

# Non-verbal predication and copular variation in Eastern Bantu

©2024

Aron D. Finholt

M.A., University of Kansas, 2021

B.A., Western Washington University, 2019

Submitted to the graduate degree program in the Department of Linguistics and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

---

Chair: John Gluckman, Ph.D.

---

Andrew McKenzie, Ph.D.

---

Jie Zhang, Ph.D.

---

Peter Ojiambo, Ph.D.

---

Claire Halpert, Ph.D.

Date defended: May 8th 2024

The Dissertation Committee for Aron D. Finholt certifies  
that this is the approved version of the following dissertation:

Non-verbal predication and copular variation in Eastern Bantu

---

Chair: John Gluckman, Ph.D.

Date approved: May 8th 2024

## Abstract

This dissertation explores the nature of the copula through a typological investigation of copular variation in the expression of non-verbal predication in five Great Lakes Bantu languages, namely Kihavu (JD52), Mashi (JD53), Kinyamulenge (JD61a), Kirundi (JD62), and Kifuliiru (JD63). I show that the form of the copula in these languages is sensitive to particular environmental factors, resulting in similar patterns of copular allomorphy across languages. Specifically, two general patterns of allomorphy are observed. First, cognate forms of the copula are shown to appear in similar *interpretive* contexts. Much like the contrast between the Spanish copulas *estar* and *ser* (Clements 1988, 2006, Maienborn 2005, Deo et al. 2017), reflexes of the proto-Bantu copula *\*-de* (e.g., *-li/-ri*) are associated with limited, context-specific interpretations of the predication relation, while reflexes of the proto-Bantu copula *\*-bà* (e.g., *-ba*) are associated with more general interpretations. Second, cognate forms of the copula are shown to appear in similar *morphosyntactic* contexts. As in many other languages, the form of the copula is sensitive to various morphosyntactic factors, particularly tense/mood environment. Though multiple copular forms are attested in the present indicative, other tense/mood environments restrict the form of the copula; in general, only reflexes of proto-Bantu *\*-de* appear in the past tense, and only reflexes of proto-Bantu *\*-bà* appear in irrealis environments. I argue that these patterns reflect a core fact about the functional nature of the copula; in addition to its morphosyntactic function as a relational element (Lyons 1968, Hengeveld 1992, Stassen 1997, Pustet 2003, Bjorkman 2011, Myler 2016, a.o.), the copula also serves a semantic function as a means of distinguishing distinct types of non-verbal predication relations.

## Acknowledgements

This dissertation is the culmination of many years of coursework and research that would not have been possible without the care and support of my friends, family, and colleagues. There are many people who have had a significant impact on this work, and I am thankful for each of them and their respective contributions. The following acknowledgements highlight only a small subset of these individuals.

First and foremost, I would like to thank all of the wonderful language consultants I have had the privilege of working with during the development of this project. Thank you to Antoinette Bifuko, Mitima Burhungane, Fudjo Jeannot, Alice Kanyage, Peter Nitizeho, Justin Ngirabakunzi, Bonke Nkurunziza, Darlene Nyasafari, and J.P. Sabimpa for your time and for teaching me about your languages. This project would not have been possible without your knowledge and generosity.

I would also like to specially thank another person who had a monumental impact on this project, namely my advisor John Gluckman. I am extremely fortunate to have been able to work with John in various capacities during my time at the University of Kansas, and I am incredibly grateful for his support as a mentor. John has had a profound impact on my academic and personal development, and I would not be where I am today without him. Thanks as well to the other faculty in the Linguistics Department at the University of Kansas for their mentorship and support over the past five years, particularly Phil Duncan, Alison Gabriele, Andrew McKenzie, and Jie Zhang. Additional thanks to Peter Ojiambo and all those affiliated with the KiSwahili language program and the Kansas African Studies Center. Thanks as well to my committee member Claire Halpert at the University of Minnesota, and the entire faculty of the Linguistics Department at my alma mater, Western Washington University. I would also like to thank the various people and audiences that have provided input on various iterations of this work. Thanks in particular to the members of KUBantu and the Research in Field and Formal Linguistics seminar at the University of Kansas

for their weekly comments and suggestions. Thanks as well to the audiences of the conferences and talks where I have shared portions of this work, especially those at the Annual Conference on African Linguistics.

I would also like to express my gratitude towards the friends and family who have directly and indirectly supported me and my work over the past few years. I am particularly grateful for the support of Elizabeth Overschmidt, who has been there for me through all the challenges of graduate school. Thanks as well to the entire Overschmidt family for the overwhelming support and hospitality. I would also like to thank my cohort mates and close friends Andrew Collins and Henry Pratt for five years of camaraderie and fond memories. Thanks as well to Nick Feroce, Hyoju Kim, Jieun Lee, and Ruben Zamora for the friendship and support. Additional thanks to my friends and family back home in Washington for supporting me during and after my move to Kansas.

Lastly, I would like to give special thanks to my late mother, Bobbi Menzie, for her unconditional love and support throughout my life. My mother played an instrumental role in my academic development, and she was incredibly proud of my accomplishments as a researcher. Though I unfortunately do not get to celebrate this milestone moment with her, I know that she would have been incredibly proud of this achievement. With this in mind, I formally dedicate this dissertation to the memory of my mother, Bobbi Menzie. This work would not have been possible without you mom, and I wish I could share it with you in person.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Language background and methodology . . . . .	5
1.1.1	Language profiles . . . . .	6
1.1.2	Limitations of this study . . . . .	8
<b>2</b>	<b>Overview of Non-verbal Predication and Copular BE</b>	<b>10</b>
2.1	What is (non-verbal) predication? . . . . .	10
2.2	Non-verbal predication strategies . . . . .	15
2.2.1	Verbal copula . . . . .	16
2.2.2	Invariant copula . . . . .	19
2.2.3	Zero copula . . . . .	21
2.2.4	Locative copula . . . . .	22
2.2.5	Pronominal copula . . . . .	24
2.2.6	“Inflectional” copula . . . . .	25
2.2.7	Possessive copula . . . . .	26
2.2.8	Other strategies . . . . .	29
2.3	Complex copular systems . . . . .	31
2.3.1	Interpretive contrasts . . . . .	32
2.3.2	Types of copular clauses . . . . .	38
2.3.3	Morphosyntactic considerations . . . . .	42
2.4	Use of BE in other domains . . . . .	47
2.5	Summary . . . . .	49

<b>3</b>	<b>Theoretical Background</b>	<b>51</b>
3.1	Revisiting the copula . . . . .	51
3.2	Semantics of copular clauses . . . . .	55
3.2.1	How many copular clause types? . . . . .	55
3.2.2	Meaning of the copula . . . . .	59
3.3	Structure of copular clauses . . . . .	67
3.3.1	Realizing the copula . . . . .	67
3.3.2	Clause types and structure . . . . .	78
3.3.3	Interpretation and structure . . . . .	85
3.3.4	Deriving HAVE from BE . . . . .	89
3.4	Summary . . . . .	93
<b>4</b>	<b>Interpretation</b>	<b>94</b>
4.1	Introduction . . . . .	94
4.2	Utterances and situations . . . . .	95
4.2.1	Situation semantics . . . . .	97
4.2.2	Pragmatic approaches to copular contrasts . . . . .	98
4.3	Interpretive contrasts among Great Lakes Bantu copulas . . . . .	104
4.3.1	<i>-li</i> vs <i>-ba</i> . . . . .	105
4.3.1.1	Beyond temporary vs permanent . . . . .	110
4.3.2	<i>-ri</i> vs <i>ni</i> . . . . .	113
4.3.2.1	Additional complications . . . . .	117
4.3.3	<i>-li</i> vs <i>-ba</i> vs <i>-tula</i> . . . . .	124
4.3.3.1	Three-way contrast . . . . .	129
4.4	Analysis . . . . .	133
4.4.1	Approach . . . . .	135
4.4.2	Formal System . . . . .	141
4.4.3	Lexical entries . . . . .	143

4.4.4	<i>-li vs -ba</i> . . . . .	147
4.4.5	<i>-ri vs ni</i> . . . . .	148
4.4.6	<i>-li vs -ba vs -tula</i> . . . . .	152
4.5	Discussion and conclusion . . . . .	154
<b>5</b>	<b>Form</b>	<b>158</b>
5.1	Background . . . . .	162
5.1.1	Non-verbal predication in Bantu . . . . .	162
5.2	Data . . . . .	164
5.2.1	Interpretation . . . . .	164
5.2.2	Copular clause type . . . . .	165
5.2.2.1	JD50 languages . . . . .	166
5.2.2.2	JD60 languages . . . . .	168
5.2.3	Tense and Aspect . . . . .	171
5.2.3.1	JD50 languages . . . . .	171
5.2.3.2	JD60 languages . . . . .	176
5.2.4	Mood . . . . .	180
5.2.5	Other . . . . .	182
5.3	Analysis . . . . .	184
5.3.1	Background assumptions and schematics . . . . .	185
5.3.2	Kifuliiru copular system . . . . .	189
5.3.3	Kihavu and Mashi copular system(s) . . . . .	191
5.3.4	Kinyamulenge/Kirundi copular system(s) . . . . .	201
5.4	Discussion and conclusion . . . . .	210
<b>6</b>	<b>Conclusion</b>	<b>219</b>
	<b>Appendix</b>	<b>242</b>



## List of Tables

2.1	Comparing French and Spanish under Roy's (2013) Analysis . . . . .	38
4.1	Summary of copular forms in present tense predicational sentences . . . . .	105
4.2	Copular forms in present predicational sentences - Kinyamulenge and Kirundi . . .	114
4.3	Distribution of copulas by tense/mood environment - Kinyamulenge and Kirundi .	114
4.4	Copular forms and interpretations in Spanish and Modern Irish (Roy 2013: 189) . .	136
4.5	Copular forms and interpretations in Kihavu and Mashi à la Roy (2013) . . . . .	137
4.6	Comparing analyses of interpretive contrasts among copulas . . . . .	141
5.1	Summary of Bantu non-verbal predication strategies . . . . .	162
5.2	Copular forms in different tense/aspect environments . . . . .	179

# Chapter 1

## Introduction

Over the years, much of the conversation regarding non-verbal predication has focused on the function of the copula. In many of the world's languages, clausal predication involving a diagnostically non-verbal predicate (e.g., a nominal, adjective, adposition, etc.) requires the addition of a functional element — i.e., the copula — in order to relate the predicate to its subject. A fairly standard assumption in the literature is that this element is fundamentally meaningless; the sole purpose of the copula is to establish a predication relation in cases where the predicate cannot directly compose with the subject (Lyons 1968, Hengeveld 1992, Stassen 1997, Pustet 2003, Bjorkman 2011, Myler 2016). The English copular verb *be* for example is often thought of as a form of inflectional support, appearing only in cases where the predicate cannot inflect for tense or agreement.

- (1) a. John \*(is) happy  
b. John \*(is) a teacher

Seemingly at odds with this assumption is the fact that many languages employ multiple distinct copular forms in non-verbal predication. In contrast to languages like English that uniformly employ a single copula, e.g., *be*, languages such as Spanish (Maienborn 2005, a.o.), Modern Irish (Stenson 1981, a.o.), Thai (Warotamasikkkhadit 1972, Wongwattana 2015), and Washo (Bochnak et al. 2011), among many others (see Devitt 1994, Stassen 1997, Curnow 2000, Pustet 2003, Gibson et al. 2019, a.o.), utilize multiple copular forms to establish a predication relation between non-verbal predicates and their subjects. The cross-linguistic prevalence of such complex copular systems is surprising if we assume that the copula serves only to host inflection in cases where the predicate cannot; unless the form of the copula is sensitive to the particular properties of its environment, there is no reason

why any language should feature more than one copular form.

Looking across languages, we do in fact find that the form of the copula is sensitive to certain aspects of its environment. In many languages, the copular form is directly determined by morphosyntactic context; different copular forms correspond to different morphosyntactic environments. Depending on the language, the form of the copula may correspond to the categorial identity of the predicate, e.g., Bambara (Vydrin 2020, Sangare and Roy 2022), the tense-aspect-mood environment, e.g., Russian (Seres and Espinal 2019, a.o.), the person features of the subject, e.g., Lamba (Doke 1922, Lanham 1953), and/or the polarity of the clause, e.g., Arabic (Choueiri 2016) (see subsection 2.3.3). Since the respective copular forms in these languages exhibit distinct distributions, the difference between them can be reduced to contextual allomorphy; the form of the copula reflects the morphosyntactic context in which it is realized. In this sense, the effect of morphosyntactic context on the form of the copula is largely consistent with the idea that the copula serves a fundamentally morphosyntactic function. Though the form of the copula may change depending on its environment, its generally purpose is to relate two elements in a predicational relationship and host inflection.

That being said, copular variation cannot always be explained by morphosyntactic context alone; in some languages, distinct copular forms may appear in the same morphosyntactic environment. In these cases, copular contrasts are often reported to reflect differences in *interpretation*. In Spanish for example, the two copular verbs *estar* and *ser* are associated with distinct interpretations of the predication relation. In (2), *estar* yields a transient, stage-like property reading, while *ser* yields a more permanent, intrinsic property reading (see chapter 4 for further discussion of this contrast).

- (2) a. Su hermano **está** bueno  
Your brother estar.PRS.3SG good  
'Your brother is in good health/good-looking'
- b. Su hermano **es** bueno  
Your brother ser.PRS.3SG good  
'Your brother is a good person'
- Spanish (Ramsey 1894: 313)

The existence of interpretive contrasts like that observed in Spanish — and many other languages (see Ramchand 1997, Greenberg 1998, Green 2000, Bochnak et al. 2011, Deo et al. 2017, a.o.)

