Variation in Bantu copula constructions

Hannah Gibson, Rozenn Guérois and Lutz Marten
SOAS University of London

Abstract

This paper provides an overview of variation in Bantu non-verbal predication and copula constructions. These constructions, it shows, exhibit a wide range of fine-grained micro-variation against a backdrop of broad typological similarity across the Bantu family. Variation is seen in the function of copulas, in their morphosyntactic properties, and with respect to the elements with which they can combine. Foremost in the system of non-verbal predication in Bantu are the use of copula forms to link a wide range of predicates, prosodically marked predication and the presence of morphologically distinct copulas, often with different interpretations and restricted distribution. It shows that nominal and adjective predication, existence, location and possession can all be expressed in Bantu without a lexical verbal head. In terms of morphology, the majority of the languages in the study employ an invariable copula – often a variant of *ni*. A defining feature of Bantu copula constructions appears also to be the presence of more than one copula in a language. After surveying copulas found in Bantu, the paper focuses on five languages – Mongo, Rangi, Digo, Swahili and Cuwabo – and shows differences in complementation options for the different copula forms, with tighter restrictions on locative, existential and possessive interpretations across the sample, as compared to identificational and attributive interpretations. The data presented in the paper are relevant for theoretical studies of copulas and morphosyntactic micro-variation in copula constructions, as well as having implications for the study of language change and language contact.

Keywords: Bantu languages, micro-variation, language change, language contact

1 Introduction

Copulas and copula constructions show considerable cross-linguistic diversity. However, comprehensive comparative and typological studies of copulas in the world's languages are comparatively recent (e.g. Hengeveld 1992, Stassen 1997, Pustet 2003, Creissels 2006: 343-359, 2014). The present paper aims to make a contribution to the cross-linguistic, comparative study of copula constructions by focussing on Bantu languages. With some 450 languages spoken across East, Central and Southern Africa, the Bantu language family provides an ideal lens through which to examine morphosyntactic variation. Bantu languages share a core set of morphosyntactic features: morphologically complex verbs and nouns, noun classes, an extensive system of agreement and pragmatically-motivated word order. However, the languages also exhibit a high degree of micro-variation, which although it has often been noted (e.g. Bearth 2003, Marten et al. 2007), remains largely unexplored in a number of areas.

Copula constructions and non-verbal predication are a particularly rich area for cross-linguistic study in Bantu. Bantu copula constructions exhibit exactly these broad typological similarities, whilst also exhibiting a wide range of fine-grained micro-variation. Across the family, different forms of copulas can be found – invariant, uninflected copulas, copulas which take some inflectional morphology such as agreement morphology, as well as prosodic copulas. Furthermore, a few salient copula types are often formally distinguished, including 'pure', possessive and locative copulas. However, different Bantu languages differ with respect to which particular copula forms are available, and their specific morphosyntactic and

interpretational restrictions. The current study presents a comparative overview of Bantu copula constructions by providing examples from a range of Bantu languages which show both common patterns and areas of variation. We will demonstrate variation in formal expression and combinatorial properties of different copula types across Bantu. Rather than focussing on one specific language, we adopt a deliberately comparative perspective, which highlights the degree of micro-variation encountered. While drawing on a range of Bantu languages throughout the paper to illustrate variation in construction types, we present a more systematic comparative case study of five Bantu languages – Mongo, Rangi, Digo, Swahili and Cuwabo – in Section 5 to show how variation plays out between the languages of the sample. Given the number of Bantu languages, and the absence of good morphosyntactic documentation for many of them, the present study is necessarily preliminary. However, it will highlight comparative, typological and theoretical aspects of Bantu copula constructions, as well as identify challenges for future research.

Before beginning our discussion, we will provide some background on the notion of copula we adopt, and note limitations of the study. As a category, copulas are not easily defined. Traditional definitions tend to consider copulas to be semantically empty elements (see, for example Hengeveld (1992:32) and Stassen (1997:65)). Pustet (2003:5) provides the following definition:

> "A copula is a linguistic element which co-occurs with certain lexemes in certain languages when they function as predicate nucleus. A copula does not add any semantic content to the predicate phrase it is contained in."

However, the issue of 'empty' copulas on the one hand and semi-copulas (also known as copula verbs or verbal copulas) and auxiliaries on the other hand is not entirely straightforward and this challenge extends to the situation in Bantu. In many Bantu languages the so-called 'semi-copulas' often convey a lexical meaning such as 'stay, remain' or 'become' and may be able to host subject and temporal information – which is not typically the case with 'true' copula forms. Although they are found across Bantu, such 'semi-copulas' are excluded from the scope of the current study, which focuses on 'true' copulas in order to facilitate cross-linguistic comparison. We follow the definition provided by Pustet (2003) and do not examine 'copula-like' elements which make their own, independent semantic contribution to the phrases in which they occur. We restrict our discussion to true copulas which appear as a predicative base, are broadly morphologically invariant and which make limited or no semantic contribution to the clause. We therefore leave to one side verbal copulas which are transparently based on or related to verbs such as 'be' or 'become' although we are aware that the line between copulas and 'semi-copulas' is often blurry and a conclusive decision in this regard in some languages may present more of a challenge than in others.

The study makes no claims as to being exhaustive but rather examines Bantu copula constructions in a small convenience sample of languages with a geographic spread across the Bantu area. A challenge for this current work exists in terms of the descriptive state of the various languages. Resources vary in the extent to which they analyse the different functions of copulas and interrogate their ability to combine with elements from various lexical classes. In many (if not all) of the cases examined in the current work, the languages are under-described. The previous work on which we draw also varies in regard to the terminology employed and the extent to which they examine comparable construction types.

However, with this in mind, the chapter develops an account of the variation found in Bantu copula constructions and assesses them in the following domains: 1) formal means of

expressing predication, 2) combinatorial properties of different copulas, and 3) restrictions on interpretation and distribution. The chapter shows that, in Bantu, nominal and adjectival predication, as well as existential, locative and possessive meanings can all be expressed without a lexical verbal head. Whilst there is a great deal of variety across the languages, a common feature of Bantu copula constructions is the presence of more than one copula form, often with the use of distinct specialised forms restricted to specific syntactic and/or interpretational contexts. The presence of a dedicated copula form, as well as the unavailability of copulas to host tense and aspect information is another defining feature of Bantu copula constructions.

The chapter is organised as follows: In Section 2 the formal means of expressing predication in Bantu copula constructions is explored. Section 3 describes the syntagmatic and combinatorial properties of the copulas, whilst Section 4 is concerned with restrictions on the distribution and interpretation of copula constructions. Section 5 presents a focused case study of five Bantu languages and compares the copula forms and constructions found in these languages. Section 6 brings out the overall typological variation found in the data investigated and probes the comparative and theoretical implications of the variation. Section 7 constitutes a summary and highlights possible avenues for future research.

2 Formal means of expressing predication

Bantu languages commonly employ a morphological copula as the basis for predication. This morphological copula can be either invariable – showing no concord with the nominal subject for which it is acting as the predicative base – or an inflected form which hosts subject information. However, non-verbal predication may also be prosodically encoded and expressed solely through tonal mechanisms. A number of languages also employ zero-marked predication or 'copula omission' strategies (Pustet 2003) in which no overt copula form is used. Whilst in the current section we adopt a descriptive perspective, the presence of the phenomenon of 'copula omission' raises questions about the syntactic status and issues for a theoretical analysis of the copula elements – an issue which is returned to in Section 6. The current section examines the formal means of expressing predication in terms of the different (morphological) forms exhibited by the languages. Whenever possible, the historical origin of the copula forms under examination is discussed, as well as possible routes of grammaticalisation that may have led to their development.

2.1 Invariant copulas

A common strategy for non-verbal predication across Bantu involves the use of an irreducible copula, i.e. a copula which is invariable for all persons and tenses. Such a one-form copula has been reconstructed for Eastern Bantu as **ní* for the affirmative copula and it also has a negative counterpart which has been reconstructed as **ti* (Meeussen 1967:115). A form that is the same as or closely resembles *ni* is found in many of the languages in the sample. Thus, in Swahili, for example, the non-inflecting copula *ni* can be used as a predicative base. Crucially, the copula does not show agreement in terms of person and number or noun class, as can be seen on examination of examples below where the copula is of the same form in

both (1a) which introduces a singular (class 1) noun and (1b) where the copula introduces a plural (class 2) noun.[1, 2, 3]

(1) Swahili (G42)

a. Juma **ni** mw-alimu.
Juma COP 1-teacher
'Juma is a teacher.'

b. Wa-toto ha-wa **ni** wa-nafunzi.
2-child DEM-2 COP 2-student
'Those children are students.'

