seen in (52).

(52) **Speech Act Participant subjects require *-ri***

a. Yohaani **ni** umunyeshuúre
John *ni* 1.student
'John is a student '

b. n-**ri** umunyeshuúre
1SG.S-*ri* 1.student
'I am a student.'

These two restrictions, I claim, fall under the same generalization: predication with *ni* is banned from contexts with INFL, (53). The tense restriction, where T = INFL, leads straightforwardly to this conclusion. Under the assumption that person-features are licensed by functional material (e.g., Béjar and Rezac 2003), the restriction of *ni* from contexts with Speech Act Participant (SAP, i.e, first- and second-person) subjects also follows from a distributional restriction under T.

(53) **Generalization on the distribution of *ni***
*[T_{INFL} *ni*]

Finally, embedded clauses must use *-ri*. This can be seen in 54. To account for this, I assume that the complementizer is selectionally restricted to TP.

(54) **Matrix vs. Embedded clauses**

a. Umurwa mukuru wa u-Bu-rúundi **ni** Gitega.
capital.city of 14.Rundi *ni* Gitega
'The capital city of Burundi is Gitega.' (Matrix specificational)

b. * N-a-vug-ye kó umurwa mukuru wa u-Bu-rúundi **ni** Gitega.
1SG.SM-PST-say-PFV C capital.city 3.of 14.Rundi *ni* Gitega
'I said that the capital city of Burundi is Gitega.' (Embedded specificational)

c. N-a-vug-ye kó umurwa mukuru wa u-Bu-rúundi **u-∅-ri** Gitega.
1SG.SM-PST-say-PFV C capital.city 3.of AUG-14-rundi 3SM-PST-*ri* Gitega
i. 'I said that the capital city of Burundi is Gitega (the city).' (Embedded spec.)
ii. 'I said that the capital city of Burundi is in Gitega (the province).' (Locational)

In this way, the restriction to matrix, present-tense predication of third-person subjects can be tied together to the obligatory presence of T. Before turning to the final factor determining the distribution of *ni*, I will spell out two possible choices for implementing this generalization. The first, taken by e.g. Zentz (2016b), is to maintain that *ni* is a copular verb (categorically a *v*, see Mikkelsen 2005, 2011) and stipulate the absence of TP as a lexical property of the *ni*. While this is adequately able to capture the generalization in (53), it does so without leveraging the distinct properties of *ni*, noted for Kirundi in early work by Meeussen (1959: p. 180-6): the inability to convey temporal information.

The second analytical option, which I will be adopting and arguing for here, is to analyze *ni* and *-ri* as categorically distinct heads: the former is a non-verbal predicator Pred and the latter is the verbal instantiation *v* of predication in the absence of a lexical verb (Bowers 1993, 2002; Adger and Ramchand 2003).[13] When there is independent need to project verbal inflectional material in order to convey temporal information, to license person-features, or to be selected by a higher complementizer, Pred is ruled out. As Pred is categorically non-verbal, it does not permit the projection of further categories in the verbal extended projection (Grimshaw 2000). In contrast, *v*, which is spelled out as the verbal copula *-ri* in the absence of a lexical verb, is unremarkable among verbs in Kirundi in requiring inflectional structure.[14] What *does* have to be stipulated in this account is that the minimal amount of structure is used. I leave investigating this final point to the future.

Finally, consider the data below, which suggests that PredP is perhaps not the same thing as small clause predication, taking this latter term to mean non-verbal predication without additional morphosyntactic material. Non-verbal predication, as seen above, must occur with one of the two copular elements, *ni* or *-ri*. Similarly, in embedded contexts such as (55), *-ri* must be present, since *ni* is independently ruled out.

(55) N-i-baz-a kó Yohaáni *(a-ri) u-mwígīsha.
1SG.SM-RFLX-ask-IPFV COMP 1.Yohani 1SM-COP AUG-1.teacher
'I think that Yohani is a teacher.'

However, non-verbal predication without any additional morphosyntactic material is licit in fronting constructions discussed here. For example, consider the sentence in (56). Crucially, compare the post-focal nominal predicate, which does not occur with the augment vowel, with the the post-predicative post-posed nominal, which obligatorily occurs with the augment vowel. A discussion of the role of the augment vowel would take us too far astray, but Kirundi nominals in argument position uniformly require augments. I take the former (56a) to be small-clause predication, whereas the latter (56b) contains a right-dislocated subject.

(56) **Small clause predication**

a. Ni Yohaáni mwígīsha
NI 1.Yohani 1.teacher
'It's Yohani who the teacher is.' (Small Clause Predication)

b. Ni Yohaáni, u-mwígīsha
NI 1.Yohani AUG-1.teacher
'It's Yohani who the teacher is.' (Right-dislocated subject)

(57) Ni Yohaáni a-ri *(u)-mwígīsha
NI 1.Yohani 1SM-COP AUG-1.teacher

[13] One alternative way to maintain the syntactic-category distinction between the two heads, and perhaps closer to the proposals cited here, is to suggest that these two categories, Pred and *v*, contextual allomorphs, where Pred=*v* when it occurs with a VP complement, and Pred=Pred when it occurs with a non-VP complement. I will avoid for reasons that will become clear in the following section: locational PP complements also require verbal inflection. This would require additionally stipulating that PPs are somehow "verbal enough" to count as triggering the *v* allomorph or to posit null verbal material.

[14] The lack of aspectual information can also be tied to the lack of lexical verb, by analyzing the aspectual suffix (the final suffix on the verb) as an instance of InnerAspect (Travis 2010), also plausibly part of the verbal extended projection. The lack of aspectual contrast was also noted by Meeussen (1959: p. 184): "Le thème -ri, *être*, n'a pas de finale, et ne présente donc pas la distinction d'aspect (imperfectif: perfectif)"

'It's Yohani who the teacher is.' (Copular clause in cleft)

To the extent that this contrast is really indicative of small clause predication, which I leave motivating to future research, this initial data provides preliminary indication that PredP predication is not necessarily the same as small-clause predication in Kirundi.

4.2 Non-verbal eventuality-denoting predicates

The final factor determining the distribution of *ni* is whether the predicate is a locational PP. Locational predicates and embedded clauses require the verbal copula.[15]

(58) **Locational (PP) predicates**

a. inká **i-ri** mu murima
9.cow 9SM-*ri* in 3.field
'The cow is in the field.'

b. * inká **ni** mu murima
9.cow *ni* in 3.field
Intended: 'The cow is in the field.'

I will argue here that this fact provides one final piece of empirical evidence that non-verbal Pred *ni* lacks verbal inflectional structure. The analysis central to this argumentation recalls a similar set of facts from Scottish Gaelic, so I will first outline the analogous facts from Adger and Ramchand (2003). The conclusion to be drawn from both sets of data are that certain predicates carry an inherent eventuality argument by virtue of their meaning: they denote spatio-temporality delimited eventualities.

Scottish Gaelic has two relevant copular constructions, termed the Substantive Auxiliary Construction (SAC) and the Inverted Copular Construction (IAC) in Adger and Ramchand (2003). The two constructions are analyzed as involving two distinct copular elements: SACs include a head Pred/*v* which binds an eventuality variable introduced by the complement, and IACs are formed from a Pred head that does not.

(59) **Scottish Gaelic copular clauses**

a. Tha Calum faiceallach/anns a'bhùth
be.PRES Calum careful/in the.shop
'Calum is (being) careful/is in the shop.' (Substantive auxiliary construction)

b. Is mòr an duine sin
COP.PRES big that man
'That man is big.' (Inverted Copular Construction)

Adger and Ramchand (2003) differentiate between the two properties by positing two distinct Pred/*v* heads. Crucially, while these heads alternate and have distinct effects on the downstream derivation, they participate in fundamentally the same clausal structure as each other. So far in this section, I have argued for a more radical distinction between non-verbal predication mediated via Pred and verbal predication mediated via *v*. The restrictions on the distribution of *ni* seen in the previous subsections suggest

[15] A similar distinction can be found in Scottish Gaelic where APs also pattern like PPs and VPs, Adger and Ramchand 2003

a clausal structure lacking TP. Here, I will argue that the inability of predication with *ni* to bind an eventuality argument shows a further distinction between the two predication strategies.