— directly challenges the idea that the copula serves a uniquely morphosyntactic function as a meaningless piece of functional material. In contrast to the distributionally distinct copular forms seen in other languages, the Spanish copulas *estar* and *ser* are meaningfully distinct, with each encoding different information about the interpretation of the predication relation. Though theoretical accounts often attempt to capture such interpretive contrasts using morphosyntactic and/or structural means (see subsection 3.3.3), the fact is that the surface form of the copula in these cases is semantically meaningful. While the copula appears to serve purely as a relational element that can host inflection in languages like English, it can also serve as a means of differentiating interpretively distinct types of predication relations in other languages.

As it turns out, a more comprehensive analysis of English copular *be* suggests that it too may interact with the semantics to some degree. In fact, there is a long philosophical tradition following Russell (1919) that assumes at least two (if not three or four) senses of copular *be* (see Kahn 1973, Devitt 1994, den Dikken 2006, a.o.). Consider for example the difference between the two English copular clauses in (3).

- (3) a. Paul is tall  
b. Clark Kent is Superman (Roy 2013: 8)

In contrast to its standard predicational use in (3a), the copula in (3b) serves to establish an identity relation. In this case *be* does not attribute the property denoted by the predicate to its subject, but rather equates the reference of two individuals. In this sense, the two copulas in (3) are semantically distinct; though both realized as *is* (*be*), the copula in (3a) serves a predicational function, while the copula in (3b) serves an equational function (see subsection 3.2.2). Though the contrast between predication and equation is not overt in English, similar distinctions are expressed via distinct copular forms in other languages, e.g., Thai (Warotamasikkhadit 1972, Wongwattana 2015). Just like the contrast between Spanish *estar/ser*, this suggests that the copula serves a semantic function as a means of distinguishing different types of predication. The difference between languages is simply whether the form of the copula overtly reflects these differences.

The point I am trying to make here — and ultimately, one of the major points I try to make

in this dissertation — is that the copula is simultaneously sensitive to two different components of the grammar. On the one hand, the copula serves an ostensibly morphosyntactic function as a relational element that hosts inflection. Correspondingly, copular variation across languages is sometimes morphosyntactically determined; in some cases the form of the copula is determined by the morphosyntactic context in which it appears. On the other hand, the copula also serves a semantic function as a means of distinguishing interpretively distinct types of predication relations. Likewise, copular variation is sometimes interpretively determined; in some cases the form of the copula is determined by the interpretive context of the predication relation it helps establish. In this sense, copular variation across languages can be seen as a response to the functional duality of the copula.

In this dissertation, I explore the nature of the copula through a typological investigation of copular variation that focuses on the effect of both morphosyntactic and interpretive environment. In an effort to extend the typology and comment on various theoretical issues regarding non-verbal predication, I present a novel, micro-typological investigation of copular variation in a select group of Great Lakes Bantu languages, specifically Kihavu (JD52), Mashi (JD53), Kinyamulenge (JD61a), Kirundi (JD62), and Kifuliiru (JD63).<sup>1</sup> As I will demonstrate in the chapters that follow, these languages exhibit extensive and highly complex patterns of copular allomorphy that are conditioned by a variety of morphosyntactic and interpretive factors, making them ideal languages for an investigation of this type. I show that the form of the copula in the languages surveyed is sensitive to different combinations of environmental factors including copular clause type, tense-aspect-mood environment, predicate category, subject person, and interpretive context. Though individual languages are found to differ in terms of the copular forms they employ and the specific combination of factors they deem relevant to copular form, I ultimately show that there are generalizable

---

<sup>1</sup>Throughout this work, Bantu languages are indexed according to the classificational system presented in Maho (2009), which serves as an expansion and evolution of Malcolm Guthrie's (1948, 1967 1971) classification of the Bantu languages. Indexical information is provided as a series of letters and numbers in parentheses following the language name. Letters serve to identify the language zone, and numbers identify distinct subgroup(s) and languages within them. Maho's classification includes the Tervuren zone J, which unifies some of the distinct zones originally proposed by Guthrie (1948, 1967 1971), e.g. Kihavu (JD52) is classified as a JD zone (or Guthrie zone D) language within the JD50 subgroup, like Mashi (JD53).

patterns of variation across languages; cognate forms tend to exhibit similar interpretive profiles and morphosyntactic restrictions in different languages. In response to the observed patterns, I propose two separate analyses of the copular variation data: the first being a semantic account that addresses interpretive differences among copular forms (chapter 4), and the second being a contextual allomorphy account that addresses distributional differences among copular forms (chapter 5). In developing these two proposals, I comment broadly on some of the major theoretical questions related to non-verbal predication and the copula, including whether the copula is a meaningless syntactic element, whether there are multiple kinds of (semantic) copulas, how many types of copular clauses there are, and what kinds of interpretive contrasts are possible.

To my knowledge, this dissertation represents one of the first efforts to provide both a semantic and contextual allomorphy account of copular variation in Bantu languages. As such, in addition to its typological and theoretical contributions to the literature on non-verbal predication, this work also makes valuable contributions to the field of African linguistics more broadly, particularly from a Bantuist perspective. First, this work addresses a salient gap in the description of Bantu languages, namely what drives copular variation; in some Bantu languages, the form of the copula is sensitive to both interpretive and morphosyntactic factors. Second, this work further highlights the tremendous degree of linguistic diversity in Sub-Saharan Africa, emphasizing yet another point of micro-variation among Bantu languages in particular; variation in the particular copular forms used in non-verbal predication can be seen even in closely-related Bantu languages spoken in the same geographical area. Finally, despite this variation, this work also demonstrates that there are generalizable patterns in the expression of non-verbal predication across Bantu languages; though Bantu languages regularly feature different combinations of copular forms, their copular systems often feature similar interpretive and distributional distinctions.

## **1.1 Language background and methodology**

This work provides a novel description of copular variation in a group of related Great Lakes Bantu languages, namely Kihavu (JD52), Mashi (JD53), Kinyamulenge (JD61a), Kirundi (JD62),

and Kifuliiru (JD63), that are spoken in the border region that comprises the East African states of Rwanda, Burundi, and the Democratic Republic of Congo. These languages belong to two classificational subgroups (e.g., the JD50 and JD60 groups) of the Great Lakes (or Lacustrine) Bantu group, which is one of the major classificational groups that make up the Northeast Bantu branch of the Bantu language family (Guthrie 1948).

Data from these languages was primarily collected as part of collaborative field research with my advisor John Gluckman and other members of KUBantu, the Bantu research and reading group at the University of Kansas. Data was collected via direct elicitation with native speaker consultants living in the Greater Kansas City area (Kansas and Missouri) and Kigali, Rwanda, both of which feature sizable populations of resettled Congolese refugees. Data collection proceeded over a total period of three years (Fall 2020 - Summer 2023). Elicitation techniques used to gather data include direct translation, storyboards, well-formedness judgement tasks, and semantic judgement tasks including truth-value and felicity-judgment tasks. Elicitation sessions primarily involved translation to and from English, however Swahili and French were also used.

Broadly speaking, all of the languages surveyed in this project exhibit prototypical morphosyntactic properties of Bantu languages, including base SVO word order (though information structure plays a role in determining surface word order), an extensive noun class system, and rich inflectional morphology, among others. Like many other Eastern Bantu languages, tone plays an important role in these languages, with most if not all languages featuring productive grammatical tone contrasts (see Marlo 2013 for discussion of Bantu tone). That being said, tone has been omitted from almost all of the novel examples presented in this work due to time and/or data limitations.

### **1.1.1 Language profiles**

The majority of the data in this work represents novel field data from five Great Lakes Bantu languages, e.g., Kihavu (JD52), Mashi (JD53), Kinyamulenge (JD61a), Kirundi (JD62), and Kifuliiru (JD63). External data from other closely related Bantu languages — most notably Kinande (JD42) and Kinyarwanda (JD61) — is sometimes discussed alongside this data for comparative or

typological purposes. In any case, I restrict much of the discussion and analysis in this work to the aforementioned five languages. Short descriptions of these languages is provided below.

Kihavu (JD52) — or Havu/Haavu — is the primary language of the Bahavu, an ethnolinguistic group based primarily on the island of Idjwi and the neighboring of Kalehe in South Kivu, Democratic Republic of Congo. Updated speaker population figures are difficult to find, however estimates from the early 2000s list approximately 500,000 native speakers of Kihavu (Ethnologue 2024). Like many other languages in the region, Kihavu is very poorly documented overall. Relatively few descriptions of the language exist in the literature, with the most notable being that of Birusha (1985). The Kihavu data presented in this work were collected from a single native speaker living in the Greater Kansas City area between Fall 2021 - Spring 2023.

Mashi (JD53) — or Shi — is the primary language of the Bashi, one of the largest ethnolinguistic groups in South Kivu, Democratic Republic of Congo centered near the shores of Lake Kivu between Kalahe and Bukavu. The exact number of native speakers is unknown, however figures from the early 1990s list a population over 600,000 (Ethnologue 2024). Few descriptions of the language can be found in the literature, with the notable exception of Polack-Bynon (1975). The language is very closely related to neighboring Kihavu (JD52), with the two languages being mutually-intelligible in some cases. The Mashi data in this work were collected from a single native speaker living in the Greater Kansas City area (not the same person as the Kihavu consultant) between Fall 2021 - Spring 2023.

Kinyamulenge (JD61a) is a language spoken by the Banyamulenge, a subgroup of the pan-Rwandan ethnolinguistic group known as the Banyarwanda that is predominantly located in modern day South Kivu province, Democratic Republic of Congo. Given the significant similarities between the two, Kinyamulenge is often considered as a “dialect” of Kinyarwanda (JD61), even by some Kinyamulenge speakers. For this reason, among others, the exact number of speakers of Kinyamulenge is unknown. Compared to other languages in the region, Kinyamulenge is particularly poorly documented, likely due to its status as a “dialect”. The Kinyamulenge data included in this work were collected from five separate consultants (two from the Greater Kansas City area, and



three from and Kigali, Rwanda) between Fall 2020 - Summer 2023.

Kirundi (JD62) is the dominant national language of Burundi having well over 10 million native speakers according to Ethnologue (2024). Unlike some other closely-related languages, limited descriptions of Kirundi can be found in the literature (see chapter 5). Unless otherwise noted, the Kirundi data in this work were collected from a single native speaker living in Kigali, Rwanda take from novel field research in Kigali, Rwanda in Summer 2023.

Kifuliiru (JD63) is a language spoken by the Bafuliiru, an entholinguistic group located near the Burundi border in South Kivu, Democratic Republic of Congo. The latest estimate from 2012 lists a native speaker population of around 400,000 (Ethnologue 2024). Once again, a limited number of descriptions of Kifuliiru can be found in the literature, most notably the two volume grammar by (Otterloo 2011a) and (Otterloo 2011b). That said, the Kifuliiru data in this work were collected from a single native speaker living in the Greater Kansas City area between Spring 2022 - Spring 2023.

### **1.1.2 Limitations of this study**

From a methodological perspective there are two major limitations of this study: the number of languages investigated, and the number of native speaker language informants consulted for each language. Regarding the first of these limitations, this work focuses narrowly on copular variation in only five Great Lakes Bantu languages. Importantly, these languages were not chosen arbitrarily. In fact, these specific languages were chosen for two reasons. The first relates to the research question(s) explored in this work; these languages exhibit highly complex patterns of copular variation and are therefore ideally situated for a study focused on non-verbal predication and copular variation. The second reason these languages in particular were chosen is more practical. The Greater Kansas City area features a large population of resettled Congolese refugees who have fled from the eastern provinces of the Democratic Republic of Congo (e.g., North Kivu and South Kivu) due to sustained periods of violence and unrest. For this reason, there are substantial speaker populations of many Great Lakes Bantu languages in the Kansas City area, including the

five investigated in this work. Since Fall 2020, I have contributed to a collaborative field research project involving my advisor John Gluckman and the other members of KUBantu, the Bantu reading and research group at the University of Kansas, that is focused on the identification and description of Eastern Bantu languages spoken in the Greater Kansas City area. This dissertation and the data presented within it were developed as part of this collaborative project, and is therefore limited to a subset of the languages focused on in that work. Ideally a larger typological study of copular variation in Eastern Bantu languages would involve far more than five languages and two classificational subfamilies, however a project of this size was not possible for this work due to time constraints. Future work is necessary to determine whether the patterns shown in this work extend broadly to Great Lakes Bantu languages and/or Bantu languages in general.

The second major methodological limitation of this study relates to the number of language informants consulted during data collection. As mentioned above, there is some variation with respect to how many native speakers were consulted in each language: data from Kinyamulenge (JD61a) were collected from five different native speakers, while data from the other four languages in this study were collected from a single, representative native speaker of each language. Ideally, a typological study of this nature would involve data collected from multiple language consultants for each language investigated, however this was ultimately not possible in this project due to time/other constraints. For this reason, it is possible that some of the judgements in this work may not align with the judgements of all other speakers' of the same language. Further work is necessary to determine whether there is variation — and if so, to what degree — across speakers of the languages in question with regard to copular variation. In any case, I take the data presented in this work to be representative of at least one variety of the languages in question, though I acknowledge that there may be differences across speakers.

## Chapter 2

### Overview of Non-verbal Predication and Copular BE

In this chapter, I explore the typological landscape of non-verbal predication through a broad survey of copular systems across languages. In section 2.1 I provide general background on some key terms and concepts related to non-verbal predication. In section 2.2 I provide a comparative overview of non-verbal predication strategies/copular forms and their diachronic sources, with a particular focus on Bantu languages. In section 2.3 I turn my attention to languages that feature multiple distinct copulas, and lay out some morphosyntactic and interpretive factors that have been shown to play a role in the distribution of copulas. I finish the chapter with a brief discussion of copulas outside of non-verbal predication contexts in section 2.4.