A similar situation is seen in Digo (E73, Kenya) and Kagulu (G12, Tanzania) which employ the invariant copulas *ni* and *no* respectively.

(2) Digo (Nicolle 2013:286)

Mu-tu hiyu **ni** daktari.
1-person 1.DEM1 COP 1a.doctor
'The person is a doctor.'

(3) Kagulu (Petzell 2008 :162)

A-niye **no** mu-nhu wa bosi ha-no.
IV-1SG COP 1-person 1.CON first 16-DEM
'I am the first person here.'

Tunen (A44, Cameroon) and Makhuwa (P31-32, Mozambique) also employ invariable copula forms, as can be seen in examples (4) and (5) below.

(4) Tunen (Dugast 1971:348)

Mɛ **lɛ** mù-ɛs.
1SG.PRO COP 1-good
'I am fine.'

[1] A defining feature of the Bantu languages is the use of noun classes. Noun classes function like grammatical genders and are realised morphologically as prefixes on nouns and as agreement markers on a wide range of dependent elements. By convention, the noun classes are numbered and with the numbers of up 10, odd numbers commonly represent singular forms whilst even numbers represent plural forms.

[2] Following Guthrie (1967-71) and Maho (2003), Bantu languages are subdivided into geographic zones by means of an alpha-numerical index, such as G42 for Swahili. The zones do not necessarily reflect language history or genetic classification.

[3] Glossing conventions follow the Leipzig Glossing Rules with the following additions: 1, 2, 3 etc. = noun class number, APPL = applicative, AUG = augment, AUX = auxiliary, CC = complement case, CJ = conjoint, CON = conjunction, COP = copula, DEM = demonstrative, F = feminine, FUT = future, FV = final vowel, HAB = habitual, INC = inceptive, INF = infinitive, LOC = locative, NEG = negative, OM = object marker, PASS = passive, POSS = possessive, PLUR = pluractional, PRD = predicative, PRF = perfective, PST = past, PRO = pronoun, RED = reduplication, SBJ = subjunctive, SM = subject marker.

(5) Makhuwa (van der Wal 2009:45)

O-tek-a	w-aw'	e-nupa	Zainale	**ti**	w-ooreer-a.
15-build-FV	15- POSS.1	9-house	1.Zainal	COP	15-be.good-FV

'Zainal's (way of) building a house is good.'

Copulas do not always appear as independent elements. In the course of their diachronic development, some Bantu copula forms have become segmentally reduced, leading to a loss of phonological and syntactic autonomy. A copula may thus appear as a clitic, attached to the element for which it is acting as the predicative base. This is the case in Yeyi (R41, Botswana), where the clitic appears as the form *ndi-* which is attached to the nominal form before the noun class prefix (6). Similarly, in Saamia (JE34, Uganda) the copula appears as a proclitic on the noun (7).

(6) Yeyi (Seidel 2008:415)

Muraliswani	**ndi**=mu-teriki.
Muraliswani	COP=1-cook

'Muraliswani is the cook.'

(7) Saamia (Botne et al. 2006:38)

N'Ochwada	yéesi	**n**'-ómú-kóyí.
and-Ochwada	1.also	COP-1-brewer

'… and Ochwada also is a beer brewer.'

In all of these languages, the use of the copula is restricted to the present indicative and the copula is not available for tense-aspect inflection (see Section 4.5 below.)

2.2 Inflectional copula constructions

Bantu languages are known for their noun class systems in which nouns are allocated to classes (or grammatical genders) on the basis of semantic and phonological factors. Membership in these classes subsequently triggers agreement across the nominal and verbal domains which appear, amongst others, as a system of subject markers and object markers. In addition to the non-inflecting, invariable copula forms, a number of Bantu languages also employ inflected copula forms which show exactly this subject agreement. For example, in both Mongo (C61, Democratic Republic of the Congo) and Zulu (S42, South Africa) the inflected copula is comprised of a copula form and the standard subject marker appropriate for the class in question ((8) and (9)).

(8) Mongo (Hulstaert 1965:340, glosses are our own)

N-kómbé	**a-le**	m-púlu.
1-kite	SM1-COP	9-bird

'The kite is a bird.'

(9) Zulu (Zeller 2013:1120)

U-Thandi	**u-ng**-u-m-fundi.
AUG-1a.Thandi	SM1-COP-AUG-1-student

'Thandi is a student.'

In contrast, in many Bantu languages, inflecting copulas are not based on the standard invariable copula *ni* but on a locative clitic or a comitative preposition, often with interpretative restrictions (see Section 3, below). This is the case, for example, in Swahili, where the subject marker appears alongside the locative clitic *-ko* (10) as well as in Herero, with the inflecting copula formed using the subject marker and the comitative *-na* 'and, with' (11).

(10) Swahili (Marten 2013:56)

Yeye	**yu-ko**	Ukerewe	mimi	**ni-ko**	Usukuma.
3SG.PRO	SM1-LOC17	Ukerewe	1SG.PRO	SM1SG-LOC17	Usukuma

'He is in Ukerewe, I am in Usukuma.'

(11) Herero (Möhlig and Kavari 2008:214)

Mbí-na	o-rűvyó	o-ru-pé.
SM1-POSS.COP	AUG-11.knife	AUG-11-new

'I have a new knife.'

The systems above show the use of a copula form, which may be derived from a locative clitic or the comitative *na*. These constructions regularly employ the standard subject marker used throughout the language. However, an additional predicative strategy sees the use of a subject marker functioning on its own (i.e. without an overt copula form or other morphological element), as can be seen in examples (12) and (13) below.

(12) Digo (Nicolle 2013:289)

Chi-tabu	**chi**	tayari.
7-book	SM7	ready

'The book is ready.'

(13) Swahili (Ashton 1947:93)

Nyumba	**i**	tupu.
9.house	SM9	empty

'The house is empty.'

In addition to the copula strategies outlined above, there are also languages in which the inflected copula forms are historically derived from demonstrative pronouns. This is the case in Cuwabo, for example, where there is no invariable copula, rather the copula is formally identical to the proximate demonstrative stems (with the exception of the class 1 copula form *ddi*).

Table 1 Copula forms and proximate demonstratives for all classes in Cuwabo

Class	1	2	3	4	5	6	9	10	14	15	16	17	18
COP	ddi	ba	bu	si	tti	ba	ji	si	bu	ku	pa	ku	pu
Demonstrative ('this')	-ddu	-ba	-bu	-si	-tti	-ba	-ji	-si	-bu	-ku	-pa	-ku	-pu

These copulas appear as morphologically independent forms before the element for which they function as the predicative base. However, these forms are restricted in terms of function and distribution. Their function is essentially presentative, and so they are not found in typical predication relations. They are typically used without an overt nominal subject, and

they agree with the following nominal complement, either a full NP (14a) or a (personal/demonstrative/interrogative) pronoun (14b-c).

(14) Cuwabo (Guérois 2014:455)

a. óbú **bu** **m̩-béni** o-a mú-yâna.
3.DEM.I COP3 3-knife 3-CON 1-woman
'This is the knife of the woman'

b. sísi **ku** **óĺle** óĺle ki namarogolo
INTER COP17 1.DEM.III 1.DEM.III EMPH 1a.hare.PRD
'Hey! Here it is, that one there is the hare'

c. kí [o-ni-mú-j-a naámbédde]$_{REL}$ **ba** **aani**.
EMPH SM1-IPFV.CJ-OM1-eat-FV 1a.maize COP2 who
'Who is eating my maize?"

The copula system in Cuwabo is therefore similar to those outlined in the current section in which the copula forms show obligatory agreement. However, the Cuwabo system differs from the inflected copula system found in other languages in several respects: first, it is semantically confined to a presentational function; second it constitutes an independent paradigm, the inflectional system of which is not linked with that found in standard verbal agreement; third, the copula agrees with the following element, i.e. the presented element (except, however, when it is involved in a certain type of adjectival predication, based on a connective construction – see subsection 3.1).[4]

Another source for an inflected copula form attested in Bantu is the combination of a copula and an inflected referential pronominal marker. Whilst this functions as a predicative base, this strategy is typically employed for emphatic purposes as can be seen in the examples from Digo (15) and Swahili (16) below. These forms are also regularly used in the formation of emphatic cleft constructions in Swahili (see e.g. Schadeberg 1992:28).