Similar to Gaelic Inverted Auxiliary Constructions, Kirundi *ni* predication cannot be used with PP predicates, as seen above in (58); unlike Gaelic IACs, Kirundi *ni* predication can be used with adjectival predication. I take this last point to be language-specific differences in the syntax of adjectives, which are already quite a small, closed class in the Bantu languages.

(60) **Gaelic IACs are banned with adjectival and PP predicates**

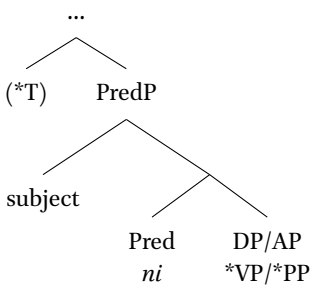
a. * Is an duine sin mòr
COP.PRES that man big
'That man is big.' (adjectival predicate)

b. * Is an cù leamsa
COP.PRES that dog with-me
'That dog belongs to me.' (PP predicate)

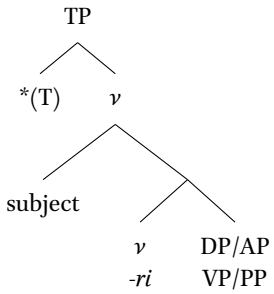
Given this similarity, I take Kirundi PPs to be a non-verbal category introducing an eventuality variable, noting their locational semantics. If Pred and *v* do indeed differ in their ability to bind the eventuality argument in their complement domain as proposed by Adger and Ramchand (2003), then the inability of Pred to occur with PP predicates follows straightforwardly.

In the final analysis, I propose that the structural configuration in non-verbal predication is a highly reduced one. In the absence of independent need to include inflectional structure, Kirundi permits a matrix clause consisting only of a PredP. If inflectional structure *is* required, *ni* is ruled out since it is not syntactically compatible with INFL.

(61) **Non-verbal predication with non-verbal syntax**



(62) **(Non-verbal) predication with verbal syntax**



This section provides additional support for the claim that Kirundi has a clause type that consists solely of the predicational core, PredP. While there are several restrictions on this predicational strategy, these restrictions can be tied to independent properties of tense and person licensing, as well as the binding of syntactic eventuality variable. Having made my argument for the structure of the cleft clause in the previous section, and supporting this proposal with a discussion of an independent context for matrix PredPs in Kirundi, I will turn now to evaluating this against possible alternatives.

5 The non-uniform structures of clefts cross-linguistically

In this final section, I compare the proposal presented above with alternative analyses made for other languages, both within Bantu and more widely. I will consider two families of analyses: in §5.1, I look at

analyses where *ni* is a left-peripheral head (as in Rizzi 1997), concluding that it cannot be identified with either Foc, Top, nor a head intermediate to the two, despite various proposed solutions; in §**??**, I consider the usual analyses for elements like *ni*, where it is treated analogously to English *be*, that is to say a *verbal copula*, concluding that this analysis predicts a much broader distribution of *ni* than we observe.

While I ultimately conclude that these analyses are not adequate for Kirundi, these alternative analyses do provide the basis for a set of typological observations that structures unified under the term "cleft" or which have been called "cleft-like" in their interpretation may in fact share some common structure. In line with these observations, which I present in §5.2, I propose that the structural configurations arising in Hungarian pre-verbal focus, the English cleft, and the Kirundi cleft share a common $\bar{\text{A}}$-fronting strategy, but diverge in two ways: firstly, in the lexical specification of the C head hosting the $\bar{\text{A}}$-fronted as either a matrix or obligatorily embedded clause, as discussed above in §3.2.2; and secondly, (in the latter case) the *verbality* of the embedding material, as discussed above in §4. I end by explicitly outlining the typology arising from these two parameters.

5.1 Non-cleft analyses of $\bar{\text{A}}$-fronting: *ni* as a left-peripheral head

One widely-adopted analysis of focus fronting constructions with similar surface properties to the Kirundi data under discussion here is the left-peripheral analysis proposed by Rizzi (1997). Under this analysis, there are three logically possible candidates for the syntactic identity of *ni*.[16] I will argue that each either fails to capture the full range of empirical data or requires non-trivial theoretical mechanisms in order to do so. The primary empirical challenge I will address on here is word-order, the distribution of *ni*, and the constituency with the following element.

5.1.1 Kirundi *ni* is not Foc

One influential analysis of phrasal fronting to a left-peripheral position is the FocusP analysis. This class of analyses typically takes the form of a Rizzian Left Periphery, wherein there is a dedicated and fixed position in the upper domain of the clause which hosts focused material (construed here as bearing a Focus feature, see Rizzi 1997). In this subsection, I will discuss the data motivating this proposal in some depth, and show that the Kirundi data does not bear out some crucial predictions regarding the surface realization of the struture (as already noted in other Bantu languages). I will then discuss three approaches that propose solutions which would permit a FocP-analysis to be maintained: the Head Adjucntion/Undermerge solution (Schwarz 2003; Yuan 2017a,b), the Q-particle solution (see Cable 2007; Branan and Erlewine 2022), and the multiple-Focus-head solution (Abels and Muriungi 2008). Regarding the first two, I present a counterargument from the lack of evidence that the derived constituency between [*ni* FOCUS] holds in Kirundi. On the final solution, I argue that there is no evidence that Kirundi's *ni* has the range of functions warranting the increased complexity of system with multiple Focus projections. I conclude here that the functions the *ni* does have strongly suggests a non-left-peripheral analysis, as presented in above.

The Rizzian approach has garnered empirical support from its application to languages with overt material accompanying this movement, such as Gungbe where the particle accompanying fronting is analyzed as an overt lexicalization of the Focus head (Aboh 2016). This is illustrated in (63) with the post-focus-constituent particle *wɛ̀*, glossed as FOC.

[16] I will ignore the possibility that *ni* is a Force or a Fin head, as they predict incorrectly *ni*-Topic and Focus-*ni* as word orders.

(63) **Gungbe focus fronting (Aboh 2016)**

a. ménù wè ɖà lésì ná Àlúkú sɔ̀?
who FOC cook rice to Aluku yesterday
'Who cooked rice for Aluku yesterday?'

b. étɛ́ wè Súrù ɖà ná Àlúkú sɔ̀?
what FOC Suru cook rice to Aluku yesterday
'What did Suru cooked rice for Aluku yesterday?'

Languages like Gungbe instantiate what appears to be the "ideal" Left Periphery in terms of overtness and word-order: the XP-particle word order seen in (63) permits a straightforward analysis as a Spec-head configuration. Other languages with similar structures include Wolof (Klecha and Martinović 2015; Martinović 2021a), Hausa (Green 2007), and the Bantu language Kinande (Schneider-Zioga 2007), illustrated in (64).

(64) **Kinande focus fronting (Schneider-Zioga 2007: p. 412)**

a. ekitabu *(kyo) Kambale a-asoma
book$_j$ that$_{focus-j}$ Kambale AGR-read
'(It's) the book (that) Kambale read.'

b. Georgine yo Kambale a-alangira
Georgine$_j$ that$_{focus-j}$ Kambale AGR-saw
'(It's) the book (that) Kambale saw.'

Kinande is a particularly interesting case in the present discussion, since it suggests a possible line of analysis where the particle *ni* that we have been investigating is the copula which accompanies fronting. Schneider-Zioga shows that the above focus fronted examples are not clefts, which are morphosyntactically distinct in two ways: firstly, clefts include the copula *ni* (analogous to Kirundi) and also require an augmented agreeing word glossed as *that* in (64).