#### 2.1 What is (non-verbal) predication?

The linguistic phenomena of *predication* has long been a central topic of inquiry at the syntax-semantics interface. As discussed in den Dikken (2006), a.o., attempts to describe the relationship between a predicate and its subject can be traced back as far as the philosophical works of Plato and Aristotle.<sup>1</sup> Over the years, numerous accounts have attempted to make sense of this relationship, both in terms of its structural representation and its semantic interpretation. Within this broader research domain, considerable attention has been paid to non-verbal predication in particular, as this has been shown to differ from standard, verbal predication in a number of ways (den Dikken 2006, Roy 2013).

In its most general sense, predication describes a relation between a property-denoting syntactic

---

<sup>1</sup>Plato first defines the sentence as the conjunction of a subject and predicate in *Sophist* (see den Dikken 2006). This idea is further explored in later works by Aristotle, notably *De Interpretatione* (see Ackrill 1963, den Dikken 2006).

constituent, the predicate, and the syntactic constituent it attributes a property to, the subject (den Dikken 2006). In a more semantic sense, a predicate is thought of as a property-denoting function which takes as its argument the subject. For example, English intransitives like *smoke* in (4) are said to be predicative in that they attribute a property to the syntactic argument they select, e.g., the subject. This is, broadly, the traditional Fregean view of predication, where predicates are characterized as functions that directly relate individuals and propositions (Frege 1897, Heim and Kratzer 1998). In a sentence like *Ann smokes*, the verbal predicate  $\llbracket\text{smoke}\rrbracket$  is a function that takes an individual,  $x$ , and asserts that  $x$  smokes.

- (4) a. Ann smokes  
b.  $[\lambda x. \text{smoke}(x)](\text{Ann})$   
c.  $\text{smoke}(\text{Ann})$

Although categorically verbal predicates like English *smoke* most clearly fit the definition of “predicate” in the Fregean sense, predicates do not need not be verbal. For example, English adjectives and nominals may also express property concepts — that is, function as semantic predicates — despite being unable to inflect for tense and person/number agreement like verbs. In examples like (5)-(6), *happy* and *teacher* act as predicates in that they denote functions that take an individual, e.g., the subject, and assert that a property holds of that individual.

- (5) a. John is happy  
b.  $[\lambda x. \text{happy}(x)](\text{John})$   
c.  $\text{happy}(\text{John})$
- (6) a. John is a teacher  
b.  $[\lambda x. \text{teacher}(x)](\text{John})$   
c.  $\text{teacher}(\text{John})$

With respect to the semantics, non-verbal predicates (e.g., *happy/teacher*) are functionally identical to verbal predicates (e.g., *smoke*) in that they are both property-denoting; they attribute

properties to individuals. Importantly however, English non-verbal predicates crucially differ from verbal predicates in terms of how they are expressed in the syntax. This is most apparent in the prototypical, fully clausal examples of predication we have seen to this point. Since adjectives and nouns cannot directly inflect for tense or agreement in English, when used as a clausal predicate, they require an additional piece of functional material — namely copular *be* — to establish the predication relation and carry inflectional information (Lyons 1968, Hengeveld 1992, Stassen 1997, Pustet 2003, Bjorkman 2011, Myler 2016). In this sense, English *be* in (5)-(6) serves a purely syntactic function — it provides a means of expressing tense and relating a non-verbal predicate to its subject, but doesn't itself contribute any meaning.

The idea that copular *be* is a meaningless piece of functional material has long been the default assumption in work on non-verbal predication. Consider the definition provided in Pustet (2003: 5), which summarizes the general position taken by those who assume a meaningless copula.

- (7) “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.”

Although Pustet's definition is relatively conservative in the sense that it does not specify what kinds of predicates copulas should and should not appear with, it is straightforwardly extendable to the use of copulas that we have seen thus far. Copulas co-occur with non-verbal predicates because non-verbal predicates function as predicate nuclei, but cannot directly compose with their subjects for some reason, e.g., they cannot host inflection. Though I will return to this definition — and the notion that the copula is “meaningless” — in later chapters (see chapter 3 and chapter 4 in particular), the key point for the time being is that the copula is thought of as an additional piece of functional material that sometimes appears in non-verbal predication.

From a typological viewpoint, it is relatively common to find languages that overtly distinguish between verbal and non-verbal predication in this way. In fact, one of the more prevalent non-verbal predication strategies we observe involves the use of a copular element, just as in English. For example, French similarly utilizes a single *be*-verb, *être*, to express non-verbal predication (Roy

2013); adjectival, nominal, and prepositional predicates in French all require the presence of the copular *be*-verb in main clauses (8)-(10).

(8) Paul **est** malade  
Paul is sick  
'Paul is sick'

(9) Paul **est** un acteur  
Paul is an actor  
'Paul is an actor'

(10) La Joconde **est** au Louvre  
the Mona.Lisa is at.the Louvre  
'The Mona Lisa is at the Louvre'

French (Roy 2013: 79)

Though languages differ significantly in how they express non-verbal predication (see section 2.2), the use of a copular *be*-verb as relator and host for inflectional features is commonly observed across the world's languages (Clarke 1978, Devitt 1994, Stassen 1997, Curnow 2000, Pustet 2003). That being said, non-verbal predicates are not unique to primary (main clause) predications in English or other languages (den Dikken 2006). Indeed, non-verbal predication can be found in a variety of other morphosyntactic contexts as well, including those that do not involve tense.

Consider so-called small clauses in English, which appear as complements to epistemic, positional, and dynamic verbs (Williams 1975, Stowell 1981, den Dikken 2006). As the name suggests, a small clause is defined as a fully propositional clause that is "smaller" than a tensed clause, i.e., that lacks tense. In contrast to primary, main clause predications like (5)-(6), the small clause predicates (in bold) in (11a-d) enter into a predication relation with the italicized subjects without the need for additional functional morphology; since there is no tense, there is no need for an overt copula.<sup>2</sup> Note that the absence of tense in small clauses also precludes them from occurring with truly verbal predicates, e.g., \*Brian considers [<sub>SC</sub> *Imogen* **smoke** ].

(11) a. Brian considers [<sub>SC</sub> *Imogen* **smart** ]

---

<sup>2</sup>At least, there is no need for an inflectional *be*-verb. Other types of copulas, e.g., invariant particles like *mo* in Kinande (Bantu; Schneider-Zioga 2018), often appear in tenseless environments like small clauses (den Dikken 2006).

- b. Brian hung [<sub>SC</sub> *his shirt on the line* ]
- c. Brian hammered [<sub>SC</sub> *the metal flat* ]
- d. Brian ran [<sub>SC</sub> *the pavement thin* ] (Adapted from den Dikken 2006: 58)

Importantly, small clauses are assumed to be structurally identical to main clause predication barring the absence of tense. That is, like their main clause counterparts, e.g., *Imogen is smart*, *his shirt is on the line*, etc., the small clauses in (11) are argued to involve a relational element that allows a non-verbal predicate to compose with the subject. This relational element is assumed to be represented syntactically by the functional head Pr/Pred (Williams 1983, Bowers 1993), which serves to mediate subject-predicate relations in main clauses and small clauses. In the case of small clauses, the Pred head is sometimes unrealized at PF as shown above, though it need not always be. For example, English small clauses sometimes allow *as* to intervene between the subject and predicate of the small clause, suggesting that a relational, predicative head, e.g., Pred, must be involved in small clauses (12).

- (12) a. Imogen regards [<sub>SC</sub> Brian \*(**as**) a nice guy ]
- b. Imogen regards [<sub>SC</sub> Brian \*(**as**) nice ]
- c. Imogen views [<sub>SC</sub> Brian \*(**as**) a nice guy ]
- d. Imogen finds [<sub>SC</sub> Brian \*(**as**) a nice guy ]
- e. Imogen considers [<sub>SC</sub> Brian (**as**) a nice guy ] (Adapted from den Dikken 2006: 34)

Though there is a sense that English *as* is functionally distinct from copular *be*, it ostensibly serves a predicative function in small clauses in the same way that *be* does in main clauses. In this sense, *as* and *be* both serve to mediate a subject-predicate relation in their respective environments, i.e., they both realize Pred.

There is considerable evidence of a functional Pred head in languages other than English as well. For example, small clauses in some Bantu languages are reported to involve a copular element between the subject and predicate. Consider the Eastern Bantu language Kinande (JD42), which

features an invariant copular particle *mo* in both small clauses and depictive secondary modification (Schneider-Zioga and Mutaka 2015a, Schneider-Zioga 2018).<sup>3,4</sup>

- (13) a. Ngálangira [sc Nadiné **mo** mu-búya ]  
 1SG.SM.see Nadine COP 1AGR-beautiful  
 ‘I find Nadine beautiful’
- b. Kám bale mwálywa [sc e-nyamá **mó** mbisi ]  
 Kambale 1SM.PST.eat AUG-9.meat COP 9AGR.raw  
 ‘Kambale ate the meat raw’ Kinande (Schneider-Zioga and Mutaka 2015a: 86-90)

Though I will not dig into the structural details here (see chapter 3), the point is that non-verbal predication relations can be found in clauses of variable sizes, and they always involve the same core structure, i.e., a subject-predicate relation mediated by a Pred head. That being said, I will focus much of the discussion in this dissertation on “primary” — or main clause — predication, as this is where we see arguably the most diversity across languages (Curnow 2000, Schneider-Zioga 2018, Gibson et al. 2019).

## 2.2 Non-verbal predication strategies

One of the central observations in previous work on non-verbal predication is that it often involves additional functional material — i.e., the copula — that serves to host grammatical features and relate the predicate to its subject. We have already seen an example of this with the distribution of the copula in English; copular *be* is used to establish a non-verbal predication relation in main clauses where tense would otherwise go unrealized (5)-(6), but not in small clauses where there is no tense (11). This sort of distribution has lead many to assume that the copula is a semantically

<sup>3</sup>In general, all glossed examples presented in this work adhere to the Leipzig Glossing Rules. However, glossing of Bantu language data includes additional language-specific conventions and terminology. Per Bantuist tradition, noun class identity is indicated in the gloss via noun class number. Proper names are morphologically unmarked, but generally pattern as class 1 nouns since their referents belong in class 1. Subject noun class agreement is marked on the verb via a subject marker prefix. Agreement with class 1/2 nouns (e.g., human nouns) is subdivided by person: 1SG/PL for first person, 2SG/PL for second person, and 1SM/2SM for third person. The full list of glossing abbreviations used in this work can be found in the Appendix.

<sup>4</sup>Although secondary predication has often been analyzed as involving a small clause (Kayne 1984, Kratzer 2005, Harley 2005, a.o.), it is unclear if this is truly the case (see Bruening 2018 for considerable evidence against a small clause analysis). In the case of Kinande, the point is simply that small clauses and secondary predication both involve an invariant copular particle where we would expect a relational Pred head to be.



empty element that is little more than a relational element that can host inflectional features (Lyons 1968, Hengeveld 1992, Stassen 1997, Pustet 2003, Myler 2018).

Though the accepted definition of the copula paints a fairly simplistic picture of its function, the true nature of the copula is significantly more complex than such a description can convey. The particular dimension of this complexity that we will discuss here concerns the form of the copula and the various means through which languages may express non-verbal predication. In the following subsections, I will highlight the typological variation we see in the expression of non-verbal predication through a survey of eight prominent non-verbal predication strategies reported across languages. In so doing, I will also broadly comment on how copulas come to be: languages often re-purpose existing morphosyntactic material to serve a predicative function. Moreover, I will show that the extent of copular variation we observe across languages is mirrored by the variation within Bantu languages alone; all of the strategies discussed below are reported in Bantu languages (Schneider-Zioga 2018, Gibson et al. 2019).

### 2.2.1 Verbal copula

As we have seen, one of the more common means of encoding non-verbal predication is via an inflectional copular verb that is marked for person and/or number agreement (Devitt 1994, Stassen 1997, Curnow 2000, Pustet 2003). Consider for example Modern Greek, which uses an inflected form of the *be*-verb *einai*, e.g., third singular *íne* (14), in standard copular clauses (Devitt 1994).

- (14) aftós                    **íne**      anthropos  
       3SG.MASC.PRO 3SG.be human  
       ‘He is a man’ Modern Greek (Devitt 1994: 50)

There are many other languages that similarly employ an inflectional *be*-verb like Modern Greek *einai* to express non-verbal predication, including a vast swathe of Indo-European languages (Stassen 1997). For this reason, the terms “verbal copula” and “*be*-verb” are sometimes assumed to be interchangeable in the literature, though this may be somewhat of an overgeneralization. In fact, there are many languages which employ a verbal copula that is ostensibly not a ‘*be*-verb’. One

salient example of this is the use of *have*-verbs in non-verbal predication (Francez and Koontz-Garboden 2017). In French for example, certain property concept predicates, e.g., *faim* ‘hunger’ (Stassen 1997), require the verb *avoir* ‘have’ to establish a predication relation with the subject.

- (15) a. Pierre is hungry
- b. Pierre a faim  
 Pierre have.PRS.3SG hunger  
 ‘Pierre is hungry’ French (Francez and Koontz-Garboden 2017: 22)

Broadly speaking, *avoir* in (15) serves the same purpose as the *be*-verb *être* that appears with other non-verbal predicates (8): it serves to host tense and relate the subject and predicate. Nonetheless, there is an intuitive sense that *avoir* and *être* fall into distinct categories of “verbal copula” since they are distinct synchronic verbs.