(15) Digo (Nicolle 2013:289)
Kulungu **ndi**-ye wangu.
1a.antelope COP-1.REF 1.POSS.1SG
'The antelope is (indeed) mine.'

(16) Swahili (Ashton 1947:93)
Hindi **ndi**-ko kw-enye nguo.
9.India COP-17.REF 17-with 10.clothes
'India is (the country) of clothes.'

2.3 Tonal marking

In addition to segmental copulas, non-verbal predication can be expressed through tonal marking. In other words, referential and predicative nominal forms may be distinguished solely by means of the associated tone patterns. Pustet (2003:44) terms such cases

[4] A similar system is also observed in Koti (P311, Mozambique) (Schadeberg and Mucanheia 2000) – a language closely related to Makhuwa and Cuwabo.

'suprasegmental copulatives'. Two configurations of tone-marked copulas are possible in Bantu. One option is that predication is expressed by a 'tonal replacive' (Welmers 1973:322f), where the tone which occurs with predication is associated with the noun class prefix or with the pre-prefix (discussed below). This is the case in Shona (S10, Zimbabwe) where nominal prefixes inherently carry a low tone. However, this low tone is replaced by a high tone to express identification, as can be seen in the example in (17) (the relevant vowels appear in bold to aid presentation).

(17) Shona (Welmers 1973:323)

a.	m**ù**-nhù	b.	m**ú**-nhù
	1-person		1-person.PRD
	'person'		'It is a person.'

A subset of Bantu languages also have an augment, also known as 'pre-prefix' or 'initial vowel' in addition to the noun class prefix. The function of the augment varies from language to language: in some languages it is merely stylistic and therefore optional whilst in other languages it is a marker of referentiality (see de Blois (1970) and Maho (1999) for more on the Bantu augment). As already noted in Section 1, in Herero (as in Shona above), predication is achieved through a distinct tonal pattern. However, in the case of Herero, this tone attaches to the augment rather than to the noun class prefix. This is the case both in nominal predicates where the high tone appears on the augment of the nominal (18b), as well as in adjectival predicates where the high tone is associated with the concordial prefix (18d.)

(18) Herero (Möhlig and Kavari 2008:122, Kavari et al. 2012)

a.	ò-tjì-hávérò	b.	**ó**-tjì-hávérò
	AUG-7-chair		AUG-7-chair
	'chair'		'It is a chair'

c.	ò-zò-ngòmbè	ò-zò-néné	d.	ò-zò-ngòmbè	**ó**-zò-néné
	AUG-10-cow	AUG-10-big		AUG-10-cow	AUG-10-big
	'big cows'			'The cows are big'	

The second strategy for tonal marking of predication is that rather than a tone being added, the tone is either lowered or deleted. This is the case in Zulu, for example, where the unmarked form of the noun involves a high tone on the augment. However, a low tone on the augment can be used predicatively on nouns and adjectives. This can be seen in the contrast between example (19a) and (19b) below.

(19) Zulu (Doke 1961:215)

a.	**í**-m-buzi	b.	**ì**-m-buzi
	AUG-9-goat		AUG-9-goat
	'goat'		'It is a goat'

A similar process also occurs in Cuwabo and Makhuwa where it has been termed 'Predicative Lowering' (van der Wal 2006, 2009, Guérois 2015a, 2015b) (glossed below as PRD). Predicative Lowering consists of the deletion of the first underlying high tone (regardless of its position), on both nominal (20) and adjectival predicates (21).

(20) Cuwabo (Guérois 2015a:126)

a. mu-sáno / mú-yaná
1-queen 1-woman
'queen' 'woman'

b. mu-s**a**no / m**u**-yaná
1-queen.PRD 1-woman.PRD
'It is a queen' 'It is a woman'

(21) Makhuwa (van der Wal 2009)

a. n-thálí mw-áńkhaáni
3-tree 3-small
'the small tree'

b. n-thálí mw-**an**khaáni
3-tree 3-small.PRD
'The tree is small'

The difference between a simple nominal (or adjectival) form and the predicative form in these two languages therefore, can be encoded purely through the use of a distinct tonal pattern.[5]

2.4 Copula omission

The omission of the copula, also known as 'copula dropping', is a common phenomenon cross-linguistically (Pustet 2003). This omission can be optional (e.g. in Cantonese, Mould (1974)) or compulsory in specific grammatical environments (e.g. in Russian and Hungarian, de Blois (1970)). However, it has been observed that copula omission usually occurs with a restricted set of grammatical categories and Pustet (2003) notes that in languages in which both nouns and adjectives can be combined with copulas, adjectives are more likely to be associated with copula omission than nominal predicates.

A number of Bantu languages also allow for the omission of the copula. This is the case in Swahili, for example, where it is possible for predication to be achieved without the presence of the copula *ni*, if the subject is an independent pronoun or demonstrative (Welmers 1973: 325) This can be seen in example (22), where the copula *ni* can optionally be omitted from the clause.

(22) Swahili (Ashton 1947:92/93)
Mimi (ni) Hamisi
1SG (COP) Hamisi
'I am Hassan.'

The same phenomenon can be seen in example (23a), and no overt copula form is present (contrast this with example (23b)). Whilst this construction employs a subject form followed by a demonstrative, this is also possible with a possessive pronoun (Ashton 1947:92). In sentences such as these however, the overt subject nominal (or pronominal) form is obligatorily present, whilst in other contexts an overt subject expression is not necessary since subject pro-drop is widespread throughout the language (and in Bantu more widely).

[5] Note that a (diachronic) relation between Predicative Lowering and the augment has also been proposed by van der Wal (2006) for Makhuwa and Guérois (2015b) for Cuwabo. However, this issue falls outside the scope of the present study and the interested reader is instead referred to the aforementioned references for a more detailed discussion of the matter.

(23) Swahili (Ashton 1947:93)

a. M-ti huu m-bovu.
3-tree 3.DEM 3-rotten
'This tree is rotten.'

b. M-ti huu ni m-bovu
3-tree 3.DEM COP 3-rotten
'This tree is rotten.'

It has also been noted that examples in which the invariable copula is omitted (such as (22) and (23) above) are commonly associated with the use of specific prosodic cues. Thus, in example (24), the predicative adjective *kuzuri* 'good' is described as receiving a 'final tonic' and the whole expression is associated with sentence intonation (Maw 1969).

(24) Swahili (Maw 1969:42)

Ku-tend-a | ku-zuri.
15-act-FV 15-good
'Acting is good.'

A similar situation is also seen in Digo – a language closely related to Swahili – where both nominal (25a) and adjectival (25b) predication can be achieved without the presence of an overt copula form (although a dedicated copula form is available in the language), presumably also with an associated prosodic pattern. This same type of copula omission also occurs in presentative clauses, as can be seen in example (25c).

(25) Digo (Nicolle 2013:286, 297)

a. Mimi sowe tu.
1SG.PRO 1a.your.father only
'I am your father.'

b. Ma-ruwa-ge mereru.
6-flower-6.3SG.POS 6.white
'Its flowers are white.'

c. Bwana Sulutani m-ganga hiyu.
lord Sultan 1-healer 1.DEM.I
'Lord Sultan, this is the healer.'

This section has shown the formal means of expressing predication as they are manifested across a subset of Bantu languages. In addition to the descriptive variation found in copula constructions, the forms also give rise to a number of cross-linguistic theoretical challenges, which we will address in more detail in Section 6. For example, how best to define a copula construction and how to distinguish between different co-occurring forms (if indeed such a distinction is necessary), and whether a 'zero' copula construction simply consists of two nominal phrases, or whether a pause or a specific intonation is necessarily associated with the construction (cf. the Swahili example in (24)). These challenges notwithstanding, the data in the present section have exemplified some broad similarities across Bantu copula constructions, as well as a range of variation. The next section discusses different restrictions on the use of the different copula types.

3 Restrictions on the interpretation and distribution of copulas

The interaction of different formal means of encoding non-verbal predication with specific syntactic configurations in which they are used often results in differences in interpretation and appropriateness in different contexts. Copulas commonly perform one (or more) of a number of functions including introducing attributive, identificational, existential, quantificational and temporal meanings (Pustet 2003:30-33; Mikkelsen 2005). Across the language family, different copula forms are often used for these functions, as well as for possessive predication, locative predication and negative predication. This also means that it is a common feature of Bantu languages to have more than just a single copula. Furthermore, different copula forms are also often restricted to the use with specific tense-aspect forms. The current section discusses the interpretation and distribution associated with copula forms in the Bantu languages under examination.