(65) **Kinande clefts are morphosyntactically distinct (Schneider-Zioga 2007: p. 420)**

a. ni-ki ekyo Kambale a-agula
be-what that Kambale AGR-bought
'What is it that Kambale bought?' (Kinande cleft)

b. ekitabu ekyo Kamable a-agula
book that Kambale AGR-bought
'the book that Kambale bought' (Kinande relative clause)

The cleft structure in Kinande seen here is similar on the surface to the Kirundi data presented above. There are crucial differences, however, between Kinande and Kirundi. Most strikingly Kirundi lacks the evidence that unifies clefts and relative clauses (the similarity between (65a) and (65b)), as well as the evidence that differentiates between focus-fronting and cleft constructions (the dissimilarity between (64) and (65a)). In other words, it appears that what looks like two distinct possible structures in Kinande is realized in a single configuration in Kirundi. One may speculate whether the differences here may be tied to the different paths of grammaticalization taken from similar material; I will not pursue this here. Nonetheless, the Kinande data shows us that the FocP account can adequately capture instances of non-cleft focus-fronting in Bantu languages which have it. Kirundi, I have argued, does not.

More to our point, the differences between Kinande and Kirundi show that across Bantu, there is substantial variation with respect to how focus is mapped onto the syntax: variation in the inclusion of a dedicated focus structure, and whether the accompanying (focus) particle precedes or follows the fronted phrase. On this last point of variation, the former configuration is unproblematic for the FocP analysis, but the latter poses linearization issues. For example, consider the Kikuyu data in (66), where the focus marker (FM) *ne* precedes fronted *wh*-words or foci, analogously to Kirundi.

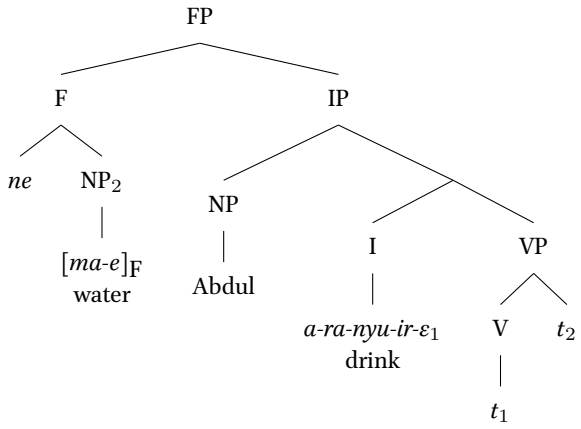
(66) **Kikuyu focus fronting (Schwarz 2003: p.54)**

a. ne ma-e Abdul a-ra-nyu-ir-ɛ
FM 6-water A SM-T-drink-ASP-FV
'It is water that Abdul drank.'

b. ne-kee Abdul a-ra-nyu-ir-ɛ
FM-what A SM-T-drink-ASP-FV
'What did Abdul drink?'

For analyses committed to the FocP analysis, where *ni* is analyzed as the Focus head, despite the linearization issues which arise, two solutions have been proposed to properly linearize the fronted phrase and the focus marker. The first solution proposed by Schwarz (2003: p. 86) for the Bantu language Kikuyu is to adjoin the fronted phrase to the Foc head. As already noted by Schwarz, this proposal is attractive insofar as "it can be fitted into a framework where this really adds to the explanatory adequacy of the account, and does not just 'get the word order right'" (p. 86). This analysis is presented below in (67). I will discuss two means to generate this configuration: Undermerge and Pied-piping of a low FP.

(67) **Structure for (66a), following Schwarz (2003: p. 86)**



The difficulty with adopting this account here rests in motivating this phrasal head-adjunction more broadly. There is not much discussion in Schwarz (2003), and this proposed account may be troubling for a theory wherein the Extension condition is active. This point is explicitly taken up and the analysis is defended against such trouble in Yuan (2017a,b), where an implementation of this is developed and discussed within the context of Undermerge. In addition, Yuan notes that this movement is motivated as the overt instantiation of Focus association by covert movement (Wagner 2006; Erlewine and Kotek 2018). However, note that the resulting structural configuration still predicts constituency between *ni* and the post-*ni* XP. As we have seen above, this constituency is not motivated for Kirundi. Under the unified hypothesis above, where I argued that the copular and the focus use of *ni* are instances of the

same Pred structure, there are selectional restrictions on *ni* which fail to arise in focus constructions. Recall that in Kirundi, prepositional phrases cannot form predicates with *ni*, as seen in (68a). However, PP foci are fully grammatical, seen in (68b). Under the constituency proposed above, this contrast remains unexplained.

(68) a. i-n-ká ***ni/i-ri** mu mu-rima.
AUG-9-cow NI/9SM-*ri* in 3-field
'The cow is in the field.'

b. Ni **kw' isoko** n-a-gīye ___ [kubēra n-kenér-ye umukâté].
NI to.store 1SG.SM-PST-walk.PFV because 1SM-need-PFV bread
'It's to the store I went because I need bread.'

These restrictions in the copular cases were tied to the requirement that verbal functional material be present; under the Undermerge approach, it is unclear why these restrictions on the complement to *ni* are relaxed, unless we abandon the unifying hypothesis and concede that the two functions of *ni* are in fact unrelated.

A plausible alternative means of deriving a similar structural configuration as (67) without positing this movement is to analyze *ni/ne* as being generated above the fronted (nominal) phrase itself and that this phrase triggers fronting to a (null) focus phrase (see Cable 2007, 2010; Branan and Erlewine 2022 for particle-based approaches to pied-piping; Yuan 2017a,b for an extension to Kikuyu). While I argue that this view is ultimately not motivated for Kirundi, one language where this may be the case is the Grassfields Bantu language Medumba (Keupdjio 2020).

(69) **Medumba focus particle (Keupdjio 2020: p. 17-18)**

a. Wàtét nɔ́ʔ sʷɛ̀n [á Nùŋgɛ̀]
Watat AUX.T2 sell FOC Nuga
'Watat betrayed Nuga$_{FOC}$.'

b. [á Nùŋgɛ̀] Wàtɛ́t nɔ́ɔ̀ʔ ⁿ-sʷɛ́ɛ̀n lá
FOC Nuga Watat AGR.AUX.T2 N-AGR.sell C.-Q
'Nuga$_{FOC}$ Watat betrayed.'

Note that the focus particle *á* appears when the focus is in-situ as well as when in is fronted. Taking the Medumba data to instantiate evidence for particle-based pied-piping, we can note some of the analytical presuppositions that such an account carries with it. While potentially unproblematic to the constellation of facts in Medumba, such a view is ultimately untenable for Kirundi.

The main challenge faced when adopting either the head-adjunction view or the particle-based pied-piping view for Kirundi (and indeed for Kikuyu) is that there is no independent motivation for the constituency between the *ni* and the immediately following phrase as seen in (68a). Furthermore, while neither Kirundi nor Kikuyu maintain the particle for in-situ foci, Medumba does, further suggesting a base-generated constituency of the particle and the following nominal in Medumba, but not in Kirundi.[17]

(70) * Yohaáni **a-a-som-ye** [ni igitabu]
Yohani 1SM-PST-read-PFV NI 7.book
Intended: 'Yohani read A BOOK.' (Kirundi)

[17] A further question is raised for Medumba, specifically on why the particle does not trigger movement uniformly. See Keupdjio 2020 for a comprehensive discussion and analysis.

In sum, maintaining a view where the single focus projection hosts the fronted phrase in its specifier leads to the inability to capture the correct word order without admitting one of two additional mechanisms: phrasal-adjunction to the focus head, or generating the Foc-particle directly above the phrase to be fronted, neither of which is descriptively adequate for the Kirundi data.

A second solution to the challenge faced by a commitment to the left-peripheral approach is given by Abels and Muriungi (2008) for the Kîîtharaka (Bantu). This solution makes use of an articulated focus structure within the left-periphery, containing thee focus projections. In addition to this added complexity within the high functional sequence, this analytical move requires a spell-out algorithm which determines which head the single exponent is spelled out in. In Kîîtharaka, the focus marker shows up in three distinct contexts: in pre-predicative position for predicate focus (71a), pre-nominal position for fronted foci (71c), and finally in contexts of successive-cyclic movement (71b).