Even if we restrict ourselves to dedicated synchronic *be*-verbs, it is quite difficult to pinpoint what exactly a “true *be*-verb” really is, if such a category exists at all. In general, copular verbs — including *be*-verbs — are assumed to have been derived through various processes of grammaticalization (Stassen 1997). For example, many copular verbs are reported to have lexical sources as dynamic verbs, e.g., ‘do/make/build’, ‘happen/occur’, ‘go/turn into/come/become’, ‘act (like)’ etc., which “designate processes through which a certain state of affairs comes about” (Stassen 1997: 93). Over time the meaning of dynamic verbs can be extended to encapsulate not just processes, but also their resultant states, at which point they may be reanalyzed as copular verbs (Langacker 1975, Stassen 1997). In some languages, the lexical origin of copular verbs is transparent. For example, (Willms 1972: 259) reports that the synchronic copula *g/eg* in the Berber languages Tamazight and Shilha is actually a verb meaning ‘make, do, become, happen’. Similarly, the copular verb *zywe* in Lisu (Tibeto-Burman) is reported to mean ‘act like, perform the function of’ (Hope 1974: 38). In other languages, the connection is less clear synchronically. For example, the Indo-European root *\*-bheu-/\*bhu-* — the source of ‘*be*’-roots in Indo-European languages, e.g., English *be*, Welsh *bod*, German 1SG.PRS *bin*, etc. — is assumed to have an origin as a dynamic verb meaning ‘to grow’ (Stassen 1997).

In addition to dynamic verbs, locational verbs are also reported as a common source of copular verbs. Indo-European *\*sta-*, e.g., Irish *ta*, Spanish *estar*, etc., is reconstructed as a posture verb meaning ‘stand’ (Stassen 1997). As discussed in Stassen (1997), there are many other diachronic pathways for copular verbs. That said, the point I am trying to make here is that the source of *be*-verbs — or copular verbs more broadly — varies significantly across languages. The relevance of these sources to the interpretation and function of individual copulas is something that I will return to extensively in later chapters. For now it suffices to say that differences in the development of copulas across languages makes it difficult to define sharp distinctions between sub-classes of copular verbs.

Returning to our broader typological discussion, we have seen that many languages employ some sort of verbal copula to express non-verbal predication. Given their ubiquity across languages, it is relatively unsurprising that copular verbs — in the form of inflectional *be*-verbs — are also widely observed among Bantu languages (Gibson et al. 2019). In general, Bantu languages tend to employ one (or more) related forms of the proto-Bantu copulas *\*-de* (or *\*-dì*) ‘be’ and *\*-bà* ‘be, become’, usually *-li* or *-ba*, respectively (Meeussen 1967, Guthrie 1967 1971, Wald 1973).

For example, Mongo (C61) employs an agreeing copular verb *-le* (*\*-de*) in pure predicational sentences like (16). As a fully inflectional copular verb, *-le* obligatorily expresses noun class agreement with the subject, e.g., class 1 *nkómbé* ‘kite’ (16a). This rich system of agreement is amenable to pro-drop (16b)-(16c), which is a commonly attested feature of Bantu languages.

- (16) a. Nkómbé **a-le** mpúlu  
 1.kite 1SM-be 9.bird  
 ‘The kite is a bird’
- b. **A-le** bɔ-nénɛ  
 1SM-be 1AGR-big  
 ‘He is big’
- c. **Ba-le** ba-laki  
 2SM-be 2AGR-teacher  
 ‘They are teachers’

Mongo (Hulstaert 1965: 340)

Though cognates of *-li* and *-ba* are the most commonly reported *be*-verb forms in Bantu, others

are reported as well. For example, both Lamba (M54; Doke 1922) and Kifuliiru (JD63; Finholt 2023) exhibit a third copular verb in addition to *-li* and *-ba*; *-ikala* (\*-yìkad; Guthrie 1967 1971) ‘remain, stay’ in Lamba (Doke 1922, Schneider-Zioga 2018), and *-tula* ‘dwell, inhabit, live’ in Kifuliiru.<sup>5</sup> As I will discuss more extensively later in this chapter, other forms of the copula — including other verbal copulas — sometimes appear in restricted morphosyntactic environments. Mongo for example exhibits another copular form *-ki* that uniquely appears in the past tense (17).

- (17) **N-ki**            bo-sáj’    ò-káé  
           2SM-be.PST 1-worker 1AGR-POSS3SG  
           ‘I was his worker’ Mongo (Hulstaert 1965: 340)

In sum, there are many different types of copular verbs employed across languages. Focusing narrowly on Bantu languages, we see a similar degree of variation; there are many different copular verbs in Bantu, sometimes even in a single language.

### 2.2.2 Invariant copula

Though many languages utilize a dedicated copular “*be*-verb” to express non-verbal predication, copular forms are not always fully verbal. In place of — or sometimes alongside — an inflectional *be*-verb, many languages employ an invariant particle to establish a predication relation. For example, one way of expressing non-verbal predication in Modern Irish is through the use of the invariant copular particle, *is*, which does not inflect for person, number, or tense (18) (Doherty 1996).

- (18) a. **Is**    duine    maith    è  
           COP    person    good    him.ACC  
           ‘He is a good person’
- b. **Is**    dh’a    rud    éagsúla    iad  
           COP    two    thing    distinct    them.ACC  
           ‘They are two distinct things’ Modern Irish (Doherty 1996: 35)

---

<sup>5</sup>The historical source of *-tula* is unclear. However, it appears to be related to Kinyarwanda (JD61) *-tura* ‘dwell, inhabit, live’, which may provide some insight into its source.

Similarly, the Mandarin invariant particle *shi* — which also serves as a focus marker — is sometimes used as a copula with nominal predicates; though optional in declarative sentences, *shi* is obligatory under negation (Hengeveld 1990).

- (19) a. Zhangsan (**shi**) yi-ge hushi  
 Zhangsan COP one-CL nurse  
 ‘Zhangsan is a nurse’
- b. Wo bu **shi** Zhongguo-ren  
 1SG.PRO NEG COP China-person  
 ‘I am not Chinese’
- Mandarin (Hengeveld 1990: 6)

Turning our attention to Bantu, we again find that this particular strategy of encoding non-verbal predication is widely attested. In fact, many Bantu languages like Digo (E73) and Swahili (G42) utilize an invariant form of the copula — generally *ni* or *nu* — to establish a predication relation between the predicate and its subject, regardless of subject noun class (20)-(21).

- (20) Mutu hiyu **ni** daktari  
 1.person 1DEM be 1.doctor  
 ‘This person is a doctor’
- Digo (Nicolle 2013: 298)
- (21) Kaka yangu **ni** m-refu  
 1.brother 1AGR.POSS1SG be 1AGR-tall  
 ‘My brother is tall’
- Swahili (Marten 2013: 62)

Though generally treated as invariant *be*-verbs, there is a growing body of literature which suggests that the copular use of *ni* — in Swahili and other Bantu languages — is derived from its use as a focus marker elsewhere (McWhorter 1994, Schwarz 2003, Güldemann 2003). Note that this mirrors what is observed of Mandarin *shi*, which functions as both a copular particle and focus marker (Hengeveld 1992). As I will see, focus turns out to be an important factor in the form of the copula in Bantu languages and beyond. I will return to discuss the relationship between focus and the copula in later chapters, particularly chapter 4.

### 2.2.3 Zero copula

While many languages utilize overt morphosyntactic material to help realize a non-verbal predication relation, some languages allow a predicate to directly compose with its subject without any additional functional material. Consider Japanese, which generally expresses non-verbal predication via the copular clitic, *da* (22).<sup>6</sup>

- (22) a. *Kore-wa sakura-no hana-da*  
this-TOP cherry-GEN flower-COP.PRS  
'This is a cherry flower'
- b. *Sakura-no hana-ga kirei-dat-ta*  
cherry-GEN flower-NOM beautiful-COP-PST  
'The cherry flower was beautiful' Japanese (Nambu 2023: 1)

Though often obligatory, Japanese *da* is reported to be optional in certain contexts, e.g., in embedded clauses headed by the quotative complementizer *to* following doxastic predicates like *kangaeru*, 'think' (Abe 2001, 2015, Tanomura 2008, Shioda 2010). When *da* goes unpronounced, a non-verbal predication relation is established seemingly with no additional functional material; in (23) the predicate *hitsuyoo* composes with the expletive subject in the embedded clause despite there being no overt copula.

- (23) *Sore-wa hitsuyoo(-da)-to omo-u*  
it-TOP necessary-COP.PRS-COMP think-PRS  
'(I) think that it is necessary' Japanese (Nambu 2023: 2)

Given that a predication relation is still established in the absence of an overt copula, examples like (23) are said to involve a "zero copula" that serves the same predicative function as other overt copular elements, but is phonologically null (Pustet 2003, Buell and De Dreu 2013). As it turns out, zero copulas are cross-linguistically well-attested, although they manifest in slightly different ways across languages (Pustet 2003, Gibson et al. 2019).

One way zero copulas appear is through optional dropping of an overt copula like in Japanese. Another example of this can be found in Mandarin, where the copula *shi* is optional in declarative

---

<sup>6</sup>See Nakayama (1988), Nishiyama (1999), Urushibara (1993) for claim that this is morphologically complex.

sentences (19a). A second way that zero copulas appear is through strict morphosyntactic restrictions on overt copular forms; a zero copula may be obligatory if overt copular morphology is banned in a certain environment. In Modern Hebrew for example, a zero copula is obligatory in present tense constructions featuring a nominal or adjectival predicate (24).

- (24) a. Sara mora  
       Sara teacher  
       ‘Sarah is a teacher’
- b. Hu xaxam  
       3SG.PRO clever  
       ‘He is clever’
- Modern Hebrew (Junger 1981: 122)

Once again, we find that the broader typological patterns we observe in copular systems are mirrored in Bantu languages, as zero copulas are also reported in some languages. For example, it is reported that copular clauses involving nominal or adjectival predicates in Swahili (G41) may be expressed with or without the invariant copula *ni* discussed previously.

- (25) a. Hamisi (ni) mpishi  
       Hamisi be cook  
       ‘Hamisi is a cook’
- Swahili (McWhorter 1994: 63)
- b. Shati (ni) chafu  
       5.shirt be 5AGR.dirty  
       ‘The shirt is dirty’
- Swahili (Matushansky and de Dreu 2009: 7)

The copula has been reported to be similarly optional in closely related Digo (E73; Nicolle 2013, Gibson et al. 2019). Copula dropping has also been reported in some Bantu languages that employ tonal marking strategies in place of a segmental copular element (see subsection 2.2.8).

## 2.2.4 Locative copula

Though some languages may utilize a zero copula in place of a dedicated copular verb, many other languages instead “recruit” (non-verbal) morphosyntactic material from other places in the grammar to express non-verbal predication (Stassen 1997). One such instance of this practice

involves the use of locative morphology as a copular element, which often (but not exclusively) occurs in locative predications. For example, Frajzyngier (1986) suggests that the copula *à* in Bolewa (Chadic), developed diachronically from a preposition, *\*a* ‘in, at’. The development from preposition to copula is argued to have been driven by the innovation of new locative case marker, *gá*. Frajzyngier’s proposal is that this innovation resulted in the regular adjacency of prepositional *\*a* and the case marker *gá*, at which point the function of the preposition became unclear. This in turn led to the reanalysis of prepositional *\*a* as a copular element that appears strictly in locative predications (26a). Correspondingly, the locative copula *à* is crucially unavailable in contexts where the preposition would not have been present diachronically, e.g., with an NP predicate (26b).

(26) a. Sùbá-nò            **\*(à)** kò réwè  
 shirt-POSS1SG is    on tree  
 ‘My shirt is on a tree’ Bolewa (Frajzyngier 1986: 372)

b. Yúsúfù (**\*à**) móy  
 Yusufu chief  
 ‘Yusufu is a chief’ Bolewa (Frajzyngier 1986: 375)

In other languages, the connection between locative markers and non-verbal predication is even clearer, as some employ synchronic locatives as a copula. In addition to the standard invariant copula *ni*, Swahili (G41) also permits the use of a locative copula in certain contexts, e.g., with locative predicates (27a), or predicates like *huru* ‘free’ (27b) (McWhorter 1994, Marten 2013). Unlike the locative copula in Bolewa shown above, locative copulas in Swahili are inflectional; they agree with the noun class of the subject (Gibson et al. 2019).

(27) a. Yeye **yu-ko**            Ukerewe  
 3PRO 1SM-LOC17 Ukerewe  
 ‘S/he are in Ukerewe’ Swahili (Marten 2013: 56)

b. Watu    wa-li-ji-ona            **wa-po**            huru zaidi  
 2.person 2SM-PST-REFL-see 2SM-LOC16 free more  
 ‘People saw themselves as being more free’ Swahili (Marten 2013: 62)

Looking across Bantu languages, it is quite common to find copulas that are either derived from a locative or used exclusively with locative predicates (Schneider-Zioga 2018, Gibson et al. 2019).



There are additionally many Bantu languages that employ a locative alongside a separate copular element to express possession and/or possessive predication (see discussion of comitative *na* in subsection 2.2.7).

### 2.2.5 Pronominal copula

Yet another example of existing morphosyntactic material being used to express non-verbal predication involves so-called pronominal copulas. As the name suggests, pronominal copulas are said to involve a pronominal element, e.g., a demonstrative, that comes to serve a predicational function in the absence of a dedicated copula. Consider for example the Russian pronominal copula *eto*. In addition to an overt copular verb and zero copula (Roy 2013), the demonstrative pronoun *eto* can be used — either in isolation (28a) or alongside another copular element (28b) — to establish a non-verbal predication relation (Pereltsvaig 2001, Markman 2008b).