3.1 Predicative copulas

A common function of copula constructions cross-linguistically and across-Bantu is to convey a predicational relation. We include in this relation different more specific readings, such as specificational, equative and identificational readings (see e.g. Citko 2011 for discussion), since we do not have at present sufficient empirical evidence to distinguish between them. In addition, these readings are expressed across Bantu by the same morphological form of copula (sometimes called the 'pure' copula), as opposed to possessive and locative copulas, discussed in the next section.

This predicative use of the copula can be seen in example (26) from Digo where the copula *ni* is used to link the first person singular subject *mino* 'I' and the nominal expression *mutu* 'person'.

(26) Digo (Nicolle 2013:298)

Mino	**ni**	mimi	mu-tu.
1SG.PRO	COP	just	1-person

'I am just a person.'

A similar situation can be seen in Rangi (F33, Tanzania) where the copula *nɪ* is used attributively and can connect an adjective, for example, with a nominal phrase.

(27) Rangi (Gibson 2012:92)
Í-kí kɪ-ɪntʊ **ní** ch-óócho
DEM-7 7-thing COP 7-true
'This thing is true.'

In Swahili, in addition to the use of the pure invariable copula (28a), the locative copula can be used to express predication (28b). However, the use of this construction appears to be highly restricted, and can occur with just a few forms, such as *huru* 'free' and *tayari* 'ready'.

(28) Swahili
a. Kaka y-angu **ni** m-refu
1.brother 1a-my COP 1-tall
'My brother is tall.'

b. Wa-tu wa-li-ji-on-a **wa-po** huru zaidi.
2-person SM2-PAST-REFL-see-FV SM2-LOCCOP16 free more
'People saw themselves as being more free.' (Marten 2013:62)

In Cuwabo, a verbless predicative relation can be formed by means of Predicative Lowering (see section 2.3). This is illustrated in (29a) and (29b) where the tonally 'lowered' predicates *mu-nddimúwa* 'be tall' and *namapuja* 'be a joker' are derived from the adjective *mú-nddímúwa* 'big' and the noun *namápûja* 'joker'.[6]

(29) Cuwabo (Guérois 2015a:452/449)
a. Íyéène **mu-nddimúwa** nínga míyo.
3SG.PRO 1-big.PRD like 1SG.PRO
'He is as tall as me.'

b. Namárógoló **namapuja**.
1a.hare 1a.joker.PRD
'The hare is a joker.'

3.2 Possessive copulas

Copulas can also be used to express possession. A number of languages employ a subject marker and the comitative preposition *na* 'and, with' in order to form possessive constructions. This can be seen in Herero (30) and Digo (31), which use the comitative *na* with a class 1 subject marker.

[6] Note that adjectival predicates may also be obtained in Cuwabo through the combination of the agreeing copula (derived from the demonstrative stem – see subsection 2.2) followed by the connective relator *-a*, in turn followed by an infinitive verb expressing a quality. This is illustrated below, with the infinitive verb *olápa* 'be big'. However, his specific use of the agreeing copula in Cuwabo is in need of more thorough description.
Cuwabo (Guérois 2015a:459)
Mú-sólro wa mú-áná óddu bu-a o-láp-a.
3-head 3.CON 1-child 1.DEM.I COP3-CON 15-be.big-FV
'The head of this child is big.' (lit. the head of this child it is of to be big')

(30) Herero (Möhlig and Kavari 2008:214)
Ú-ná omú-tímá omu-wá.
SM1-POSS.COP 3-heart 3-good
'S/he has a good heart.'

(31) Digo (Nicolle 2013:287)
Yuya mw-ana w-ao m-mwenga **ka-na** chi-tu
1.DEM2 1-child 1-3PL.POS 1-one SM1.NEG-COM 7-thing
'One of their sons did not have a thing'

Similarly, Rangi employs the copula form *-rɪ* which hosts subject agreement, together with the preposition *na* 'and, with' to encode a possessive reading (32).

(32) Rangi (Gibson 2012:101)
Ndí-**rɪ** na njala.
SM1SG-COP PREP 9.hunger
'I am hungry.' Lit.: I am with hunger.'

3.3 Locative copulas

A number of Bantu languages employ a dedicated copula form in locative constructions. This is the case in Rangi where a locative reading is achieved through the use of an inflected form of the locative copula *mwáá-rɪ*. The form consists of the prefix *mwáá-* which is thought to be a relic of the class 18 locative marker (which is no longer productive in the language) and the pure copula *-rɪ* (33a). Crucially, such locative constructions cannot be formed using the pure invariant copula *ní* (33b).

(33) Rangi (Gibson 2012:201)
a. ʊhʊ mʊ-waanga a-mwáá-rɪ?
1-DEM 1-doctor SM1-LOC-AUX
'Is the doctor here?'

b. *ʊhʊ mʊ-waanga nɪ kʊra
1-DEM 1-doctor COP there
Intd. 'The doctor is there.'

As was also shown in Section 2.3, one of the copula forms found in Swahili is comprised of a subject marker and a locative clitic. Typically this copula form encodes a locative interpretation, where it serves as the predicative base for the combination of a noun phrases and a location (34). The pure copula *ni* cannot be used to express locative predications on Swahili, as can be seen on examination of the examples (34a) and (34b) which can instead only be rendered using one of the locative copula forms (34c).

(34) Swahili (Marten 2013:56, 62)
a. Yeye **yu-ko** Ukerewe mimi **ni-ko** Usukuma.
3SG.PRO 1SP-LOC17 Ukerewe 1SG.PRO 1SG.SP-LOC17 Usukuma
'He is in Ukerewe, I am in Usukuma.'

b. **Wa-ko** wapi wa-toto w-angu?
SM2-LOC17 where 2-child 2-my
'Where are they, my children?'

(35) Swahili (Marten 2013:56, 62)

a. *Yeye **ni** Ukerewe
3SG.PRO COP Ukerewe
Intd.: 'He is in Ukerewe.'

b. *Yeye **ni** katika nyumba kubwa
3SG.PRO COP in 9.house 9.big
Intd. 'S/he is in a big house'

c. Yeye **yu-ko** katika nyumba kubwa
3SG.PRO SM1-LOC17 in 9.house 9.big
'S/he is in a big house'

In contrast, in Mongo, both the pure copulas *-le* and *ni* can be used to encode a locative meaning (36), and in Digo the pure copula form *ni* can also be used in locative constructions (37).

(36) Mongo (Hulstaert 1965:340, glosses are our own)
Nyangó **a-le** ndá li-sála.
1.mother 1-COP PREP 5-field
'The mother is at the field.'

(37) Digo (Nicolle 2013:302)
Nkaamb-w-a m-lungu-ni **ni** phapha.
SM1SG.PAST.tell-PASS-FV 3-heaven-LOC COP here
'I was told heaven was here.'

The close relationship between locative and existential readings can also be observed in copula constructions in which similar strategies for encoding both readings are present. In Swahili, for example, both a locative clitic constructions and a locative possessive constructions (cf. Section 4.2) can be used to express existence in a place or more abstract existence (38).

(38) Swahili (Marten 2013:57, 47)

a. **U-po** u-husiano kati y-a elimu na
SM11-LOC16 11-relation between 9-GEN 9.education CONJ
ki-pato.
7-earnings
'There is a relation between education and earnings.'

b. **Ku-na** ma-endelo sana.
SM17.LOC-COM 6-development much
'There is a lot of development.'

However, there are considerable differences between these two Swahili locative constructions in terms of complementation and interpretation. The clitic construction, as in (39a), is not only used in locative-existential contexts, but can also be used in true locative contexts (cf. (34), above) as well as in some predicational contexts. Furthermore, the order between the copula and subject of predication (as well as the optional locative phrase) is free, although sometimes associated with interpretational differences:

(39) Swahili (Marten 2013)

Locative clitic construction

a…maskini **wa-po,…** lakini na u-tajiri pia **u-po**
10.pauper SM2-LOC16 but CONN 11-wealth also SM11-LOC16
'…there are poor people, but wealth too, is there.'

b. **Zi-ko** sababu m-bili zi-li-zo-fany-a ki-tabu
SM10-LOC17 10.reason 10-two SM10-PAST-REL10-make-FV 8-book
hi-ki ki-andik-w-e.
DEM-8 SM8-write-PASS-SBJV
'There are two reasons which made this book be written.' [Jen Fal 1]

In (39a), the subject of the predication *maskini* 'poor people' precedes the copula (which shows class 2 animate plural agreement with it), while in (39b) the copula precedes the subject.