(71) **Kîîtharaka focus marker (Abels and Muriungi 2008: p. 690, 716)**

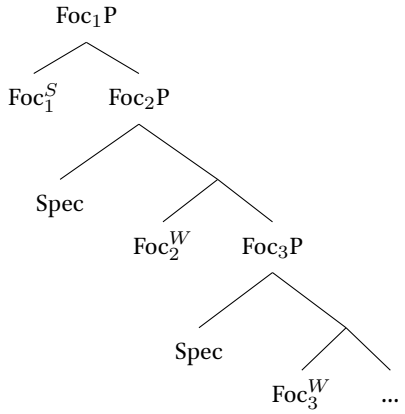
a. Maria n-a-gûr-ir-e î-buku
1.Maria FOC-1.SM-buy-PFV-FV 5-book
'Maria bought a book.'

b. I-mbi$_1$ John (*n)-a-ug-ir-e [Pat *(n)-a-ug-ir-e [Maria
FOC-what 1.John FOC-1SM-say-PERF-FV 1.Pat FOC-1SM-say-PERF-FV 1.Maria
*(n)-a-gûr-ir-e ___$_1$]]
FOC-1SM-buy-PERF-FV
'What did John say Pat said Maria bought?'

c. N-Aana a-gûr-ir-e î-buku
FOC-1.Ana 1.SM-buy-PFV-FV 5-book
'Maria bought a book.'

Their analysis of these posits a cline of Foc heads, where pre-predicative focus marking contains a single Foc head, successive-cyclic marking structures have two, and pre-nominal focus marking has all three. A positioning algorithm, which crucially requires reference to the "strength" of the heads, marked with diacritics S or W in (72), is used to restricted the spell-out of *n-/i-* to either the lowest or highest Foc heads. In other words, in a sequence containing only Foc2 and Foc3, Foc3 is spelled out and Foc2 is null; in a sequence containing all three, Foc1 is spelled out due to its strength.

(72) **Three-headed Approach (Abels and Muriungi 2008: p. 721)**



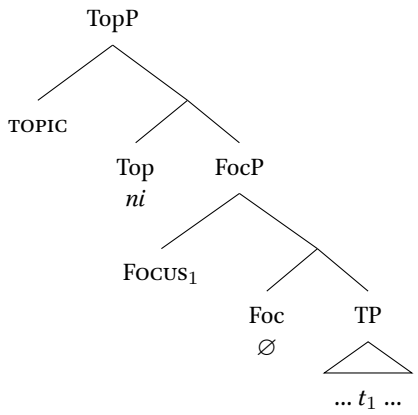
While the three heads line-up with the three contexts where FOC is used (pre-nominal focus, pre-predicate focus, and successive cyclic movement), this one-to-one mapping comes at the expense of substantially increasing the complexity of the left-periphery. It is unclear how such an account will transfer onto languages like Kirundi, where the language has a much more distributionally restricted particle, and furthermore the consequences for such a proliferation of Focus heads has for the cross-linguistic picture more broadly.

Having ruled out the proposed solutions for *ni* being a Focus head, I will turn to two other approaches that maintains a left-peripheral syntax for *ni*, but proposes identification with another head.

5.1.2 Kirundi *ni* is not Top

A plausible alternative structure which trivially derives the correct surface word-order is to assume that *ni* is a Top head, selecting the FocP which hosts the $\bar{\text{A}}$-fronted constituent. As Kirundi does not have any overt segmental material co-occuring with this movement, the Foc head is phonologically null.[18] The hypothetical structure is given in (73). While this structure trivially derives the word-order seen in Kirundi, I will go through two arguments against this view briefly here.

(73)



The structure in (73) incorrectly predicts two facts about the distribution of *ni*: firstly, it predicts that Topics obligatorily co-occur with *ni*. In fact, the opposite is true: Topics never occur with *ni*, unless there is also an element in the post-*ni* position. The example in (74b) is ungrammatical under the parse that *ni* and Yohani is not $\bar{\text{A}}$-fronted. Compare this with (75a), which shows that the identical string is grammatical, however, when the post-verbal position of *ni* is filled with the $\bar{\text{A}}$-fronted subject. The position may also be filled with a pronoun co-referential with the topic, as seen in (75b).

(74) **Topic and *ni* do not obligatorily co-occur**

a. Ico gitabu, Yohaáni a-a-*(gi)-som-ye
7DEM 7.book Yohani 1SM-PST-7OM-read-PFV
'This book, Yohani read it.

b. * Ico gitabu ni, [Yohaáni a-a-gi-som-ye]
7DEM 7.book NI Yohani 1SM-PST-7OM-read-PFV
Intended: 'This book, Yohani read it.

(75) ***ni* must be followed by "focused" constituent**

[18] Languages that do include overt phonological material were cited above, such as Gungbe and Kinande.

a. Ico gitabu [ni Yohaáni$_1$ [t_1 a-a-gi-som-ye]]
7DEM 7.book NI Yohani 1SM-PST-7OM-read-PFV
'This book, it's Yohani who read it.

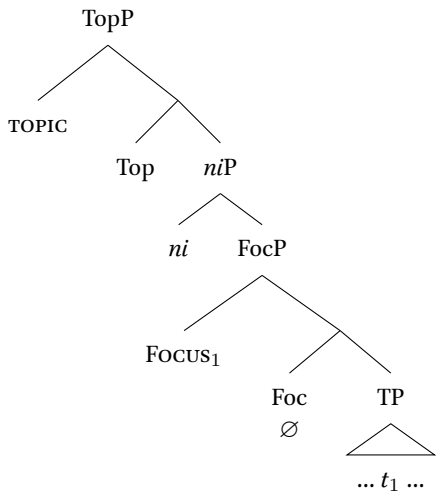
b. Ico gitabu$_1$ [ni co$_1$ [Yohaáni a-a-gi-som-ye]]
7DEM 7.book NI 7.PRON Yohani 1SM-PST-7OM-read-PFV
'This book, it's that which Yohani read it.

Secondly, there is no immediate way to capture the obligatory co-occurence of *ni* with fronted constituents (to Spec,FocP). The sole way to ensure the obligatory surfacing of *ni* with fronted constituents is to stipulate the requirement that TopP is projected whenever FocP is, even when the specifier of TopP is phonologically and possibly syntactically empty. I will take these two challenges and the stipulations required to surmount them, to be indicative of such an accounts inadequacy for Kirundi.

5.1.3 Kirundi *ni* is not an intermediate left-peripheral head

One final option is to assume that *ni* is a head intermediate to the Foc and Top heads. While I will remain agnostic on the function this putative head may have, see Wasike (2007) for the view that this left-peripheral head has a predicative function *within* the left-periphery.[19] This alternative structure is illustrated in (76), where *ni* is simply a head intervening between Top and Foc.

(76)



This account faces two challenges. Firstly, as in the previous Topic head analysis, there is no immediate way to capture the obligatory co-occurence of *ni* with fronted constituents. While the account proposed above ties the obligatory *ni* to the fact that the CP hosting the fronted constituent is obligatorily embedded, the account of *ni* as an independent head in the left-periphery requires independent motivation; I do not see a clear, motivated way to tie the projection of FocP to the obligatory projection of a distinct head.

Secondly, there is no straightforward way to capture the observation that embedded clefts are structurally more complex. Consider the data in (77–78), where a cleft is embedded under a matrix verb. In

[19] While I ultimately propose a similar analysis, where *ni* is identified with Pred, the crucial difference is the syntax of this head. Wasike (2007) places it squarely in the left-periphery, thus deriving a mono-clausal structure, whereas my use of the label Pred is analogous to *v*, which thereby delimits the boundary of the lower clause without requiring the root clause to be syntactically headed by a verb.

such a case, *ni* is not permitted, and instead the cleft is formed with the copular verb *-ri*. As a result, a tense contrast is possible in this context, as expected from the discussion of non-verbal predication in §non-v-pred-sec above. Furthermore, since the copula agrees with the φ-features of its subject, we can verify in (78b) that the fronted constituent is not the subject: the copula has class 1 agreement rather than first-person-singular agreement.

(77) **Embedded clefts use the copula *-ri***

a. Yohaáni a-a-vug-ye [kó a-ri Kagabo a-a-som-yé igitabu]
Yohani 1SM.PST-say-PFV COMP 1SM-COP Kagabo 1SM.PST-say.EMB-PFV 7.book
'Yohani said that Kagabo read the book.'

b. …kó a-a-ri Kagabo a-a-som-yé igitabu
…COMP 1SM-PST-COP Kagabo 1SM-PST-say.EMB-PFV 7.book
'Yohani said that Kagabo had read the book.'