- (28) a. Cookie **eto** tolstaja koshka  
 Cookie this fat cat.NOM  
 ‘Cookie is the fat cat’
- b. Misha **eto (byl)/∅** nash doctor  
 Misha this was/is our doctor.NOM  
 ‘Misha was/is our doctor’
- Russian (Markman 2008b: 366)

Similar pronominal copulas — or pronominal clitics as they are sometimes called — are also reported in Semitic languages like Arabic (Eid 1983, 1991, Benmamoun 2000, Choueiri 2016, Camilleri and Sadler 2019, a.o.). For example, in Lebanese Arabic a pronominal copula is one of (at least) two predicational strategies available in sentences featuring a definite NP predicate and present time reference (Choueiri 2016, Camilleri and Sadler 2019). Importantly, the pronominal copula *huwwe* in (29) is morphologically identical to third person strong pronouns (Camilleri and Sadler 2019).

- (29) Sami **∅/huwwe** mudīr l-madrāse  
 Sami COP.SG.MASC director.SG.MASC DEF-director  
 ‘Sami is the school director’
- Lebanese Arabic (Adapted from Choueiri 2016)

As with many of the other predicative strategies we have seen to this point, pronominal copulas (or something similar) have been reported in many Bantu languages as well (Schneider-Zioga 2018). Pronominal copulas in Bantu languages tend to be derived from an agreeing pronominal stem, e.g., *-o* in Nyakusa-Ngonde (M31; Meinhof 1906, Persohn 2017), and can often be found in predicational sentences that involve focus (Schneider-Zioga 2018).<sup>7</sup>

- (30) a. I-mi-piki mi-nywamu  
 AUG-4-tree 4AGR-big  
 ‘These trees are big’
- b. I-mi-piki i-gi gyo mi-nywamu  
 AUG-4-tree AUG-4DEM\_PROX 4DEM\_REF 4AGR-big  
 ‘These trees, they are big’ Nyakyusa-Ngonde (Adapted from Persohn 2017: 305)

In some Bantu languages the agreeing pronominal copula is used alongside another copular element, usually invariant *ni/ndi*, to establish a predication relation. As before, the use of a pronominal copula in these contexts tends to yield a focus or emphatic interpretation (31).

- (31) Kulungu ndi-ye wa-ngu  
 1.antelope COP-1.REF 1AGR-POSS1SG  
 ‘The antelope is (indeed) mine’ Digo (Nicolle 2013: 289)

## 2.2.6 “Inflectional” copula

Another common way of expressing non-verbal predication is to mark the non-verbal predicate with inflectional morphology as if it were a verb. In languages that employ this strategy, the predicate is marked by verbal inflectional morphology, but is still diagnostically non-verbal; such predicates generally differ from true verbal predicates in that they cannot undergo verbal derivations (Campbell 1985, Curnow 2000).

An example of an inflectional copula can be found in Pipil (or Nawat), a Nahuatl language spoken in El Salvador (Campbell 1985). Though Pipil also exhibits an overt copular verb and a zero copula, non-verbal predication is sometimes realized by marking the predicate with a verbal prefix

<sup>7</sup>See discussion of Higgins (1979), Mikkelsen (2005, 2011), Roy (2013) in section 2.3 for more on predicational sentence types and focus.

that encodes the person/number of the subject; in (32a), the non-verbal predicate takes the same subject prefix *ni* that appears with the verbal predicate in (32b).

- (32) a. **ni**-ta:kat  
 1SG.SM-man  
 ‘I am a man’ Pipil (Campbell 1985: 54)
- b. na            **ni**-k-elna:miki  
 1SG.PRO 1SG.SM-it-remember  
 ‘I remember it’ Pipil (Campbell 1985: 107)

Though perhaps reported less often cross-linguistically than other copular elements/predicational strategies, inflectional copulas are attested in other languages as well, e.g., Kalaallisut (West Greenlandic; Fortescue 1984); see Curnow (2000) for further discussion. Once again, inflectional copulas of this sort are also found in Bantu languages (Schneider-Zioga 2018, Gibson et al. 2019). Alongside its invariant copula (20), Digo (E73) also allows a subject marker in isolation to serve a predicative function, even without the presence of an overt locative clitic or *be*-verb (33a).

- (33) a. Chitabu **chi** tayari  
 7.book 7SM ready  
 ‘The book is ready’ Digo (Nicolle 2013: 289)
- b. Gafula ná-sikir-a            chitu **chi**-na-vum-a  
 suddenly 1SG.PST-hear-FV 7.thing 7SM-PRS-roar-FV  
 ‘Suddenly I heard something roar’ Digo (Adapted from Nicolle 2015: 24)

Like “inflectional” copulas in other languages, the copular strategy like that in (33a) is distinct from similar constructions involving a true verbal predicate. Unlike fully verbal predicates, “inflectional” copulas are often limited to specific morphosyntactic environments, e.g., present tense. Verbal predicates on the other hand are generally available in a much broader set of environments.

## 2.2.7 Possessive copula

It has often been reported that non-verbal predication and clausal possession are expressed in similar ways in some languages (Freeze 1992, Kayne 1993, Ritter and Rosen 1997, Harley 1998, a.o.). In

fact, we have already seen an example of a possessive *have*-verb being used to express non-verbal predication in French; with certain adjectival predicates, the *have*-verb *avoir* obligatorily serves as the copula. There is also considerable evidence of this relationship in the opposite directionality; many languages are reported to employ an overt copular element in possessive constructions in lieu of a dedicated lexical HAVE-verb. Specifically, many languages employ a copular *be*-verb in combination with a locative element to encode possession (Freeze 1992, Kayne 1993, Ritter and Rosen 1997, Harley 1998).

To illustrate, consider the similarity between (non-verbal) locative predication, existentials, and clausal possession in Russian (34)-(36) (Levinson 2011). Importantly, the possessive structure in (36) employs the same copular *be*-verb form that appears in the non-verbal predication and existential sentences before it. Unlike the others however, the possessive includes a pre-copular locative preposition, *u*, that assigns genitive case to its DP complement and marks it as the possessor (36).<sup>8</sup>

(34) *Kniga            byla            na stole*  
 book.NOM.FEM be.PST.FEM on table.LOC  
 ‘The book was on the table’ (Locative Predication)

(35) *Na stole        byla            kniga*  
 on table.LOC be.PST.FEM book.NOM.FEM  
 ‘There was a book on the table’ (Existential)

(36) *U menja        byla            sestra*  
 at 1SG.GEN be.PST.FEM sister.NOM.FEM  
 ‘I had a sister’ (Clausal Possession) Russian (Freeze 1992: 553)

Similar locative possessive strategies, i.e., constructions involving a copula and locative element, are reported in a variety of other languages including Hindi, Tagalog, and Finnish (see Freeze 1992 for further discussion and examples).

As with previous predicational strategies, locative possessives involving the use of an overt

---

<sup>8</sup>Hallman (2022) describes the Russian-type possessive pattern as a “dependent marking strategy” — the locative identifies a dependent of the possessive relation (e.g., the possessor) rather than the possession relation itself. Hallman compares this to the so-called “head marking strategy” employed in languages like Rangi (F33), where the locative element serves to relate the dependents of a possessive relationship rather than mark one of them (37)-(38).

copular *be*-verb are reported in many Bantu languages (Halpert and Diercks 2014, Myler 2016, Gibson 2012, Gibson et al. 2019). Though there is significant variability in how clausal possession is expressed across Bantu languages, the most common strategy involves the use of an overt *be*-verb alongside the comitative particle — usually *na/no* — meaning ‘and/with’ (Gibson et al. 2019). In Rangi (F33) for example, clausal possession (37) and possessive predication (38) are expressed via the combination of the overt agreeing copula *-ri* and the comitative particle *na*.

(37) **A-ri na** vijombulo vikuulu  
 1SM-be COM 8.calves 8.AGR.big  
 ‘S/he has big calves’ Rangi (Gibson 2012: 101)

(38) **Ndi-ri na** njala  
 1SG.SM-be COM 9.hunger  
 ‘I am hungry’ (Lit., ‘I have hunger.’) Rangi (Gibson 2012: 92)

Though the use of a locative possessive like that in Rangi is typologically common across Bantu languages, it is not the only possessive marking strategy that is observed (Gibson et al. 2019). One particularly relevant strategy to our discussion of non-verbal predication is the use of an overt *have*-verb in languages like Mashi (JD53). Unlike other so-called HAVE languages (Levinson 2011), Mashi (among other JD50 languages) exceptionally employs two distinct lexical *have*-verbs, *-dwiire* and *-jira*, to express possession and/or possessive predication (39).

(39) Johne a-<sup>✓</sup>**dwiire**/<sup>✓</sup>**jira** iwazo linja  
 John 1SM-have-DWIIRE/have-JIRA 5.idea 5.good  
 ‘John has a good idea’ Mashi (Finholt 2024)

As discussed in Finholt (2022, 2024), the distribution and interpretation of the two Mashi *have*-verbs is found to be quite similar to the two *be*-verbs in the language, *-li* and *-ba*. On the basis of this overlap, Finholt (2022, 2024) argues that the *be*- and *have*-verbs in Mashi are derivationally related. That said, the important point here is that languages differ in how they express clausal possession, and how closely this resembles the expression of non-verbal predication.

### 2.2.8 Other strategies

In addition to the overt segmental marking strategies outlined above, there are many other ways of marking a non-verbal predication relation. Given the focus of this dissertation, I will limit my discussion to two additional non-verbal predication strategies that are commonly reported in Bantu languages, namely tonal marking and marking via the augment.

Another common predication strategy reported in Bantu languages involves the use of a distinct tone pattern to mark a constituent (usually a noun) as a predicate (Welmers 1973, Pustet 2003, Schneider-Zioga 2018, Gibson et al. 2019). Consider the tonal alternation between the two nominal forms in Shona (S10) in (40). As illustrated in (40a), nominal prefixes in Shona carry an inherent low tone; a bare NP in isolation will always surface with a low tone class marker, e.g., class 1 *mù-*. In some contexts however, a nominal prefix may surface with a high tone, in which case it yields a predicational reading of the noun (40b).

- (40) a. *Mù-nhù*  
1-person  
'person'
- b. *Mú-nhù*  
1-person.PRD  
'It is a person'
- Shona (Welmers 1973: 323)

Given their morphological identity, the interpretive difference between the two examples above suggests that tone alone — namely a left-edge high tone — may serve to mark a constituent as a predicate in Shona (Welmers 1973, Matushansky and de Dreu 2009, Gibson et al. 2019). It is sometimes reported that tonal marking strategies of this kind exist in complementary distribution with overt copular elements. In Lamba (M54), for example, tonal marking of the predicate — with a low tone on the left edge — is limited to constructions involving a third person subject, and co-occurs with a zero copula; With first and second persons, the copular verb *-li* is used (Doke 1922, Schneider-Zioga 2018).

- (41) a. **U-li**            mu-ntu  
           2SG.SM-be 1-person  
           ‘You are a person’
- b. **Mù**-ntu  
           1-person (lowered tone)  
           ‘S/he is a person’

Lamba (Doke 1922: 94)

Schneider-Zioga (2018) reports that similar tone marking strategies have been attested in many other (particularly southern) Bantu languages, including Umbundu (R11; Schadeberg 1986), Zulu (S42; Matushansky and de Dreu 2009), and Mahkuwa (P31; Van der Wal 2006).

Another method of marking nominal predicates in particular among Bantu languages is through the deletion of the augment.<sup>9</sup> Much like tonal patterns may be used to distinguish a bare NP from a predicative NP, full deletion of the augment has been observed as a predicate marking strategy in languages like Lusoga (JE16; Van der Wal 2006: 231); the augmentless nominal in (42b) is interpreted as a predicate and receives an identificational reading.

- (42) a. o-mú-géni  
           AUG-1-guest  
           ‘guest’
- b. mú-géni  
           1-guest  
           ‘It is a guest’

Lusoga (Van der Wal 2006: 231)

Though this marking strategy is limited to a subset of Bantu languages that feature the augment synchronically, predicate augment-deletion is quite prevalent among Eastern Bantu languages, particularly those of the JD group (Schneider-Zioga 2018). As Schneider-Zioga (2018) reports in

---

<sup>9</sup>The augment — or the pre-prefix — is the leftmost element in the nominal domain of certain Bantu languages. In morphosyntactic contexts where it is licit, the augment generally corresponds to a vowel that precedes the noun class prefix (Gambarage 2019).

- (1) a. o-mú-nwa  
           AUG=3-mouth  
           ‘a/the mouth’
- b. *Nominal domain template:*  
           AUGMENT-NOUN CLASS-STEM

Kinande (JD42), predicate nominals generally surface without an augment in copular sentences in JD languages despite surfacing with an augment in other contexts.

- (43) a. *ibó ni bagalímu*  
2PRO COP 2.teacher  
'They are teachers'
- b. \**ibó ni a-bagalímu*  
2PRO COP AUG-2.teacher  
'They are teachers'

Kinande (Schneider-Zioga 2018: 10)

As we will see in later chapters (see chapter 5 in particular), many other languages of the JD group exhibit similar augment deletion patterns in non-verbal predication.

### 2.3 Complex copular systems

As we have seen, there are a variety of different strategies that languages may use to realize a non-verbal predication relation. To this point I have shown that these strategies can vary radically across languages; though one language may employ an overt copular verb, another may employ an invariant copular particle, or even a zero copula. Although this sort of copular variation has been framed as a comparative issue thus far, that is not to say that a single language may not itself employ multiple distinct predicational strategies. In fact, there are many examples of languages that feature multiple distinct copular forms. For example, we have seen that Bolewa features both a locative copula and a zero copula. Likewise, Mandarin and Modern Greek feature an overt copular verb and a zero copula. Among the Bantu languages surveyed thus far we have seen similar variation, with Digo featuring an invariant copula, pronominal copula and inflectional copula, and Swahili featuring an invariant copula, locative copula and a zero copula.