In contrast, possessive locative copulas are only used in locative-existential contetxs, and only allow one order, that in which the subject follows the copula:

(40) Swahili (Marten 2013)

Possessive locative copula construction

a. **Ku-na** m-oshi!
SM17.LOC-COM 3-smoke
'There is smoke!'

b. *M-oshi **ku-na**
3-smoke SM17.LOC-COM
Intd.: 'There is smoke'

3.4 Negative copulas

Bantu languages also often have specific invariant negative copulas. In addition to not showing subject agreement in the language in question, these are also often transparently synchronically related and often take the form of some variant of *si*, as can be seen in the example from Digo (41). In Rangi, in addition to the negative copula *si*, sentential negation obligatorily employs the negative marker *tʊkʊ* which appears post-verbally or clause-finally (42).

(41) Digo (Nicolle 2013:287)

Mino	**si**	dza	a-tu	a-njina.
1SG.PRO	NEG.COP	like	2-people	2-other

'I am not like other people.'

(42) Rangi (Gibson 2012:95)

ɪkɪ	ki-kombe	**sí**	ch-aanɪ	tʊkʊ,	nɪ	ch-ááchwe.
DEM-7	7-cup	NEG	7-my	NEG	COP	7-his/her

'This cup is not mine, it is his/hers.'

Mongo is one of the few Bantu languages which employs a dedicated negative copula which shows subject concord. This copula form hosts the standard subject markers seen throughout the language, as can be seen with the form *-fa* (43).

(43) Mongo (Hulstaert 1965:340)

Tó-fa	ba-laki.
1PL-NEG.COP	2-teacher

'We are not teachers'

3.5 Copulas and tense-aspect information

In Swahili, the copula *ni* can only be used in the present indicative. In other tenses it is replaced by the verbal form *kuwa* 'to be' which can carry temporal and aspectual information. Indeed, across Bantu languages, copulas are typically not available for the standard tense-aspect inflection found in the verbal domain. However, in some languages, copulas are compatible with distinct tense-aspect interpretations. For instance, in Digo, the time reference of a clause containing the invariable copula *ni* may be determined solely by the context. This means that whilst *ni* does not host temporal inflection it can be used with past time reference, as in example (44) where the past time reference is recoverable from the preceding context.

(44) Digo (Nicolle 2013:287)

[Hipho kare kpwahenda mutu na mchewe achivyala ana airi alume. Phahi yuya mutu na mchewe, asagala, achisagala lakini achikala] yuya mwana wao mmwenga kana chitu,] ni mchiya na mmwenga ni tajiri sana.

'Long ago there was a man and his wife and they had two sons. So that man and his wife, they stayed, and they stayed but he was one of their sons did not have a thing,] he was poor, and the other [lit: and one] was very rich.'

ni	m-chiya	na	m-mwenga	**ni**	tajiri	sana
COP	1-child	COM	1-one	COP	wealthy	very

'…he was poor, and the other was very rich.'

Digo also has two dedicated aspectual copulas: the persistive marker *chere* 'still' and the negative inceptive marker *dzangbwe* 'yet' (used in a negative polarity contexts). Both of these copulas host subject markers and convey these specific aspectual meanings (45).

(45) Digo (Nicolle 2013:293)

a. A-**chere** moyo.
3SG-PERS 3.heart
'S/he is still alive'

b. Wakati w-angu ta-u-**dzangbwe**.
11.time 11-1SG.POS NEG-11-INC
'My time has not yet come.' Lit. 'My time is not yet.'

Other languages employ a different strategy, with the general copula being available for aspectual modification. This is the case, for example, in Zulu where the copula *ngu* can receive aspectual marking and can host the persistive prefix *se-* (46).

(46) Zulu (Buell and de Dreu 2013:427)
Ngi-se-**ngu**-mfundisi.
1SG.SP-PRST-COP-1.teacher
'I am still a teacher.'

Mongo is the only language in the current study in which the copula itself can encode specific temporal information. Mongo employs a dedicated copula form in past tense reference. This past tense copula also conveys a further distinction between degrees of past. The hodiernal past tense (used to refer to the past tense but in the present day) is encoded by the copula *-ki* which carries a low tone (47a), whilst the hesternal past tense (used to refer to time before today, i.e. yesterday and beyond) is encoded by *-kí* which carries a high tone (47b).

(47) Mongo (Hulstaert 1965:340, glosses are our own)

a. **ń-ki** ɛ-kɔ́ la nkɛ́sá.
1SG-PST.COP LOC-DEM DEM 9.morning
'I was there this morning.'

b. **ń-kí** bo-sáj' ǒ-káé.
1SG-PST.COP 1-worker 1-POSS.1
'I was his worker.'

To summarise, this section has provided an overview of the different interpretations with which copulas can be associated in Bantu languages. A number of the languages under examination in the current study employ dedicated copula forms to convey locative meanings, which are often closely linked to (or identical to) existential forms. A number of languages also use copula constructions to express a possessive meaning, often employing the comitative *na* 'and, with' to encode this meaning. Finally, all of the languages under examination have a dedicated negative copula. Whilst in the majority of languages this negative copula is invariable and is the counterpart to the affirmative invariable copula, Mongo has an inflecting negative copula form.

4 Combinatorial properties

There is a great deal of cross-linguistic variation with regard to the different parts-of-speech with which copulas can combine (Pustet 2003:7). While there is an on-going discussion regarding the challenge of defining and distinguishing between parts-of-speech (see for example, Anward et al. (1997), Croft (1999), Vogel and Comrie (2000)), we adopt here a largely notional approach to parts of speech such that prototypical nouns designate things or entities, prototypical adjectives refer to attributive properties, whilst prototypical verbs convey actions or events (cf. e.g. Lyons 1977), coupled with morphosyntactic characteristics specific to Bantu languages.

In Bantu, nouns govern a series of agreement markers on dependent modifiers while verbs have agglutinative morphology marking categories such as person/class, tense, aspect, modality, and polarity through a series of prefixes and suffixes. Adjectives in Bantu languages typically constitute a small closed lexical class and agree with the noun they modify, most commonly taking the same noun class morphemes as nouns. The challenge comes therefore in Bantu languages with distinguishing between nouns and adjectives, many of which employ the same morphological features. The Swahili form *-fupi* 'short', for example, could be considered adjectival in nature. However, the addition of the noun class 1 prefix *m-* results in the form *mfupi* for which it is not possible to distinguish morphologically between the class 1 noun 'short person' and an adjective 'small' modifying a class 1 noun. However, adjective stems can be distinguished from nominal stems syntactically as modifying a head noun, and distributionally since they are, in contrast to nouns, not lexically restricted to a specific noun class.

Putting to one side the challenges involved in distinguishing between the formal properties of different parts of speech in Bantu, this section examines the elements with which copulas can combine.

Across the Bantu language family, copulas can combine with a wide range of elements, including nouns, adjectives, infinitives and wh-question words. In Swahili, the invariant copula *ni* can combine with nominal forms (48a), adjectival forms (48b), infinitives (48c) and wh-questions (48d). However, the copula *ni* cannot be associated with non-finite verb forms which host inflectional information pertaining to a variety of categories, including noun class, person/number distinctions and tense (48e).

(48) Swahili

a. Juma **ni** mw-alimu
Juma COP 1-teacher
'Juma is a teacher.'

b. Juma **ni** m-fupi
Juma COP 1-short
'Juma is short.'

c. Kazi y-angu **ni** ku-pand-a m-begu
9.work 9-my COP INF-plant-FV 10-seed
'My work is to plant seeds.'

d. Huyu **ni** nani?
1.DEM COP who
'Who is this?'

e. *Mimi **ni** ni-li-kuw-a m-toto
I COP SM1SG-PST-be-FV 1-child
Intd. 'I was a child.'

In Saamia the copula *ni-* is used to link a subject noun phrase (either a full overt nominal phrase or a personal pronoun) and a predicative element, which can be a noun (49a), an adjective (49b), an adjective followed by an infinitive (49c), or a relative construction (49d).

(49) Saamia (Botne et al. 2006:38)

Predicative noun

a. Lilá **n**'-éxaande lyaáxá
5.that COP-5.knife 5.new
'That is the new knife.'

Predicative adjective

b. Óxudeexá xwaang'iná **n'**-óxulaí.
15.cook 15.link.mother COP-15.good
'His/her mother's cooking is good.'