(78) **Embedded clefts do not agree with post-copular constituent**

a. …kó a-ri jēwé n-a-som-yé igitabu
…COMP 1SM-COP 1SG.PRON 1SG.SM-PST-say.EMB-PFV 7.book
'Yohani said that I read the book.'

b. * …kó n-a-ri jēwé n-a-som-yé igitabu
…COMP 1SG.SM-COP 1SG.PRON.STR 1SG.SM-PST-say.EMB-PFV 7.book
'Yohani said that I read the book.'

Note the word order for these is strict. When the fronted constituent appears before the copula and in (79), the requirement to have a filled post-copular position holds, as in matrix clauses: it must be filled by "focused" material. Once more, the agreement on the copula demonstrates that the pre-copular constituent is not the subject: that is, in (79b).

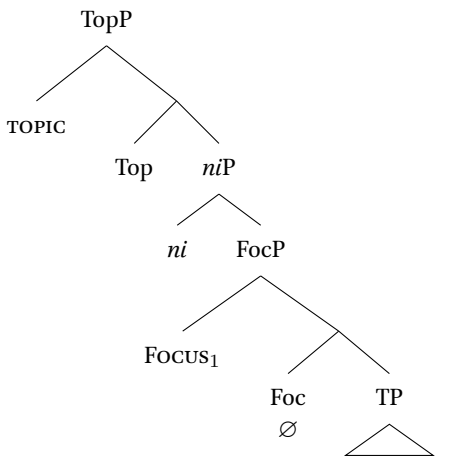
(79) **Embedded clefts have an expletive subject**

a. …kó Kagabo a-ri *(we) a-a-som-yé igitabu
…COMP Kagabo 1SM-COP 1.PRON 1SM.PST-say.EMB-PFV 7.book
'Yohani said that, Yohani, he read the book.'

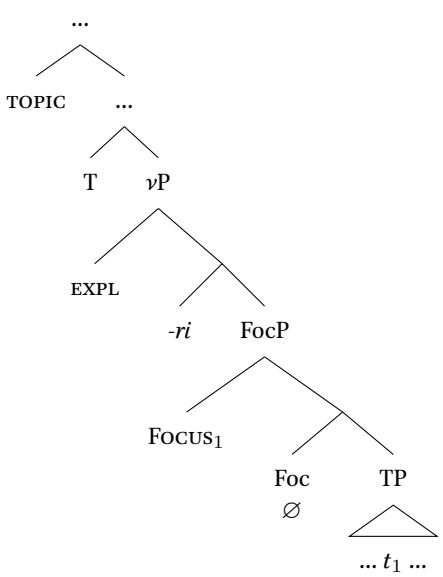
b. …kó jēwé a-ri *(je) n-a-som-yé igitabu
…COMP 1SG.PRON.STR 1SM-COP 1SG.PRON 1SG.SM.PST-say.EMB-PFV 7.book
'Yohani said that, me, I read the book.'

To capture such data, the left-peripheral analysis must propose that the alternation between *ni* and *-ri* is in fact an alternation in the upper portion of the left-periphery.

(80) a.



b.



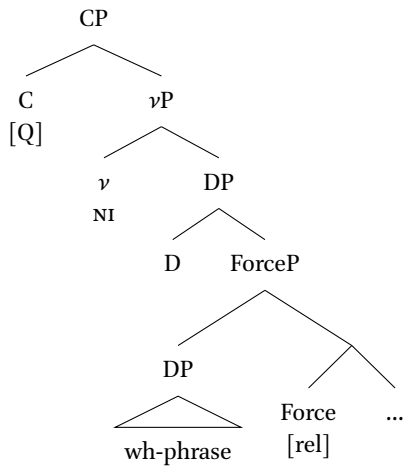
The one substantial difference between the two structures is that the left-periphery of the sole clause in structures with *ni* (80) is smaller than the left-periphery of the lower clause of the structure with *-ri* (80b). It is unclear to me at this moment whether this makes crucially different empirical predictions when compared to the cleft analysis I have proposed. On the basis of the similarity between the functions of *ni* and *-ri*, however, I believe the uniform cleft-structure I have proposed can capture both the alternation in embedded clauses as well as the same alternation in non-verbal sentences in a more natural way, by tying it to the verb-hood of the copula and the non-verb-hood of *ni* as revealed through their distribution across the two constructions.

5.1.4 Verbal-copula analysis

I would now like to consider a family of analyses which are substantially similar to the one argued for above, but which differ in one crucial way: the syntactic nature of the copula. These approaches take the copular use of the particle to be demonstrative of their syntactic role as a verbal element in non-verbal predication. I show here that the implicit assumption that the bi-clausality (which goes hand-in-hand with what I have been calling the cleft analysis) entails bi-verbal is not a necessary one. I show that analyses which make this assumption for Bantu are faced with the need to stipulate some of the structural requirements of the cleft particle. I will once again consider two analyses within this general approach. In the first, the particle directly instantiates some verbal category, typically *v*. In the second, the particle accompanies a null copula. While both these analyses have the benefit of unifying non-verbal predication and cleft structures, they require additional stipulations to the latter which I argue are avoided under the proposal advocated for in the previous section.

Turning to the first analysis, we can consider the particle is an instantiation of *v*, which selects for a CP (among other constituents in non-verbal predication). The main downside to this analysis is that it requires a stipulation that a single element, *ni*, which is structurally a verb and which is inflectionally and distributionally restricted. The structure in (81) is a proposal along these lines for Shona wh-clefts by Zentz (2016b).

(81) ***ni* as a *v*P copular clause (after Zentz 2016b)**



This analysis is similar to the intuition that leads to familiar analyses of English clefts: a copular verb is used in clefts because it is semantically bleached. However, the nature of the copula in Kirundi and English differ substantially, in a way that is often noted and just as often disregarded. While the English copula is clearly verbal, sharing a subset of the distribution and inflectional possibilities of other lexical verbs in English, the Kirundi *ni* is non-verbal. Not only does *ni* obligatorily lack inflection, it has a distribution wholly unlike verbs in Kirundi. This asymmetry between *ni* and verbs is not captured in this analysis; neither is the highly restricted distribution of *ni*. On the other hand, to capture the lack of inflection with *ni*, Zentz (2016b: 199) proposes a reduced structure which is stipulated to lack INFL. In inflectional contexts, the difference in form is captured by contextual allomorphy rather than as inflection (Zentz 2016b: p. 160).

The second analysis can be seen as an extension of the FocP analysis to non-verbal predication. In these analyses, *ni* obligatorily accompanies a separate, null particle. While I will not discuss these analyses in detail, it is worth noting that several proposals have been forwarded which take this general shape though often without a full discussion. One proposal is forwarded by Abels and Muriungi (2008: p. 690fn) in the context of extending their multiple FocP analysis to non-verbal predication. In cleft-analyses, the *ni* might be analyzed as the instantiation of functional material in the absence of an overt lexical verb (as in bergvall 1987, cit. in schwarz-2003, p. 70f). For similar null-copula analyses, see Gibson et al. (2019), who discuss this in the context of cross-Bantu variation in copular (non-verbal predication) constructions.

To end this subsection, I will note that the challenges posed in this section for the verbal copula analysis are primarily ones that interrogate the implicit assumption that copulas are uniformly verbal. Even if not framed precisely in those terms, this implicit assumption is often borne out in the lack of clear correspondence between the claim that the Bantu copula is non-verbal and the structural representation it is given. Indeed, the structure given in (81) is unclear insofar as it represents (what I argue to be) a syntactically non-verbal element within a (defectively) verbal structural configuration. This lack of clarity, I suggest, obscures the phrase-structural implications of the insight that *ni* is a unique element in the grammar, which researchers have long noted about it in its various instantiations across the Bantu languages. The next subsection is dedicated to spelling out some of these implications.