As such, in order to capture a more accurate typological picture of non-verbal predication, our survey needs to address both crosslinguistic and intralinguistic variation. One way to address these two levels of variation is to categorize languages according to the complexity of their copular systems. What number and combination of copular forms do languages exhibit? In what morphosyntactic or interpretive contexts do languages use different copular forms?



At one extreme, there are languages (like English) that exhibit a single copular form used in all (monoclausal) non-verbal predications. However, as we have seen, not all languages are limited to a single copular form in non-verbal predication. While languages like Mandarin and Modern Greek may employ different multiple different types of copulas, e.g., an overt copular verb and zero copula, some languages exhibit two distinct forms of the same copular strategy, e.g., two distinct copular *be*-verbs. Moreover, languages may employ their respective copular forms in different ways; different forms of the copula may be associated with distinct interpretations or morphosyntactic environments.

The roadmap for the following subsections will be as follows. I will begin with a discussion of interpretive effects and attested contrasts in copular systems in section 2.3.1. In section 2.3.2 I will discuss different copular clause types and their effect on copular choice cross-linguistically. Finally, I will address a range of other morphosyntactic factors, e.g., category of the predicate, tense-aspect-mood environment, subject person, etc., that have been shown to influence the form of the copula in section 2.3.3.

### **2.3.1 Interpretive contrasts**

With the observation that languages may employ multiple copulas, the question naturally arises as to when each copular form is used. This subsection gets at this question by addressing differences in their interpretation: distinct copular forms may coincide with distinct interpretations of the predication relation.

From a typological perspective, it is relatively common to find languages that make semantic distinctions using different forms of the copula — particularly different copular *be*-verbs — in non-verbal predication. There is a substantial literature suggesting that copulas may reflect or encode the semantic properties of the predication relation they help realize (Ramchand 1997, Greenberg 1998, Green 2000, Bochnak et al. 2011, Deo et al. 2017, a.o.). For example, many languages are reported to distinguish between stage- and individual-level predication using different copular *be*-verbs. Consider the Spanish copular *be*-verbs *estar* and *ser*, which are widely reported to exhibit

distributional and interpretational differences akin to the stage/individual contrast (Milsark 1974, Carlson 1977, Kratzer 1995, Diesing 1990, 1992b, Arche 2006 a.o.), with *estar* being preferred with properties that hold of stages, i.e., transient, temporary states (44a), and *ser* being preferred with properties that hold of individuals, i.e., permanent, long-term states (44b).<sup>10</sup>

- (44) a. El artista  $\checkmark$ está/#es presente/ausente/lejos  
 The artist estar/ser.PRS.3SG present/absent/far away  
 ‘The artist is present/absent/far away’ Spanish (Maienborn 2005: 173)
- b. Maria #está/ $\checkmark$ es inteligente/altruista  
 Maria estar/ser.PRS.3SG intelligent/altruistic  
 ‘Maria is intelligent/altruistic’ Spanish (Deo et al. 2017: 8)

In other languages however, such interpretive contrasts may take on different flavors. Take for example the copular system in Modern Irish. As we have already seen, Modern Irish exhibits an invariant copula *is* in copular sentences (18). Like Spanish, Modern Irish also exhibits a second invariant copula *bí* that is interpretively distinct from *is* (Stenson 1981, Roy 2013), however the interpretive contrast between the two is not identical to the traditional stage/individual contrast in the sense of (Kratzer 1995). Unlike previous accounts of stage/individual contrast in other languages, both *is* and *bí* are reported to be compatible with individual-level predicates, suggesting that any interpretive difference between them is likely unrelated to the stage/individual contrast (Carnie 1995, Roy 2013). As for their interpretive effects, Stenson (1981) suggests that the difference between the two copulas is that *is* describes a “defining” characteristic of the subject, whereas *bí* simply describes a state. As such, *is* is used to describe priesthood as a core characteristic of the subject in (45a), while *bí* instead describes being a priest as something the subject does for work (see Stenson (1981) and Roy (2013) for further discussion).

- (45) a. **Is** sagart é mo dheartháir  
 be<sub>IS</sub>.PRS priest him 1SG.POSS brother  
 ‘My brother is a priest’ (he is a member of the set of priests)
- b. **Tá** mo dheartháir ina sagart  
 be<sub>BÍ</sub>.PRS 1SG.POSS brother in.AGR priest  
 ‘My brother is a priest’ (his chosen career) Irish (Stenson 1981: 94)

<sup>10</sup>See Carlson (1977) and Kratzer (1995) for discussion of these terms.

We find a similar situation in the Eastern Bantu language Mashi (JD53), which also exhibits a binary distinction between its two *be*-verbs *-li/-ba*. While both available with a predicate like *-bulee* ‘blue’ (46), the two *be*-verbs yield distinct interpretations of the predication relation. Though similar to a stage-individual contrast (e.g., like Spanish *estar/ser* above), I argue in previous work that Mashi *-li/-ba* actually distinguish between “temporary” and “permanent” predication (Finholt 2022); *-li* yields a temporary, stative interpretation, while *-ba* yields a permanent interpretation (see chapter 4 for further discussion).

- (46) a. O-ma-lunga ga-**li** ga-bulee  
 AUG-6-sky 6SM-be<sub>TEMP</sub> 6AGR-blue  
 ‘The sky is blue’ (talking about the weather; outside right now)
- b. O-ma-lunga ga-**ba** ga-bulee  
 AUG-6-sky 6SM-be<sub>PERM</sub> 6AGR-blue  
 ‘The sky is blue’ (a property of the sky; a fact about the world) Mashi (Finholt 2022)

Though the above data suggest that Spanish, Irish, and Mashi encode subtly distinct interpretive contrasts in their copular *be*-verbs, in reality these contrasts are quite difficult to neatly categorize. For example, while generally discussed as the canonical example of the stage/individual contrast, Spanish *estar/ser* are reported to exhibit distributional and interpretive properties that are much more in line with a temporary/permanent contrast like that in Mashi (Deo et al. 2017). Specifically, it is shown that both *estar* and *ser* may be used with stage/individual-level predicates depending on the context; with the adjective *ajustada*, ‘tight’, *estar* yields a context specific reading (47a), whereas *ser* yields an “inherent” reading (47b).

- (47) a. La chaqueta le **está** ajustada en los hombros  
 The jacket CL-IO estar.PRS.3SG tight.FEM on the shoulder.PL  
 ‘The jacket is tight on the shoulders’ (it doesn’t fit me)
- b. La chaqueta **es** ajustada y muy moderna  
 The jacket ser.PRS.3SG tight.FEM and very fashionable.FEM  
 ‘The jacket is tight and very fashionable’ (it is tight by design) Spanish (Deo et al. 2017)

Given this distributional flexibility, Deo et al. (2017) suggest that *estar/ser* do not truly reflect a stage/individual contrast, but rather a distinction between a “contextually-bounded” reading (*estar*)

and a general, “unbounded” reading (*ser*). Other accounts of the contrast broadly arrive at the same conclusion; *estar* seems to encode a situationally-dependent interpretation, whereas *ser* describes a more general, permanent property (Maienborn 2005, Roy 2013).

In addition to expressing stage/individual-level predication, there are many other interpretative effects reported in copular systems, particularly in Bantu languages. Though not all Bantu languages employ multiple distinct *be*-verbs, most exhibit at least one copula derived from the proto-Bantu copulas *\*-de* ‘be’ and *\*-bà* ‘become, be’ (Meeussen 1967, Wald 1973). We have already seen one example of a language that employs reflexes of both; Mashi (JD53) encodes a temporary/permanent contrast in its two copular *be*-verbs *-li* (*\*-de*) and *-ba* (*\*-bà*). Though there is significant variation across Bantu languages (see chapter 4), reflexes of these copulas often exhibit similar interpretive and distributional restrictions. For example, cognates of *-li* (*\*-de*) are often reported to encode stative or locative interpretations (Schneider-Zioga 2018, Gibson et al. 2019). In Kinyarwanda (JD61) for example, the copular verb *-ri* (*\*-de*) is reported to strictly coincide with locative descriptions, and is therefore unavailable in examples like (48a) (Jerro 2015).

- (48) a. Kyle **ni/\*a-ri** mu-nini  
 Kyle be<sub>NI</sub>/1SM-be-<sub>RI</sub> 1AGR-big  
 ‘Kyle is big’
- b. Umw-aana **a-ri** i Kigali  
 1-child 1SM-be-<sub>RI</sub> in Kigali  
 ‘The child is in Kigali’ Kinyarwanda (JD61) (Adapted from Jerro 2015)

Likewise, reflexes of proto-Bantu *\*-bà* are often reported to coincide with inchoative interpretations (among others). In Zulu (S42), the use of *-ba* yields an inchoative, “becoming” interpretation rather than a predicative interpretation (Posthumus 2000, Schneider-Zioga 2018).

- (49) a. U muhle  
 1SM 1AGR.beautiful  
 ‘You are beautiful’
- b. U-ba gotho  
 1SM-be-<sub>BA</sub> honest  
 ‘You are becoming honest’ Zulu (S42) (Adapted from Posthumus 2000)

While much of the discussion on copular contrasts has focused on binary interpretive distinctions, e.g., stage vs individual, temporary vs permanent etc., others have suggested more fine grained distinctions. Based on data from French, Irish, and Russian, Roy (2013) posits three syntactically and interpretively distinct types of non-verbal predicate (see subsection 3.3.3 for more concrete discussion).

(50) *Three predicate types presented in Roy (2013)*

- a. *Dense* (situation-descriptive) predicates range over mass (non-atomic) eventualities, yield “temporary” state interpretations, and are syntactically associated with a bare predicate XP.
- b. *Non-dense* (characterizing) predicates range over atomic eventualities, allow for interpretive “gaps”, and are associated with an additional projection above XP, the classifier phrase (CIP).
- c. *Maximal* (defining) predicates range over maximal eventualities, yield “permanent” state interpretations, and are uniquely associated with the presence of a number phrase (NumP) above CIP and XP.

Importantly, Roy (2013) suggests that languages differ in how they morphosyntactically express these three predicate types. In languages with only one copular *be*-verb, this three-way contrast is directly expressed by the predicate. In French, which generally employs the *be*-verb *être* in non-verbal predication, the distinction between *maximal* and *non-dense* nominal predicates is encoded via the presence/absence of the indefinite article *un(e)*, which Roy (2013) argues realizes the head of NumP.<sup>11</sup> When a nominal predicate like *acteur* ‘actor’ occurs with an indefinite article, it serves to identify or “define” an individual (51a), whereas the bare predicate simply attributes a property to an individual (52b). As such, we see a stark contrast in the possible answers to *who?* and *what?* questions in French. When answering a *who?* question, only a maximal predicate is available, since this is used to describe a defining property of an individual. With a *what?* question

---

<sup>11</sup>Note that French does have other predicational strategies — possessive predication with *avoir* ‘have’ (section 2.2.7) and *comme* ‘as’ in small clauses (den Dikken 2006).

however, only a non-dense predicate is used, as this merely describes a property that holds of an individual, but doesn't necessarily define them (e.g., a job).

(51) Qui est Raymond?  
 who is Raymond  
 'Who is Raymond?'

a. *Maximal predicate*

Raymond est un acteur  
 Raymond is an actor  
 'Raymond is an actor'

b. *Non-dense predicate*

#Raymond est acteur  
 Raymond is actor  
 'Raymond is an actor'

(52) Qu'est Raymond?  
 what.is Raymond  
 'What does/is Raymond?'

a. *Maximal predicate*

#Raymond est un acteur  
 Raymond is an actor  
 'Raymond is an actor'

b. *Non-dense predicate*

Raymond est acteur  
 Raymond is actor  
 'Raymond is an actor'

The third predicate type, *dense* predicates, are expressed just like *non-dense* predicates in French; both involve a bare predicate XP after the copula. That said, the two predicate types again differ in terms of the interpretation. *Dense* predicates yield a situation-descriptive interpretation, and therefore describe temporary states, e.g., *ivre* 'drunk' in (53a). *Non-dense* predicates on the other hand yield characterizing interpretations; they describe general properties that apply over longer periods of time, e.g., *ivrogne*, 'drunkard' in (53a). As such, only the former predicate type (*dense*) may be used to answer a question about a specific situation (53).

(53) Qu'est ce qui s'est passé dehors, c'est quoi tout de bruit?  
 what.is it that REFL.is happening outside it.is what all of noise  
 'What is going on outside, what's all of that noise?'

a. *Dense predicate*

Paul est ivre, il a renversé la poubelle  
 Paul is drunk he has knocked.over the trash  
 'Paul is a drunk, he knocked over the trash'

b. *Non-dense predicate*

#Paul est ivrogne, il a renversé la poubelle  
 Paul is drunkard he has knocked.over the trash

‘Paul is a drunkard, he knocked over the trash’

As for languages with more than one copula, e.g., Spanish, Roy (2013) suggests that the three-way *dense/non-dense/maximal* contrast is expressed through both copular and predicate contrasts. Whereas French utilizes the same copula with all three predicates, Spanish utilizes *estar* with dense predicates, and *ser* with non-dense and maximal predicates (see section 3.3.3 for further discussion).