Predicative adjective + infinitive

c. Emyóógo kinó **n'**-émyáángú óxweedeexá.
4.cassava 4.this COP-4.easy 15.REFL.cook
'This cassava is easy to cook.'

Relative construction

d. Sinó **ní**-sy-o síí-nj-ixalá=xo.
7.this COP-7-REL 7-1SG-sit=LOC17
'This is the one that I sit on.'

The three copula forms of Mongo (the present copula *-le*, and the copulas *-ki* and *-kí* which encode yesterday's past and today's past respectively) can combine with nouns (50a, b) and adjectives (50c, d), but also with a prepositional phrase (in this case conveying a locative meaning) (50e).

(50) Mongo (Hulstaert 1965:340)

Nominal

a. Ba-**le** ba-laki.
2-COP 2-teacher
'They are teachers.'

b. N-**kí** bo-sáj' ŏ-káé.
1SG-PST.COP 1-worker 1-POSS3SG
'I was his worker.'

Adjective

c. A-**le** bɔ-nɛ́nɛ.
1-COP 1-big
'He is big.'

d. Bo-támbá bó-**ki** bó-ló ndá nkŏkɔta.
3-tree 3-PST.COP 4-difficult PREP 15.cut
'The tree was difficult to cut.'

Locative

e. Nyangó a-**le** ndá li-sála.
1.mother 1-COP PREP 5-field
'The mother is in the field.'

Recall that in Makhuwa, Cuwabo and Ekoti (three closely related languages spoken in Mozambique), one of the strategies for achieving non-verbal predication is through the use of Predicative Lowering. This process sees the deletion of the first high tone, resulting in a low tone. In Makhuwa for instance, this Predicative Lowering strategy can be used with nouns (51b), adjectives, infinitives and most interrogatives (51c).

(51) Makhuwa (van der Wal 2009: 122)

Predicative noun

a. nakhúku
'crow'

b. mwaánúni ulá nakhukú
'This bird is a crow.'

Interrogatives

c. Esheeni íyo?
9.what.PL 9.DEM.II
'What is that?' lit.: 'It is what, that?'

However, predicative lowering cannot be used with a number of other types of complement such as personal and demonstrative pronouns (52a), cleft-questions with the wh-word *paní* 'who' (52b), questions with 'which one,' connective constructions (52c), or the relative participial modifier (52d). In these instances, the overt invariable copula ***ti*** must be used instead.

(52) Makhuwa (van der Wal 2009:122)

Personal and demonstrative pronouns

a. Mí-wwa iye **t'** iiye tsi-ki-hom-ak-ants-e.
4-thorn 4.DEM.III COP 4.DEM.III SM4-OM1SG-sting-DUR-PLUR-PERF.REL
'Those thorns are the ones that stung me.'

Cleft-questions with 'who'

b. **Ti** paní o-ni-m-vút-ááwe menínu.
COP 1.who SM1-OM1-pull.REL-POSS.1 1.boy
'Who is it that the boy pulls?'

Connective constructions

c. E-paarti e-kush-iy-e **ti** ya a-neene.
9-bucket SM9-carry-PASS-PERF.REL COP 9.CONN 2-boss
'The bucket which is carried belongs to the boss.'

Relative participial modifier

d. Ni mí també **t'** í-n-úu-him-eery-áaka.
and 1SG.PRO also COP SM9-PRS-OM2SG-say-APPL.REL-POSS.1SG
'…and this is also what I say to you.'

To summarise, copulas in Bantu languages can combine with a wide range of elements. The examples from Swahili above show that the invariant copula can combine with nominal and adjectival forms, as well as infinitives and wh-questions, although it cannot combine with non-finite verb forms. A similar situation is seen in Saamia where the copula can combine with nouns, adjectives and relative constructions. Mongo and Digo are the only languages of our sample in which the copula combines with locative complements, in addition to nouns and adjectives. In Makhuwa, Predicative Lowering can apply to nouns, adjectives, infinitives and most interrogative forms but cannot be used with personal or demonstrative pronouns, cleft questions formed using the wh-word *paní* 'who', 'which one,' connective constructions or the relative participial modifier, which instead can only combine with the variant copula *ti*. Thus, Makhuwa represents a good example of the restriction in terms of combinational properties with different copula forms.

The next section draws out the comparative observations in more depth, through an examination of the interpretation and combinatorial properties of copulas in five of the languages in our sample. It also discusses theoretical questions that are raised by the observations relating to Bantu copula constructions.

5 Comparative and theoretical implications

Copulas and copula constructions show considerable cross-linguistic diversity, and the variation in Bantu copula constructions is no exception to this observation. In this section, we provide a more detailed discussion of the comparative and typological aspects of our study, and draw out implications for the theoretical analysis of copulas in Bantu. We do so with recourse to a small subset of five Bantu languages for which we have sufficient information for a comparative overview. The languages included in this sample are Mongo, Rangi, Digo, Swahili, and Cuwabo. Mongo is a Western Bantu language spoken in the DRC, and Cuwabo is a South-eastern Bantu language spoken in Mozambique. The remaining languages are all Eastern Bantu languages, spoken in Kenya and/or Tanzania. While Rangi is spoken in central Tanzania, both Digo and Swahili are closely related coastal languages with a long history of language contact between the two. A summary of the copula systems of the five languages is provided in Table 1.

[**See Appendix**]

The table shows a number of similarities and differences between the five copula systems. In terms of formal inventory, Mongo, Rangi, and Cuwabo distinguish three copula forms, while Digo has six forms, and Swahili has seven. The majority of copula forms in the sample have segmental content, but Cuwabo also includes a prosodic (tonal) copula, and both Digo and Swahili allow empty copulas (which, at least in Swahili, is associated with specific prosodic cues, cf. Section 2.3). Four languages have a specific negative copula – a distinction which is absent only in Cuwabo. It is noteworthy that the copula systems of Digo and Swahili are very similar – in fact they are identical except for the presence of a locative copula in Swahili which is not found in Digo: As noted above, the two languages are closely related and have been in sustained contact. Digo and Swahili are also the languages with the largest copula inventory, and we will discuss below whether the systems can be reduced by adopting a more abstract theoretical analysis. However, before looking at theoretical implications of the typology, we provide a note on the comparative aspects of the data with respect to (formal) restrictions on complementation and the (semantic) restrictions on interpretation.

5.1 Restrictions on complementation

Following our discussion in the earlier sections of this chapter (see in particular Section 4), we distinguish between nominal, adjectival, locative and infinitival complements. We showed that there are differences between the five languages of the sample in this regard, and that for some languages, there are different restrictions in terms of the complementation patterns of the different copulas.

In Mongo and Rangi, there is no differentiation at all with respect to complement type – all copula forms can be used with all four complement types. However, in Digo, Swahili and Cuwabo, there are restrictions on which copula form can be used with which complement(s). In part, this might be related to the larger formal copula inventory of Digo and Swahili (distinguishing six and seven forms, respectively), but this potential correlation is not confirmed by Cuwabo, where only three forms are distinguished, the same number as in Mongo and Rangi.

Among the three languages which do have specialised copulas – Digo, Swahili and Cuwabo – nominal complements appear to be the least restricted. Possibly with the exception of the case of Swahili subject marker predication (i.e. in which predication is achieved through use of the subject marker only, see Section 2.2), all forms can take nominal complements. Adjectival complements are possible with most copula forms, however, there is one form in each of the three languages which does not allow adjectival complements: the copula related to the demonstrative in Cuwabo, and the possessive copulas formed using the comitative *-na* in Digo and Swahili. On the other hand, the availability of adjectival complementation with the Rangi possessive copula *-rɪ* shows that the combination of possessive copula and adjectival complementation is possible. In contrast to nominal and adjectival complements, locative complements are more restricted. In Digo, locative complements pattern with adjectival complementation (that is, they are not possible with the possessive copula *-na*). However, in Swahili, which has a dedicated locative copula *-ko*, locative complements are only possible with this locative copula, or with an empty copula construction. In Cuwabo, locative complementation is found only with the copula *-li*, and is not possible with the other two copulas. Finally, infinitival complements are highly restricted. In fact, they are not possible in any copula constructions in Digo, and are restricted to two of the three copulas in Cuwabo, and two of the seven in Swahili.

5.2 Restrictions on interpretation

In terms of interpretation, we identify identificational, attributive, existential, locative, and possessive interpretations (cf. Section 4), and we also note a number of temporal restrictions on the use of copulas in our sample (see Table 1).