5.2 Towards a more fine-grained structural typology: non-verbal clefts, verbal clefts, and non-clefts

The central claim made above regarding the lack of verbal material in the matrix clause makes a clear typological prediction with respect to the possible structures we might find in cleft constructions. Here, I will outline briefly the general characteristics of this structural typology, and tie this back to variation in what is sufficient to count as a clause in the language. This variation, I suggest, is ultimately related to a distinction made by Pustet (2003) between verbal copulas and (non-verbal) particle copulas, a distinction not taken-up in the generative literature as far as I am aware. As we saw in §4, the two strategies co-exist in Kirundi and are in complementary distribution

In languages such as English, the only available copula for main clause predication is a verbal copula *be* (though see den Dikken 2006 on *as* and *of* as similar elements in the nominal domain). In contrast to the English-type system, a wide literature has grown regarding multi-copular systems. A particularly well discussed example is the Spanish distinction between *ser* and *estar*; another is the Na-Dené language Tłı̨chǫ Yatıì (Welch 2012 describes this in therms of eventiveness; see also Adger and Ramchand 2003 on Scottish Gaelic; see Gibson et al. 2019 for an overview).

(82) **Spanish multiple copulas are both verbal (Maienborn 2005, among others)**

a. Maria **es** rubia
Maria is$_{ser}$ blond
'Maria is blond.'

b. Maria **está** rubia
Maria is$_{estar}$ cansada
'Maria is blond.'

(83) **Tłı̨chǫ Yatıì multiple copulas are both verbal (Welch 2012: p. 6)**

a. Ekwǫ̀ **elı̨**.
Ekwǫ̀ ∅-**lı̨**
caribou IMPF.3SBJ-COP1
'S/he/it is a caribou.' (e.g., a role in a play) (eventive predicate)

b. Ekwǫ̀ **hǫt'e**.
Ekwǫ̀ **ha**-**l**-t'e
caribou THM-IMPF.3SBJ-COP2
'S/he/it is a caribou.' (in a characterizing sense) (non-eventive predicate)

To my knowledge, this literature on multicopular systems is generally limited to languages which have multiple verbal copulas. In other words, these are languages with multiple copulas falling within the same "copularization" strategy (to use terminology from Pustet 2003). My claim in §4 was that the two copulas *ni* and *-ri* are structurally distinct: *-ri* instantiates a verbal head *v* whereas *ni* instantiates a non-verbal Pred. In other words, the two strategies (verbal and particle copularization) from Pustet corresponds to a structural distinction, one which has been relatively neglected in the generative literature but which I hope to demonstrate has predictive consequences for the theory of cleft structures.

In what follows, I will explore the typological predictions for cleft structures raised by this account, concluding that there are two parameters giving rise to three types of cleft possible structures. In sum, this proposal has the upshot of separating "cleft" as a descriptive term for a means of dividing the proposition into a salient information-structural partition (typically understood to be a focus-presupposition

bipartition (Chomsky 1971; Jackendoff 1972). The proposal takes this basic form, to be exemplified below: firstly, cleft clauses, which involve $\bar{A}$-movement, can be constructed in the main clause or in an embedded clause, giving rise to a distinction between mono-clausal and bi-clausal constructions. I take this to be particular to the lexical properties of the C-system in the language. Secondly, the independent availability of a non-verbal particle copula in the language results in a mono-verbal yet bi-clausal cleft. This proposal is summarized in (84).

(84) **Two parameter typology of cleft structures**

	Cleft clause is ...	
	Matrix clause	Embedded clause
No copula	**Mono-clausal focus** Hungarian, Wolof	N/A
Verbal copula	N/A	**Bi-verbal cleft** English
Particle copula	N/A	**Mono-verbal cleft** Kirundi

Consider firstly the mono-clausal focus construction, exemplified by Hungarian and Wolof. These constructions express a cleft-like interpretation but within a single clause. Each of these constructions is the result of $\bar{A}$-movement of one constituent into the left periphery of the clause, bolded in the examples below. The resulting construction need not be embedded, standing alone as a matrix clause.

(85) **Hungarian mono-clausal focus construction (É. Kiss 1998: p. 249)**

a. Mari **egy kalapot** nézett ki magának
Mary a hat.ACC picked out herself.ACC
'It was a hat that Mary picked for herself'

b. $[_{\text{TopP}}$ Mari $[_{\text{FP}}$ **[egy kalapot]**$_j$ nézett$_i$ $[_{\text{VP}}$ t_i ki magának $t_j]]]$

(86) **Wolof mono-clausal focus construction (Martinović 2021a)**

a. Man, Yusu Nduur la a gis
1S.STR Youssou N'Dour C_{Wh} 1SG see
'Me, it's Yousouu N'Dour that I saw.'

b. $[_{\text{TopP}}$ Man $[_{\text{CP}}$ **Yusu Nduur** la $[_{\text{IP}}$ a gis $]]]$

Consider now the bi-clausal structures of English clefts. I will roughly follow the FP/raising analysis of clefts proposed by É. Kiss (1998) (but see Hedberg 2000 for alternatives). Crucially, I will diverge from the proposal in É. Kiss (1998) slightly, assuming that the English FP is null and selected for by the copula rather than the copula being generated in F. Under this view, the cleft clause is derived by movement into a dedicated Focus phrase in the left periphery of the clause; the resulting structure is an embedded clause, which is obligatorily selected for by embedding material.

(87) **English bi-clausal, bi-verbal cleft (É. Kiss 1998)**

a. It was to John that I spoke

b. $[_{\text{IP}}$ It was $[_{\text{FP}}$ [to John]$_i$ F $[_{\text{CP}}$ that $[_{\text{IP}}$ I spoke $t_i]]]$

The available embedding material in English is limited to the verbal copula *be*. As a result of the requirements for matrix predicates in English to be tensed, and the syntactically verbal nature of the copula, a range of inflectional possibilities are available. Crucially, the resulting cleft is has two syntactically verbal parts: the cleft clause contains the semantically substantive predicate, and the matrix clause contains a semantically expletive verbal copula.

Finally, turning to the bi-clausal but mono-verbal cleft structure proposed for Kirundi, we have argued here that the structure of the cleft clause roughly mirrors the FP analysis above: a constituent is $\bar{\text{A}}$-moved into the left periphery of an embedded clause. The difference between the English and Kirundi structures arises as a result of the richer predicative strategies of Kirundi. In addition to having a verbal copula *-ri*, Kirundi has a non-verbal copula *ni* which can function as a matrix clause, albeit a syntactically deficient one. Specifically, it lacks tense entirely. Rather than stipulate this lack of inflection, this deficiency is instead tied to the non-verbal nature of *ni*; since it is not a verb, no part of the verbal functional structure is licensed (that is, INFL and C; see e.g., Grimshaw 2000 on extended projections).

(88) **Kirundi bi-clausal mono-verbal cleft**

a. Ni **igitabu**$_1$ [Yohaáni a-a-som-yé ___$_1$]
NI 7book Yohani 1SM-PST-read-PFV.REL
'It's THE BOOK that Yohani read.'

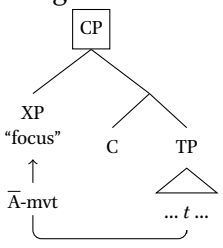
b. [$_{\text{PredP}}$ *pro* ni [$_{\text{CP}}$ **igitabu**$_1$ C [$_{\text{TP}}$ Yohaáni yasomyé t_1]]]

One final piece of evidence for this view is that, when the entire cleft construction is tensed, such as in the case of providing information that no longer holds, the verbal copula may be used. The resulting cleft is apparently uncommon, with speaker favouring marking tense on the embedded predicate, and is not discussed in previous work on Kirundi clefts. Nonetheless, insofar as it is acceptable, it crucially differs in that it requires inflection and agreement with a default class, presumably with a discourse-salient *pro* (rather than the clefted constituent).