	<b>Dense</b> (Situation-Descriptive)	<b>Non-dense</b> (Characterizing)	<b>Maximal</b> (Defining)
French	<i>être</i> + XP	<i>être</i> + CIP ( $\emptyset$ )	<i>être</i> + NumP ( <i>un</i> )
Spanish	<i>estar</i> + XP	<i>ser</i> + CIP ( $\emptyset$ )	<i>ser</i> + NumP ( <i>uno</i> )

Table 2.1: Comparing French and Spanish under Roy’s (2013) Analysis

While not overtly discussed in her analysis, Roy’s three-way predicate dichotomy leaves open the possibility that a language could encode this contrast using three distinct copulas. While there are many languages that employ three distinct strategies, three-way contrasts between copulas of the same type, e.g., *be*-verbs, are relatively uncommon. That said, there are at least two Bantu languages I am aware of that employ three distinct copular verbs, those being Lamba (M54; Doke 1922), and Kifuliiru (JD63; Finholt 2023). In both languages the three copular verbs are interpretively distinct, albeit in slightly different ways. In Lamba, *-li* (\*-de) is stative, *-wà* (\*-bà) is inchoative, and *-ikala* means ‘remain/stay’ (Doke 1922, Schneider-Zioga 2018). As I will discuss in chapter 4, I characterize the three copular verbs in Kifuliiru as encoding similar meanings to the three predicate types presented in (Roy 2013): *-li* is situation-descriptive, *-(mu)ba* is characterizing, and *-tula* is defining.

### 2.3.2 Types of copular clauses

Another way to think about copular contrasts relates to the four-way classification of copular clauses established in Higgins (1979). Broadly, work on this topic has focused on the syntactic relationship between the “subject” and the “predicate”, and how they differ interpretively. Central to this work is the observation that nominal predicates like English *doctor* may sometimes precede the copula

(54)-(55).

(54) Susan is the doctor

(55) The doctor is Susan

(Adapted from Mikkelsen 2005)

Building on this observation, Higgins (1979) provided a four-way classification of copular clauses based on differences in interpretation and information structure. The first category, (“pure”) *predicational* clauses, describe the standard case of non-verbal predication we have seen thus far; the post-copular predicate ascribes a property to the pre-copular element, i.e. the subject (56a). In so-called *specificational* clauses however, the predicate is expressed before the copula rather than after. Interpretively, the pre-copular predicate denotes a property that defines a set of individuals, and then the post-copular NP picks out a single member of that set (56b). For Stassen (1997), this prototypically involves a label (here, *the winner*) that is presumed to be unfamiliar to the addressee; the job of the speaker is to provide an “exhaustive” definition of that label. In a similar vein, *identificational* clauses involve a post-copular subject that serves to identify the reference of a pre-copular deictic element, e.g., *that woman* in (56c). Finally, *equational* clauses serve to equate the reference of two referential NPs (56d). For Stassen (1997), this equates to the collapse of two once-distinct categories. That is, the example in (56d) is compatible with a context in which the addressee has “filed” the NPs as separate categories, [*John’s mother*] and [*Susan*], that refer to distinct individuals. The speaker’s goal is simply to elicit a collapse of those two categories.

(56) *Types of copular clauses*

- a. Predicational: *Susan is a doctor*
- b. Specificational: *The winner is Susan*
- c. Identificational: *That woman is Susan*
- d. Equational: *John’s mother is Susan*

(Adapted from Mikkelsen 2005)

Though Higgins’ taxonomy of copular clauses focused primarily on English contrasts, the existence of these four predication types has found considerable support in other languages as well.



In fact, unlike English, many languages overtly distinguish these different types of predication using distinct morphosyntactic strategies. For example, the form of the copula in Thai is reported to be sensitive to predication type, among other morphosyntactic factors (Warotamasikkhadit 1972, Wongwattana 2015). In general, the copula *pen* is obligatory in “pure” predicational clauses in Thai (57), unless the predicate is a locative/temporal PP, in which case the copula *jù:* is employed instead (58). In all other copular clauses however, *pen* (and *jù:*) are unavailable; *k<sup>h</sup>i:* is used in specificational, identificational, and equational clauses (59)-(61).

(57) *Predication*

K<sup>h</sup>ǎw **pen**/\***jù:** dèk  
 s/he COP child  
 ‘S/he is a child’

(58) *Predication (PP<sub>LOC/TEMP</sub>)*

Mɔ:tə:saj **jù:/\*pen** naj bâ:n  
 motorcycle COP in house  
 ‘The motorcycle is in the house’

(59) *Specification*

P<sup>h</sup>û: tɕ<sup>h</sup>á<sup>?</sup>ná **k<sup>h</sup>i/\*pen** buak<sup>h</sup>ǎ:w po: pràmúk  
 person win COP Buakaw Por Pramuk  
 ‘The winner is Buakaw Por Pramuk’

(60) *Identification*

Nî: **k<sup>h</sup>i/\*pen** ma:rí<sup>??</sup>ô  
 this COP Mario  
 ‘This is Mario’

(61) *Equation*

Sà<sup>?</sup>jǎ:m **k<sup>h</sup>i/\*pen** prà<sup>?</sup>t<sup>h</sup>ê:t t<sup>h</sup>aj  
 Siam COP country Thai  
 ‘Siam is Thailand’

Thai (Wongwattana 2015)

Similar distinctions have been observed in many Niger-Congo languages, including Mande languages like Bambara (Vydrin 2020, Sangare and Roy 2022), and a number of Eastern Bantu languages (Schneider-Zioga 2018). For example, the Bantu language Kinande (JD42) is reported to utilize distinct strategies for expressing pure predication and specification/equation/identification (Schneider-Zioga 2018). While an invariant copula (*ni*) is used in pure predicational clauses, a pronominal copula (*-o*) is used in all other types of predicational clauses (62)-(65).

(62) *Predication*

Kambale **ni** mwimbi  
 Kambale be<sub>NI</sub> 1.thief  
 ‘Kambale is a thief’

(63) *Specification*

ómwira w-age **k’** ákákekulú k’ omo kisomó  
 1.friend 1AGR-1SG.POSS 12.FOC 12.old woman 12LNK 18LOC 7.church  
 ky-etu  
 7AGR-1PL.POSS  
 ‘My (best) friend is a little old lady from our church’

(64) *Identification*

omúlumy’ óliá **yó** omukolo owe’ departement  
 1.man 1DEM 1.FOC 1.head 1LNK 1.department  
 ‘That man is the head of the department’ (pointing)

(65) *Equation*

Eririma **ky’** ekihugo (Matthew 13:38)  
 5.field 7.FOC 7.world  
 ‘The field is the world’

Kinande (Schneider-Zioga 2018)

The pattern observed in Kinande is ultimately one that is observed in a variety of Bantu languages, particularly among those in the JD40/JD50 groups (Schneider-Zioga 2018). I will return to this pattern in chapter 5. What is important here is simply that not all types of non-verbal predication are treated the same; the relationship between the subject and predicate may play a role in how the copula is realized. Thus we find differences between languages in how they morphosyntactically encode these distinct relationships.

### 2.3.3 Morphosyntactic considerations

A common feature of complex copular systems is that their component copulas have distinct distributions. As I will show in this subsection, copulas are sometimes limited to certain morphosyntactic environments, and are often sensitive to the grammatical category of the predicate, tense, aspect, mood, polarity, and person.

Perhaps the most straightforward distributional distinction observed in copular systems is a sensitivity to the category of the predicate. In Thai for example, only the copula *jù:* appears with PP predicates; the copula *pen* is used in all pure predicational clauses except those involving a locative/temporal PP predicate, which require *jù:* (58). A similar distinction is made in the copular system of Bambara (Niger-Congo), which features four distinct copulas (Vydrin 2020, Sangare and Roy 2022). One copula, *ká*, is broadly used in all pure predicational clauses (66a). However, like Thai, a dedicated copula is used with locative or adverbial predicates (66b).

- (66) a. Búuru **ká** kálan  
bread be warm  
'The bread is warm' Bambara (Sangare and Roy 2022: 2)
- b. N **bé** yàn  
1SG.PRO be here  
'I am here' Bambara (Vydrin 2020: 81)

Similar observations are reported in Bantu languages as well, particularly with locative predicates. We have already seen that reflexes of the proto-Bantu copula *\*-de* are often limited to locative or temporal descriptions (Schneider-Zioga 2018), e.g., Kinyarwanda *-ri* being limited to morphosyntactically locative predicates in the present tense (67) (Jerro 2015).

- (67) a. Mukamana a-**ri** \*mu-nini/\*umw-alimu/√mu-rugo  
Mukamana 3SM-be<sub>RI</sub> 1-big/1-teacher/18LOC-house  
'Mukamana is big/a teacher/at home'
- b. Kyle **ni** √mu-nini/√umw-alimu/\*mu-rugo  
Kyle be<sub>NI</sub> 1-big/1-teacher/18LOC-house  
'Kyle is big/a teacher/at home' Kinyarwanda (Jerro 2015: 93-96)

Another common partition observed in copular systems is tense-based: distinct copular forms appear in distinct tense environments. Consider the alternation between the Russian copular verb *byt'* and the zero copula (Seres and Espinal 2019). In the pure predicational clauses in the present tense, only the zero copula (or pronominal copula) is available (68a). However, in the past tense, only the inflected form of *byt'* is available (68b).

- (68) a. Segodnja reka spokojna  
 today river.FEM calm.FEM.SF  
 'Today the river is calm'
- b. Ivan byl goloden  
 Ivan be.PST.MASC hungry.MASC.SF  
 'Ivan was hungry' Russian (Adapted from Roy 2013: 119)

Similar tense systems have been reported in a range of other languages, including Bantu languages. In many Eastern Bantu languages the distribution of the invariant copular particle *ni* is sensitive to tense, among other factors. Consider the distribution of *ni* in Kinande (JD42). Though *ni* cannot bear tense, it is available in pure predicational clauses in the present where tense is not segmentally marked (69a). In the past tense however, *ni* is unavailable regardless of whether it is the sole copular element (69b) or accompanied by a copular verb that can bear tense (69c).

- (69) a. Magulú ní mw-íbi  
 Magulu be<sub>NI</sub> 1-thief  
 'Magulu is a thief'
- b. Magulú á-<sup>✓</sup>byá/\*ní mw-íbi  
 Magulu 1SM-be<sub>-ba</sub>.PST/be<sub>NI</sub> 1-thief  
 'Magulu was a thief'
- c. \*Magulú á-byá ní mw-íbi  
 Magulu 1SM-be<sub>-ba</sub>.PST be<sub>NI</sub> 1-thief  
 'Magulu was a thief' Kinande (Adapted from Schneider-Zioga 2018: 20)

That being said, the distribution of *ni* changes slightly in other copular clause types. For example, while *ni* is shown to be unavailable in past tense pure predicational clauses (69b-c), it is available in past tense specificational clauses so long as it is accompanied by an inflectional copular verb,

and the post copular referent is a proper name (70c).<sup>12</sup> This example illustrates an important point that I will return to throughout this dissertation: the form of the copula is often sensitive to multiple morphosyntactic factors.

- (70) a. O-mw-íbi    **ni**    Magulú  
           AUG-1-thief be<sub>NI</sub> Magulu  
           ‘The thief is Magulu’
- b. O-mw-íbi    á-**\*byá/\*ní**            Magulú  
           AUG-1-thief 1SM-be<sub>ba</sub>.PST/be<sub>NI</sub> Magulu  
           ‘The thief is Magulu’
- c. O-mw-íbi    á-**byá**            í-**ni**            Magulú  
           AUG-1-thief 1SM-be<sub>ba</sub>.PST 1AGR-be<sub>NI</sub> Magulu  
           ‘The thief is Magulu’                            Kinande (Adapted from Schneider-Zioga 2018: 20)

Looking generally across Bantu languages, we find a number of broad correlations between copular forms and the tense/aspect environments they appear in (Schneider-Zioga 2018). For example, derivatives of proto-Bantu *\*-de* are often reported to be the sole copula employed in past tense or perfective environments, e.g., *-li* in Lamba (M54; Doke 1922) and Chichewa (N31b; Kiso 2012), and *-ri* in Kinyarwanda (JD61; Jerro 2015). This correlation is strong enough that the reflex of *\*-de* in Swahili (G41) has been reanalyzed as the past tense marker *-li*, which explains the absence of a *-li* copula in the language (Jerro 2015). Derivatives of proto-Bantu *\*-bà* exhibit a much broader distribution, however they are still often limited to certain tense/aspect/mood environments, e.g., *-wà* limited to past and future tense in Xhosa (S41; Matushansky and de Dreu 2009), *-ba* limited to future tense/irrealis environments in Kinyarwanda (JD61; Jerro 2015). Additionally, some Bantu languages employ dedicated copular verbs for the future tense, e.g., Chichewa (N31b; Kiso 2012) *-khala* ‘become, be, stay’ (from *\*-kad* ‘sit, dwell, stay’).

Yet another factor that has been observed to play a role in the realization of the copula is polarity. For example, in addition to its standard copular verb *bul-*, Bashkir (Turkic) is reported to employ at least one dedicated negative copula, *tügel* (Baranovaa and Mishchenkob 2021). This negative

<sup>12</sup>Why *ni* appears in this environment instead of the focus marker *-o* is unclear (Schneider-Zioga 2018). That said, the nature of the post copular NP clearly plays an important role, as *ni* only appears with proper names; *-o* appears with all other nominals.

copula appears in all negated non-verbal predication relations involving a “referential predicate” (71).

- (71) Min jaǰ-ə-w-sə tügel/tügel-men  
 1SG.PRO write-NMLZ-AG NEG.COP/NEG.COP-1SG  
 ‘I am not a writer’ Bashkir (Baranovaa and Mishchenkob 2021: 412)

Similar “negative copulas” are reported in a variety of other languages, including Arabic (Choueiri 2016), Movima (Haude 2018), and a significant number of Bantu languages (Schneider-Zioga 2018, Gibson et al. 2019, a.o.). In general, dedicated negative copulas in Bantu languages tend to be invariant copular particles in the sense that they do not exhibit subject concord, much like the invariant copula *ni/ndi* (Gibson et al. 2019). These invariant negative copulas often take a morphological shape akin to the Swahili (G41) negation marker *si*, e.g., *sí* in Rangi (F33; Gibson 2012), *si* in Digo (E73; Nicolle 2013).