Similar to complementation, Mongo does not place any restrictions on the different interpretations distinct copulas can be used for (at least for the cases where we have evidence). Mongo is the only language in our sample for which the copula system exhibits this amount of freedom, as in all other languages various different copula forms show restrictions with respect to their interpretational properties. This difference may reflect a wider typological difference between the copula system of northwestern and eastern Bantu languages, but our sample is too small to confirm this idea.

Among the remaining languages, identificational and attributive are the least restricted interpretations, while existential, locative and possessive interpretations are restricted to a smaller set of copulas. In fact, attributive interpretations are possible with all three copula

forms in Rangi, and all but one form in the remaining languages, and identificational interpretations are only marginally more restricted.

In contrast, existential, locative and possessive interpretations are typically only possible with one or two forms of the set of copula forms. In Cuwabo, for example, existential and locative interpretations are only possible with *-li* but not with the two other copula forms. The parallelism between existential and locative interpretations seen in Cuwabo and often noted in the literature (e.g. Lyons (1967), Freeze (1992)) is also found in the other languages of the sample, where Rangi, Swahili, and Digo (with slightly less clear data) show the same restrictions for the two interpretations. However, there is a difference between Digo and the remaining languages. Whereas in Digo, several copulas can express locative (and possibly also existential) interpretations, in Rangi, Swahili and Cuwabo, these interpretations are restricted to one or two forms. In fact, only in Digo can the invariable copula be used to express a locative meaning. It is interesting to note, in this respect, that while the copula systems of Digo and Swahili are largely identical, they differ precisely in the absence in Digo of a copula corresponding to the Swahili locative copula *-ko*, which may explain that locative interpretations are restricted to two forms in Swahili (*-ko* and *-na*), but quite unrestricted in Digo.

The expression of possessive meaning is the most restricted interpretation in the copula systems investigated. The (partial) data in our sample indicate that if this meaning can be expressed by a copula construction at all, it is possible with one copula from only. In Cuwabo, possessive meaning cannot be expressed by any of the three copulas. In Rangi, Digo, and Swahili, possessive meaning can be expressed, but it is only one copula form which allows this interpretation: *-rɪ (na)* in Rangi, *-na* in Digo, and *-na* in Swahili. However, all these copulas can also be used for the expression of other interpretations: While they are the only forms which can express possession, they are not dedicated forms in the sense of expressing this meaning alone.

Finally, across the board, copulas cannot normally be inflected for temporal information and are typically restricted to present use only. However, Mongo has three copulas which distinguish tense reference – the present copula *-le* the past copulas *-ki* (today's past) and *-kí* (remote past).

5.3 Theoretical implications

The comparative study of Bantu copula constructions also raises a number of questions for theoretical analysis of which we would like to briefly discuss two.

First, we have noted the close similarity between the copula systems of Digo and Swahili, which in their inventory differ only with respect to the presence of a locative copula in Swahili (although as we have shown, this may have wider consequences for the expression of locative interpretations in the two languages). The similarity between Digo and Swahili may reflect the genetic proximity of these languages – of those included in the study these are the two which are most closely related. However, it may also reflect convergence through language contact since these languages are spoken in the same area and are in sustained contact – Swahili being used as a lingua franca throughout East Africa. For the study of language contact and language change, and in particular of the relation between language contact and grammaticalisation (see Heine and Kuteva (2005)), examples such as the case of the Digo and Swahili copulas provide a promising challenge for disentangling the different possible underlying causes for similarities and differences. That is to say, to what extent are the apparent similarities in the Digo and Swahili copula systems the result of convergence processes resulting from contact-influenced language change, and to what extent do they

reflect language internal processes of change? Recall that the only difference between these two systems is the presence of the locative copula in Swahili which is not present in Digo. Many Digo speakers are also fluent in Swahili, and so copula (and possibly other morphosyntactic) structures in Digo may well have been introduced by bilingual Digo/Swahili speakers. Under this scenario the present Digo copula system would have replaced an earlier stage of the system with less influence from Swahili and hence once which would look more distinct from that found in Swahili. However, a more detailed study is needed to investigate this question further.

The second question can also be explored through an examination of the Swahili and Digo systems. These systems also pose a challenge for a formal analysis of copula constructions. In particular, as noted above, there might be a case for proposing that the richness of the copula inventories in the two languages is somewhat spurious, and only a surface impression, and that underlyingly the systems are in fact much simpler. An analysis along these lines would be based on proposing an empty copula for languages such as Digo and Swahili – indeed we postulated such an empty copula for both languages, but only in cases where there was no other morphological material (i.e. in cases of copula omission see Section 2.4). However, under the assumption that the empty copula is more widespread than this, the copula system of Swahili, for example, could be reduced to essentially two forms – the overt copula *ni* and the empty copula *Ø*.[7] The five other putative copula forms are, under this view, not independent copula forms, but rather complex expressions combining the empty copula with the relevant morpheme – e.g. the negation marker *si-* (53a) or the subject marker (53b):

(53) Swahili

a. Juma si-Ø mw-alimu.
1.Juma NEG-COP 1-teacher
'Juma is not a/the teacher.'

b. Nyumba i-Ø tupu.
9.house SM9-COP empty
'The house is empty.'

Such an analysis is supported by the independent existence of the relevant formatives to encode negation (*si-*), subject agreement (SM), possession and comitative marking (*-na*), and location (*-ko*), as well as by the appearance of the verb *-kuwa* 'be' in predication beyond the present tense:

(54) Swahili

Juma a-li-kuw-a na vi-tabu.
1.Juma SM1-PST-be-FV COM 8-book
'Juma had books.'

The verb *-kuwa* is found as replacing pure, locative, and possessive copulas in Swahili whenever the predication holds of non-present contexts. It could thus be analysed as replacing an empty copula as a suppletive form. An analysis along these lines would explain

[7] This is an idea that was also forwarded by McWorter (1994) who examines the diachronic development of the Swahili copulas. In fact, McWhorter (1994) proposes that the zero copula in Swahili was actually the older 'original' form on the basis that copula omission seemed to be even more widespread, even in written varities, in what he denotes as Early Modern Swahili than it is in present-day Swahili.

why it is the two languages which allow empty copulas on the surface that have the largest copula inventory. And under the assumption that the different forms are not independent copula forms, the copula systems of Digo and Swahili would become much smaller, and more similar to the other systems of our sample. However, we leave a detailed and more formalised exposition of this idea to a future occasion.

6 Conclusions

This paper has provided an overview of a number of key areas in which variation is found in Bantu copula constructions. This variation is seen in the function of copulas but also in their morphosyntactic properties, the elements with which they can combine, as well as their distribution in discourse. Foremost in the system of non-verbal predication in Bantu are the use of copula forms to link a wide range of predicates, prosodically marked predication and the presence of morphologically distinct copula forms, often with different interpretations and restricted distribution. Nominal and adjective predication, existence, location and possession can all be expressed in Bantu without a lexical verbal head. In terms of morphological form, the majority of the languages in the study employ an invariable copula form – often a variant of the copula *ni*. However, a defining feature of Bantu copula constructions appears also to be the presence of more than one copula form in a language.

After surveying different copula forms found in Bantu and the different constructions in which they can occur, we focussed on a small convenience sample of five Bantu languages – Mongo, Rangi, Digo, Swahili and Cuwabo – and highlighted comparative, typological and theoretical aspects of copula systems of the sample. The comparative approach revealed a number of patterns and correlations – for example, differences in complementation options of different copula forms, and the tighter restrictions on locative, existential and possessive interpretations across the sample, as compared to identificational and attributive interpretations. Based on this we highlighted the relevance of the data for the study of language change and language contact and their relation to each other, as well as for theoretical studies of copula, where we postulated an underlying empty copula as a means of providing a more parsimonious approach to the copula systems of Digo and Swahili.

Avenues for future research may involve the identification of more fine-grained parameters with a view to teasing out the differences between different copula forms and their corresponding interpretations. An approach that is accurately able to capture high level of variation found in the Bantu copula system may require additional questions and areas of research which are not covered in the present study. In the same vein, this area of investigation is limited by the varying descriptive status of the languages.

An additional avenue for future research would involve extending the study to include the so-called ‘semi-copulas’ and ‘copula verbs’ which have been excluded from the current paper in the interests of clarity and to maintain a strict focus on the ‘true copulas’. However, given the way in which these copula verbs often interact with the copula system as a whole, enabling for example, a wider range of tense-aspect interpretations than are possible with the invariable copula forms, extension of this research to this additional area would add to the richness of the study and further our understanding of how this domain works across the language family.