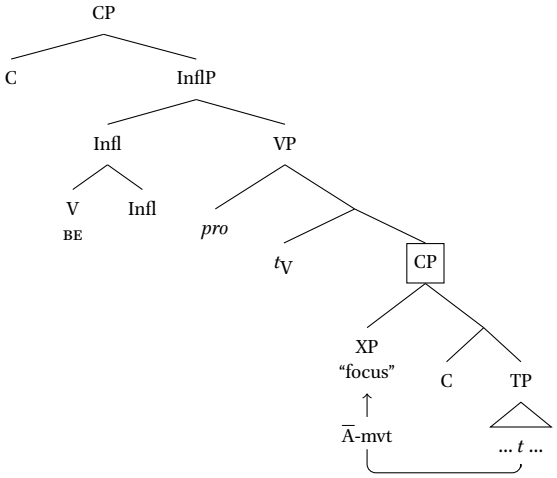
(89) a-a-ri **igitabu**$_1$ [Yohaáni a-a-som-yé ___$_1$]
1SM-PST-COP 7book Yohani 1SM-PST-read-PFV.REL
'It was THE BOOK that Yohani read.'

The upshot of this proposal, to reiterate, is that the term "cleft" must be understood as subsuming a wider class of structural configurations than previously recognized. That is, while "cleft" is a useful shorthand for a set of correspondences between a re-ordering of constituents relative to some information structurally neutral form and the information structural effects, the syntactic means languages have to achieve this correspondence is determined by language-specific lexical properties. Nonetheless, I propose that the constructions can be unified under the shared $\bar{\text{A}}$-movement structures, framed in the trees in (90).

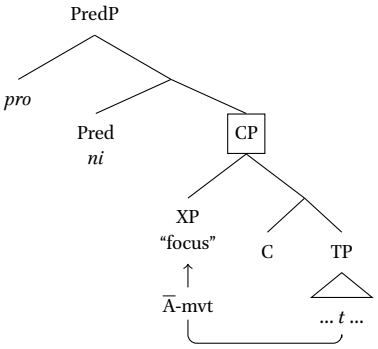
(90) a. **Hungarian monoclausal (="cleft-like") focus construction**



b. **English biverbal biclausal cleft**



c. **Kirundi monoverbal biclausal cleft**



As seen above, this typology rests on two (ultimately lexical) distinctions. I take this two-parameter typology be exhaustively exemplified by the languages represented here by Hungarian, English, and Kirundi. That is, Kirundi represents a third member of a typological system already implicit in the work of É. Kiss (1998).Firstly, the cleft/non-cleft distinction such as those traditionally made between English clefts and Hungarian pre-verbal focus are instead a result of whether the high functional structure involved in these constructions (say, FP) is a licit matrix clause in the language, or whether it must be embedded. I take this to be an independently needed lexical specification to rule out, for instance, matrix clauses headed by *that*. In English, the FP-headed clause is obligatorily embedded, whereas Hungarian FP-headed clauses need not be.

The second parameter is the verbality of the matrix clause, and is motivated here by the distinct properties of Kirundi clefts presented in this paper. The parameter makes use of the verbal/particle copula distinction from Pustet (2003) to show that, even for bi-clausal constructions, there are distinct structural possibilities. Ultimately, this rests on the language-specific availability for non-verbal structure (i.e., the particle-copula headed PredP) to be a matrix clause *without* additional functional structure surmounting it. Kirundi *ni*, I claimed, is an instance of such a language that permits this.

This final point on multiple copulas having distinct categorical specifications, furthermore, demonstrates that there are multiple ways to build a multicopular system. While the typical example of a multicopular system involves two clearly verbal elements (though these elements may be defective in some way), I argue that Kirundi shows that this is not the only state of affairs. Instead, multicopular systems might be the result of multiple instances of the same "copularization strategy" (i.e., Spanish has two verbal copulas), or the result of multiple "copularization strategies" entirely. This suggests the need to

revisit cases of languages multiple copulas in order to establish whether we are truly dealing with multiple members of the same syntactic kind, or with two different syntactic kinds.

6 Conclusions and outlook

In this paper, I have presented a novel analysis of Kirundi cleft constructions and non-verbal predication, each revolving around the central idea that they are both instances of a non-verbal root clause. More specifically, I have argued for the view that such non-verbal root clauses is syntactically highly deficient, lacking inflectional categories present in clauses with a verbal predicate. This proposal rests on the claim that this categorical distinction underlies the *ni* and *-ri* alternation, in both the cleft uses and particle copula contexts.

If this proposal is on the right track, I have shown that there are cleft-like constructions across languages fall into at least three types, but are unified by the common core of $\bar{\text{A}}$-movement deriving the classical bipartition between a 'focussed' element and the presupposed content. If the result of this movement is licit as a matrix clause, which I take to be determined by the language-specific lexical item which heads the CP to which the constituent $\bar{\text{A}}$-moves, then the result is a mono-clausal focus construction such as Hungarian or Wolof. If the CP is obligatorily embedded and the only embedding material in the language is syntactically verbal, then the cleft construction derived is bi-verbal, consisting to two articulated verbal extended projections. Finally, in the language has a syntactically non-verbal predicator like *ni*, then the result is still a bi-clausal cleft construction, but one in which the root clause is defective.

From this conclusion, there are two lines of further research that appear to be relevant. Firstly, there has been some debate on the mono-/bi-clausal status of similar constructions across the Bantu languages. I hope to have shown that the distinction is too coarse-grained, and needs to be separated into two separate questions: (i) whether the lower clause has the properties of a root clause or an embedded clause, and (ii) whether the root clause contains a fully articulated verbal extended projection. Secondly, I have relied on the notion of a particle copula from work by Pustet (2003), wherein it is shown that the term "copula" may in fact apply to structurally distinct items. In much generative work on these constructions, however, the implicit assumption appears to be that the copula is uniformly a verbal element, albeit a defective one. I hope to have made the case that this is not a necessary assumption, and that in fact there are empirical and analytical reasons to dispense with it for languages where, unlike with English *be*, the copula is morphosyntactically differentiable from the rest of the verbs in the language.

References

Abels, Klaus, and Peter Muriungi. 2008. The focus marker in Kîîtharaka: Syntax and semantics. *Lingua* 118:687–731.

Aboh, Enoch. 2016. Information structure: A cartographic perspective. In *The oxford handbook of information structure*, ed. Caroline Féry and Shinichiro Ishihara, 147–164. Oxford University Press.

Adger, David, and Gillian Ramchand. 2003. Predication and equation. *Linguistic Inquiry* 34:325–359.

Akmajian, Adrian. 1970. On deriving cleft sentences from pseudo-cleft sentences. *Linguistic Inquiry* 1:149–168.

Béjar, Susana, and Milan Rezac. 2003. Person licensing and the derivation of PCC effects. In *Romance linguistics: Theory and acquisition*, 49–62. Amsterdam/Philadelphia: John Benjamins.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591–656.

Bowers, John. 2002. Transitivity. *Linguistic Inquiry* 33:183–224.

Branan, Keynon, and Michael Yoshitaka Erlewine. 2022. Anti-pied-piping. Unpublished Manuscript, Leibniz Center for General Linguistics and the National University of Signapore.

Büring, Daniel, and Manuel Križ. 2013. It's that, and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites). *Semantics and Pragmatics* 6:1–29.

Cable, Seth. 2007. The grammar of Q: Q-particles and the nature of Wh-fronting, as revealed by the Wh-questions of Tlingit. Doctoral Dissertation, Massachusetts Institute of Technology.

Cable, Seth. 2008. Wh-fronting (in Hungarian) is not focus-fronting. Unpublished Manuscript, May 2008.

Cable, Seth. 2010. *The grammar of Q: Q-particles, Wh-movement, and Pied-Piping*. Oxford: Oxford University Press.

Chaperon, Brandon. to appear. Unravelling the 'not': left peripheral blocking of negation in Kirundi. *Toronto Working Papers in Linguistics* 45.

Chomsky, Noam. 1971. Deep structure, surface structure, and semantic interpretation. In *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*, ed. D. D. Steinberg and L.A. Jakobovits, 183–216. Cambridge: Cambridge University Press.

Diercks, Michael. 2010. Agreement with subjects in Lubukusu. Doctoral Dissertation, Georgetown University, Washington, D.C.

den Dikken, Marcel. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas*. Number 47 in Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.

É. Kiss, Katalin. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74:245–273.

Edenmyr, Niklas. 2001. Focus constructions in Kirundi. Master's thesis, Stockholm University, Stockholm.

Erlewine, Michael Yoshitaka, and Hadas Kotek. 2018. Focus association by movement: Evidence from tanglewood. *Linguistic Inquiry* 49:441–463.

Fominyam, Henry, and Radek Šimík. 2017. The morphosyntax of exhaustive focus: A view from Awing (Grassfields Bantu). *Natural Language and Linguistic Theory* 35:1027–1077.

Gibson, Hannah, Rozenn Guérois, and Lutz Marten. 2019. Variation in Bantu copula constructions. In *The Grammar of Copulas across Languages*, ed. Maria Arche, Antonio Fábregas, and Farael Marin, 213–242. Oxford: Oxford University Press.

Gluckman, John. 2022. Copulas and complementizers in Kinyamulenge. Paper presented at Syntax and Semantics of African Languages – SASAL 1, DGfS 2022, February 24 2022.

Green, Melanie. 2007. *Focus in hausa*. Hoboken, NJ: Wiley.

Grimshaw, Jane. 2000. Locality and extended projections. In *Lexical specification and insertion*, ed. Peter Coopmans, Martin Everaert, and Jane Grimshaw, volume 197 of *Current Issues in Linguistic Theory*, 115–133. Amsterdam/Philadelphia: John Benjamins.

Hartmann, Katharina, and Malte Zimmermann. 2012. Focus marking in Bura: semantic uniformity matches syntactic heterogeneity. *Natural Language and Linguistic Theory* 30:1061–1108.

Hedberg, Nancy. 2000. The referential status of clefts. *Language* 76:891–920.

Horvath, Julia. 2005. Is 'focus movement' driven by stress? In *Approaches to Hungarian 9*, 131–158. Szeged: JATEPress.

Horvath, Julia. 2007. Separating "focus movement" from focus. In *Phrasal and clausal architecture: Syntactic derivation and interpretation*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 108–145. Amsterdam/Philadelphia: John Benjamins.

Horvath, Julia. 2013. Focus, exhaustivity, and the syntax of wh-interrogatives: the case of Hungarian. In *Approaches to Hungarian 13: Papers from the 2011 Lund conference*, ed. Johan Brandtler, Valéria Molnár, and Christer Platzack, 97–132. Amsterdam/Philadelphia: John Benjamins.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jerro, Kyle. 2013. Argument structure and the typology of causatives in Kinyarwanda: Explaining the causative-instrumental syncretism. Master's thesis, The University of Texas at Austin, Austin.

Keupdjio, Hermann. 2020. The syntax of A'-dependencies in Bamileke Medumba. Phd dissertation, University of British Columbia.

Klecha, Peter, and Martina Martinović. 2015. Exhaustivity, predication and the semantics of movement. In *Proceedings of the BLS 41*, ed. Anna E Jurgensen, Hannah Sande, Specer Lamoureux, Kenny Baclawski, and Alison Zerbe, 267–286. Berkeley Linguistics Society.

Lafkioui, Mena B., Ernest Nshemezimana, and Koen Bostoen. 2016. Cleft constructions and focus in Kirundi. *Africana Linguistica* 22:71–106.

Maienborn, Claudia. 2005. A discourse-based account of Spanish: A discourse-based account of ser/estar. *Linguistics* 43:155–180.

Martinović, Martina. 2021a. Exhaustivity and predication: Non-discourse function of the left periphery. Paper presented at the Move Agree Forum, University of British Columbia/McGill University, June 4 2021.

Martinović, Martina. 2021b. Reversibility in specificational copular sentences and pseudoclefts. *Natural Language and Linguistic Theory* .

Meeussen, Achille Emille. 1959. *Essai de grammaire rundi*. Tervuren: Musée Royal du Congo Belge.

Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*, volume 85 of *Linguistik Aktuell/Linguistics Today*. Amsterdam/Philadelphia: John Benjamins.

Mikkelsen, Line. 2011. Copular clauses. In *Semantics: An international handbook of natural language meaning*, ed. Klaus von Heusinger, Claudia Maienborn, and Paul Portner, volume 2 of *Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science* [*HSK*], chapter 68, 1805–1829. Berlin: De Gruyter Mouton.

Ndayiragije, Juvénal. 1999. Checking economy. *Linguistic Inquiry* 30:399–444.

Nshemezimana, Ernest, and Koen Bostoen. 2017. The conjoint/disjoint alternation in Kirundi (JD62): A case for its abolition. In *The conjoint/disjoint alternation in bantu*, ed. Jenneke van der Wal and Larry M. Hyman, chapter 14, 390–425. Berlin: De Gruyter Mouton.

O'Neill, Teresa. 2019. The support copula in the left periphery. In *The Grammar of Copulas across Languages*, ed. Maria Arche, Antonio Fábregas, and Rafael Marín, 48–68. Oxford: Oxford University Press.

Pustet, Regina. 2003. *Copulas: Universals in he categorization of the lexicon*. Oxford: Oxford University

Press.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer.
Rochemont, Michael S. 1986. *Focus in Generative Grammar*, volume 4 of *Studies in Generative Linguistic Analysis*. Amsterdam/Philadelphia: John Benjamins.
Safir, Ken. 2019. The A/$\overline{\text{A}}$-distinction as an epiphenomenon. *Linguistic Inquiry* 50:285–336.
Schneider-Zioga, Patricia. 2007. Anti-agreement, anti-locality and minimality: The syntax of dislocated sujects. *Natural Language and Linguistic Theory* 25:403–446.
Schwarz, Florian. 2003. Focus marking in kikuyu. *ZAS Papers in Linguistics* 30:41–118.
Schwarz, Florian. 2007. Ex-situ focus in Kikuyu. In *Focus strategies in african languages*, ed. Enoch Aboh, Katharina Hartmann, and Malte Zimmermann, volume 191 of *Trends in Linguistics. Studies and Monographs* [*TiLSM*], 139–159. Berlin: De Gruyter Mouton.
Torrence, Harold. 2013a. The morphsyntax of Wolof clefts. In *Cleft structures*, ed. Katharina Hartmann and Tonjes Veenstra, 187–223. Amsterdam/Philadelphia: John Benjamins.
Torrence, Harold. 2013b. A promotion analysis of Wolof clefts. *Syntax* 16:176–125.
Travis, Lisa. 2010. *Inner aspect: The articulation of VP*, volume 80 of *Studies in Natrual Language and Linguistic Theory*. Berlin: Springer.
van der Wal, Jenneke. 2017. What is the conjoint/disjoint alternation? parameters of crosslinguistic variation. In *The conjoint/disjoint alternation in bantu*, ed. Jenneke van der Wal and Larry M. Hyman, volume 301 of *Trends in Linguistics. Studies and Monographs* [*TiLSM*], chapter 2, 14–60. Berlin: De Gruyter Mouton.
Wagner, Michael. 2006. Association by movement: Evidence from NPI-licensing. *Natural Language Semantics* 14:297–234.
Wasike, Aggrey. 2007. The left periphery, wh-in-situ and a-bar movement in Lubukusu and other Bantu languages. Doctoral Dissertation, Cornell University, Ithaca, NY.
Welch, Nicholas. 2012. The bearable lightness of being: The encoding of coincidence in two-copula languages. Doctoral Dissertation, University of Calgary.
Yuan, Michelle. 2017a. More on Undermerge: Phrasal and head movement interactions in Kikuyu. In *A Pesky Set: Papers for David Pesetsky*, ed. Claire Halpert, Hadas Kotek, and Coppe van Urk. Cambridge, MA: MIT Working Papers in Linguistics.
Yuan, Michelle. 2017b. Movement to complement in Kikuyu and the syntax of focus association. Unpublished, generals paper, MIT.
Zentz, Jason. 2016a. The biclausal status of Shona clefts. Paper presented at the LSA 90, January 9 2016.
Zentz, Jason. 2016b. Forming Wh-questions in Shona: A comparative Bantu perspective. Doctoral Dissertation, Yale University, New Haven, CT.
Zorc, R. David, and Louise Nibagwire. 2007. *Kinyarwanda and Kirundi comparative grammar*. Hyattsville, MD: Dunwoody Press.