The study could also be extended to include additional languages. The languages surveyed here represent a convenience sample with languages chosen to reflect broad geographic diversity, as well as those which represent typological variation. However, with some 350 languages in the Bantu family, the study could be extended to yet more languages in order to

approach a (more) representative cross-section and to ensure a more equal geographical distribution of sample languages.

Further light could no doubt also be shed on the Bantu copula systems through examination and broader comparative work on the typology of copulas and non-verbal predication cross-linguistically. It may also be hoped that this kind of work would provide additional insights into the common pathways of development and grammaticalisation of copula forms and copula constructions, as well as the relationship between their origins and distributional properties. A more thorough descriptive understanding of copulas in Bantu may also facilitate the development of more fine-grained and comprehensive formal accounts of copula constructions from a cross-linguistic perspective.

References

Anward, Jan, Edith A. Moravcsik and Leon Stassen. 1997. 'Parts of speech: a challenge for typology.' *Linguistic Typology* 1: 167–183.

Ashton, Ethel O. 1947. *Swahili Grammar*. London, Longman.

Bearth, Thomas. 2003. 'Syntax'. In: *The Bantu Languages*. D. Nurse and G. Philippson (ed.) London, Routledge**:** 121–142.

Botne, Robert, Ochwada Hannington and Michael Marlo. 2006. *A Grammatical sketch of the Lusaamia Verb*. Köln, Rüdiger Köppe.

Buell, Leston and Merjin de Dreu. 2013. 'Subject raising in Zulu and the nature of PredP.' *The Linguistic Review* 2: 423-424.

Citko, Barbara. 2011. 'Small clauses.' *Language and Linguistics Compass* 5(10): 748-763.

Creissels, Denis. 2006. *Syntaxe générale, une introduction typologique 1: catégories et constructions*. Paris, Hermès.

Creissels, Denis. 2014. 'La prédication : une approche typologique.' *Verbum* 36(2).

Croft, William. 1999. 'Some contributions of typology to cognitive linguistics'. In: *Cognitive Linguistics: Foundations, Scope, and Methodology*. T. Janssen and G. Reker (ed.) Berlin, New York, Mouton de Gruyter**:** 61-93.

de Blois, K. F. 1970. 'The augment in Bantu languages.' *Africana Linguistica* IV.

Doke, Clement Martyn. 1961. *Textbook of Zulu Grammar*. Cape Town, Longmans.

Dugast, Idelette. 1971. *Grammaire du tunen* Paris, Edition Klincksieck.

Freeze, Ray. 1992. 'Existentials and other locatives.' *Language* 68: 553–595

Guérois, Rozenn. 2014. 'Locative inversion in Cuwabo.' *ZAS Papers in Linguistics* 57: 49-71.

Guérois, Rozenn. 2015. *A grammar of Cuwabo (Mozambique, P34)*. PhD. University of Lyon 2.

Guérois, Rozenn. 2015. Tonological evidence of the augment in Cuwabo. Annual Meeting of the Linguistics Association of Great Britain (LAGB). University College London.

Heine, Bernd and Tania Kuteva. 2005. *Language contact and grammatical change* Cambridge, Cambridge University Press.

Hengeveld, K. . 1992. *Non-verbal predication: Theory, Typology, Diachrony*. Berlin, Mouton de Gruyter.

Hulstaert, Gustaaf. 1965. *Grammaire du lomongo. Deuxième partie: Morphologie*. Tervuren, Musée royal de l'Afrique centrale.

Kavari, Jekura U., Lutz Marten and Jenneke van der Wal. 2012. 'Tone cases in Otjihereo: Head-complement relations, linear order and information structure.' *Africana Linguistica* 18: 315-353.

Lyons, John. 1967. 'A note on possessives, existential and locative sentences.' *Foundations of Language* 3: 390-396.

Lyons, John. 1977. *Semantics* Cambridge, Cambridge University Press.

Marten, Lutz. 2013. 'Structure and interpretation in Swahili existential constructions.' *Italian Journal of Linguistics*, 25(1): 45–73.

Maw, Joan. 1969. *Sentences in Swahili: A study of their internal relationships*, School of Oriental and African Studies, University of London.

McWhorter, John. 1994. 'From Focus Marker to Copula in Swahili'. Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: Special Session on Historical Issues in African Linguistics.

Meeussen, Achille Emille. 1967. *Bantu grammatical reconstructions*. Tervuren.

Möhlig, Wilhelm J. G. and Jekura U. Kavari. 2008. *Reference Grammar of Herero (Otjiherero)*. Cologne, Köppe.

Mould, M. 1974. 'The syntax and semantics of the initial vowel in Luganda'. Paper presented at the Third Annual Conference on African Linguistics.

Nicolle, Steve. 2013. *A Grammar of Digo: A Bantu Language of Kenya and Tanzania*, SIL International.

Petzell, Malin. 2008 *The Kagulu language of Tanzania: grammar, texts and vocabulary*. Köln, Rüdiger Köppe Verlag.

Pustet, Regina. 2003. *Copulas: Universals in the Categorization of the Lexicon: Universals in the Categorization of the Lexicon*. Oxford, Oxford University Press.

Schadeberg, Thilo C. 1992. *A Sketch of Swahili Morphology*. Cologne, Koeppe.

Schadeberg, Thilo C. and Francisco U. Mucanheia. 2000. *Ekoti: The Maka or Swahili Language of Angoche*. Cologne, Köppe.

Seidel, Frank. 2008. *A grammar of Yeyi: a Bantu language of southern Africa*, Köppe.

Stassen, Leon 1997. *Intransitive Predication*. Oxford, Clarendon Press.

van der Wal, Jenneke. 2006. 'Predicative tone lowering in Makhuwa.' *Linguistics in the Netherlands* 224-236.

van der Wal, Jenneke. 2009. *Word order and information structure in Makhua-Enahara*. PhD. University of Leiden.

Vogel, P. M. and Bernard Comrie. 2000. *Approaches to the Typology of Word Classes*. Berlin, New York, Mouton de Gruyter.

Welmers, William Everett. 1973. *African Language Structures*. Berkeley, University of California Press.

Zeller, Jochen. 2013. 'Locative inversion in Bantu and predication.' *Linguistics* 51(6): 1107–1146.

Appendix

Table 1: Features of non-verbal predication in a subset of Bantu languages

Language	Mongo			Rangi			Digo						Swahili							Cuwabo		
Form	*-le*	*-fa* NEG	*-ki*	*ni*	*sí* NEG	*-ri*	*Ø*	*ni*	*ndi-*	*SM*	*si(-)* NEG	*-na*	*Ø*	*ni*	*ndi-*	*SM*	*si* NEG	*-ko*	*-na*	DEM	*-li*	PL
Nominal complement	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No?	Yes	Yes	Yes	Yes	Yes	Yes
Adjectival complement	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	?	Yes	Yes	No	Yes	Yes	?	Yes	Yes	Yes	No?	No	(Yes)	Yes
Locative complement	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	?	Yes		No	Yes	No	No	No	No	Yes	No	No	Yes	No
Infinitival complement				Yes	Yes	Yes	(No)		?	?		(No)	Yes	Yes	?	No	(No)	?	?	No	Yes	Yes
Identificational interpretation				Yes	Yes	No	Yes	Yes	Yes	?	Yes	No	Yes	Yes	Yes	No	Yes	No?	No?	Yes	(No)	Yes
Attributive interpretation				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	(Yes)	No	(No)	Yes	Yes
Existential interpretation	Yes	Yes	Yes	No	No	Yes			?	?		Yes		No?	No?	No	No	Yes	Yes	No	Yes	No
Locative interpretation	Yes	Yes	Yes	No	No	Yes	Yes	Yes	?	Yes	Yes	(No)	No	No	No	No	No	Yes	Yes	No	Yes	No
Possessive interpretation				No	No	Yes	No	No	No	No	No	Yes	No	No	No	No	No	No	Yes	No	No	No
Temporal restrictions	Yes (Prs)	Yes	Yes (Pst)	Yes (Prs)	Yes (Prs)	No	?	No	(No)			No	?	Yes (Prs)	(No)	Yes (Prs)	Yes (Prs)	Yes (Prs)	Yes (Prs)	Yes (Prs)	Yes (Prs/ Pst)	Yes (Prs)