- (72) Mino **si** dza a-tu a-njina  
 1SG.PRO NEG.COP like 2-people 2-other  
 ‘I am not like other people’ Digo (Nicolle 2013: 287)

In contrast, a select few Bantu languages employ a dedicated inflectional negative copula. A borderline case of this can be found in Zulu (S42), which sometimes employs a negative copular form *si* — like the invariant form discussed above — that co-occurs with subject marking morphology and a pre-verbal negation marker, e.g., *ku-* (Poulos and Msimang 1998, Buell and De Dreu 2013).

- (73) Ka-ku-**si**-muntu lona, yi-silwane  
 NEG-17SM-NEG.COP-1person COP 7-animal  
 ‘This isn’t a person; it’s an animal’ Zulu (de Dreu 2008: 438)

A much clearer example of this can be seen in Mongo (C61), which exhibits a negative copula *-fa* that takes subject noun class agreement morphology (Hulstaert 1965, Gibson et al. 2019).

- (74) Tó-**fa** ba-laki  
 1PL-NEG.COP 2-teacher  
 ‘We are not teachers’ Mongo (Hulstaert 1965: 340)

Though not all Bantu languages employ a dedicated negative copula, negation often still plays a role in determining the form of the copula. For example, negative copular sentences in Lamba (M54) feature the copular verb *-li*, while their affirmative counterparts need not (Doke 1922, Schneider-Zioga 2018). In (75a), the predication relation is marked by a low tone on the left edge of the predicate (a process referred to as “predicate lowering”; see Van der Wal 2006), however its negative counterpart in (75b) exhibits an overt copular verb that co-occurs with a preceding negation marker.

(75) a. **mù**-ntu  
 1-person (lowered tone)  
 ‘S/he is a person’

b. ta-**li**            mu-ntu  
 NEG-be-<sub>LI</sub> 1-person  
 ‘S/he is a person’

Lamba (Adpated from Doke 1922: 94-95)

Though there are other morphosyntactic factors that may play a role in determining the form of the copula cross-linguistically, I will conclude this subsection with a brief discussion of one factor that is of particular relevance to Bantu languages, namely the person features of the subject. It has been observed that many Bantu languages distinguish third person subjects from first and second person subjects in non-verbal predication (Doke 1922, Lanham 1953, Schneider-Zioga 2018). Consider once again the copular system of Lamba (M54). As previously mentioned, Lamba employs both an overt copular verb *-li* and a tonal marking strategy in affirmative copular sentences, however, the “choice” between one strategy and the other is crucially tied to the person of the subject; *-li* is used with a second (or first) person subject (41a), but tonal marking of the predicate occurs with a third person subject (41b).

(41) a. u-**li**            muntu  
 2SG.SM-be-<sub>LI</sub> 1.person  
 ‘You are a person’

b. **mù**-ntu  
 1.person (lowered tone)  
 ‘S/he is a person’

Lamba (Doke 1922: 94)

Similar contrasts between first/second person and third person subjects have been reported in

many Eastern Bantu languages. For example, in both Kinande (JD42; Schneider-Zioga and Mutaka 2015a) and Kinyarwanda (JD61; Jerro 2015), the contrast between the invariant copula *ni* and the verbal copula *-li/-ri* is neutralized with first and second person subjects. As illustrated in the Kinyarwanda data below, *ni* is limited to sentences involving a third person subject and a nominal or adjectival predicate (76d). In contrast, *-ri* is used for all predications involving a first or second person subject, as well as with locative descriptions of third person subjects (76a-c).

- (76) a. N-**di**            ✓*mu-nini*/✓*umu-zungu*/✓*mu-rugo*  
 1SG.SM-be<sub>RI</sub> 1-big/1-foreigner/18LOC-house  
 ‘I am big/a foreigner/at home’
- b. U-**ri**            ✓*mu-nini*/✓*umu-zungu*/✓*mu-rugo*  
 2SG.SM-be<sub>RI</sub> 1-big/1-foreigner/18LOC-house  
 ‘I am big/a foreigner/at home’
- c. Mukamana a-**ri**        \**mu-nini*/\**umw-alimu*/✓*mu-rugo*  
 Mukamana 3SM-be<sub>RI</sub> 1-big/1-teacher/18LOC-house  
 ‘Mukamana is big/a teacher/at home’
- d. Kyle **ni**        ✓*mu-nini*/✓*umw-alimu*/\**mu-rugo*  
 Kyle be<sub>NI</sub> 1-big/1-teacher/18LOC-house  
 ‘Kyle is big/a teacher/at home’
- Kinyarwanda (Jerro 2015: 93-96)

## 2.4 Use of BE in other domains

Given the topic of this dissertation, the discussion in this chapter has focused solely on copular variation as it pertains to the expression of non-verbal predication. That being said, “copular variation” can be broadly extended to patterns of variation involving copular elements used outside of non-verbal predication as well.

In addition to their predicative function as copulas, copular elements — including *be*-verbs, *have*-verbs, invariant particles, etc. — are often grammaticalized as auxiliaries across the world’s languages (Meillet 1912, Dik 1987, a.o.). As in non-verbal predication, copulas — or rather *be*/*have*-auxiliaries — generally serve as a morphological host for inflectional features that have been isolated from or are otherwise unable to be unite with the predicate (Bjorkman 2011). Consider the



familiar example of English copular *be*, which is used as an auxiliary verb in passive and progressive constructions (Bjorkman 2011). In both cases, *be* serves as obligatory morphological host for inflectional features.

(77) *English progressive*

- a. *Finite*: The children \*(**were**) eating the cake
- b. *Non-finite*: The children want to \*(**be**) eating the cake
- c. *Perfect*: The children have \*(**been**) eating the cake

(78) *English passive*

- a. *Finite*: The cake \*(**was**) eaten
- b. *Non-finite*: The cake seemed to \*(**be**) eaten
- c. *Perfect*: The cake has \*(**been**) eaten

English (Bjorkman 2011: 23)

Similar auxiliary uses of copular elements are attested in a variety of other Indo-European languages and beyond. Auxiliary uses of *belhave* are attested in both Latin and modern Romance languages (Bjorkman 2011), Hindi (Mahajan 1994), Basque (Laka 1993, Arregi 2000), and Arabic (Benmamoun 2000), among many others. More important for this dissertation is the fact that auxiliary uses of *belhave* are also reported in many Bantu languages, including Great Lakes Bantu languages (Botne 1986).

The most common auxiliary strategy in Bantu languages involves the use of a copular *be*-verb, usually *-li/-ri* (\*-*de*) or *-ba* (\*-*bà*). This can be seen in Kinyarwanda (JD61), which utilizes both copular verbs as auxiliaries (79).

(79) a. a-**ri**            mu-kw-eng-a  
           1SM-be-<sub>RI</sub> 18LOC-15INF-brew  
           ‘S/he is (in the midst of) brewing’

b. na-**bâ**-ye                            nĩ-ga            n-tégere-je            Mihigo  
           1SM-be-<sub>BA</sub>.PST-COMPL 1SM-study 1SM-wait-COMPL Mihigo  
           ‘I studied while waiting for Mihigo’                            Kinyarwanda (Botne 1986: 313-315)

Now, unlike what has been reported of auxiliary *be*-verbs in other languages, the auxiliary uses of *-li* and *-ba* are reported to be meaningfully distinct, i.e., they do not just serve as morphological hosts for inflection (Botne 1986). Specifically, Botne suggests that *-li* and *-ba* function as “shifters” that “refer to a second locus of orientation” that is “situated with respect to the speech event” (Botne 1986: 308). Without digging into this analysis too deeply, the general idea is that *-li* and *-ba* index different relationships between events. With *-li*, one event is interpreted relative to a particular moment in another event; at the exact moment of speaking in (79a), the subject is involved in an event of brewing. In this sense *-li* identifies a particular moment in an event to serve as the basis for interpreting another event. In contrast, *-ba* relates two events more broadly, yielding a durative or continuous reading; in (79b), “studying” occurs throughout the event of waiting, with no reference to a specific moment of an event. As I will discuss in the coming chapters, this difference in the auxiliary uses of *-li* and *-ba* is not unlike the interpretive contrast between *-li* and *-ba* in non-verbal predication in many Bantu languages, and may reflect a more general interpretive difference between reflexes of the proto-Bantu copulas *\*-de* and *\*-bà*.

## 2.5 Summary

In this chapter I presented a typological overview of non-verbal predication that illustrates the significant variation we see in its expression across languages. One way that this variation manifests is in the form of the copula (i.e., an additional functional element used to establish a predication relation); languages grammaticalize copular elements from a variety of grammatical sources, resulting in different “types” of copulas that are defined by their distinct sources. Another way this variation manifests is in how many copular forms are attested in a given language. So-called “complex copular systems” consisting of multiple distinct forms of the copula are observed in a variety of languages, and are reported to be sensitive to various morphosyntactic and interpretive factors, e.g., copular clause type, the interpretation of the predication relation, the person features of the subject, etc. Extending beyond variation in the expression of non-verbal predication, this chapter also explored other functions of the copula, e.g, their use as auxiliaries in complex verbal

constructions. Given these different ways of thinking about variation, the main takeaway from this chapter is that copular variation is extensive and multi-faceted: languages differ not only in terms of the form of the copula, but also in how many copulas they employ, and their various functions.

## Chapter 3

### Theoretical Background

#### 3.1 Revisiting the copula

In chapter 2, I introduced Pustet's (2003) definition of the term *copula* (repeated below) as a working description for the class of functional elements that help realize non-verbal predication relations.

- (7) "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."

This definition gets at one of the recurring claims about copulas in the literature, namely that they are meaningless (see "meaningless copula of predication" in Russell 1919). As subsequently mentioned however, this definition is also relatively broad in scope, as it (perhaps intentionally) makes no reference to the status of the predicate or the "predicative" nature of the copula.

On the one hand, this broad definition is quite amenable to uses of copulas outside of canonical non-verbal predication. Since it does not impose any restrictions on the nature of the predicate, e.g., verbal vs non-verbal, it can straightforwardly account for instances of copulas occurring with verbal predicates as reported in languages like Bambara (Niger-Congo, Mande; Pustet 2003).

- (80) ne            **be**    taa  
      1SG.PRO COP leave  
      'I am leaving' Bambara (Pustet 2003: 65)

On similar grounds, such a definition could reasonably account for some "auxiliary" uses of copulas in English as well, as these involve a similar surface structure to the Bambara example above (Arche et al. 2019).

That being said however, the definition presented in (7) still faces a number of challenges. In fact, one of the more significant of these is exactly what makes it generalizable to auxiliaries/verbal predication in the first place: it doesn't impose any restrictions on the predicate. As such, the definition as written faces an overgeneration problem; if copulas broadly co-occur with lexemes that function as predicates, then why do verbal predicates in English appear without copulas, e.g., *John reads*? Likewise, why are copulas often absent in small clauses, which also involve subject-predicate relations? Of course, we have seen that copulas can sometimes appear with verbal predicates, e.g., Bambara (Pustet 2003), and be found in small clauses, e.g., Kinande (Schneider-Zioga and Mutaka 2015a), however the definition in (7) provides no clear way to predict whether a copula will or will not appear in a given predicational context. In this sense, Pustet's definition lacks predictive power; it provides a broad description of what copulas are, but fails to capture their distribution or explain their function.

In an effort to better capture its true nature, many have proposed much more concrete — and often language-specific — definitions of copulas. One particularly popular approach that avoids the issue of overgeneration is what I will refer to as the “inflectional support” account. Under this approach, the core function of the copula — at least in some languages — is to host inflectional material that the predicate cannot bear (Hengeveld 1990, Carnie 1995, Baker 2003, Bjorkman 2011, Roy 2013, a.o.). English *be* is often cited as the prototypical example of an “inflectional support” copula (Baker 2003, Bjorkman 2011, O'Neill 2019, a.o.). For these accounts, English *be* is not a verb, but rather the realization of inflectional features associated with a functional head, e.g., T/Infl. In this sense, the copula serves only to provide “inflectional support” by ensuring that inflectional features have a morphological host — i.e., the presence of the copula ensures that the Stray Affix Filter (Lasnik 1981, 1995) is satisfied. Assuming that copulas serve only to realize “stranded” inflectional features (see Bjorkman 2011), it follows that copulas regularly appear with non-verbal predicates which cannot host inflection, but only rarely appear with verbal predicates which can. This approach also nicely accounts for why small clauses often don't feature overt copulas in many languages; since there is no tense, there is no copula.

An alternative approach to the issues faced by Putset’s definition is to explicitly define the copula as an element that is used to “predicativize” a lexeme. Consider the definition provided by Arche et al. (2019: 6) below (81).

(81) “A copular element is an element needed to define a predication structure.”

The definition provided by Arche et al. (2019) successfully sidesteps the overgeneration issue faced by Pustset’s definition because it restricts the distribution of copulas to predication structures that would otherwise be untenable. Under this definition, the reason why copulas generally do not appear with verbal predicates is because they do not need to; verbal predicates can establish a predication relation without an additional functional element. In contrast, non-verbal predicates cannot establish a predication relation alone, and need the copula to help “link” them to the subject. In many ways this resembles the inflectional support approach discussed above — if a predicate is unable to host inflectional features alone, a copula is required (at least in languages like English). However, the “predicational” definition in (81) has one distinct advantage over the inflectional support approach: it can account for copulas that are non-inflectional and/or occur in tenseless environments.

As shown in section 2.2, many languages employ copulas that are non-inflectional and/or do not pattern like verbs. For example, many Eastern Bantu languages like Kinande employ an invariant particle *ni* in certain copular clauses (Jerro 2015, Schneider-Zioga and Mutaka 2015a, Schneider-Zioga 2018, Gibson et al. 2019, a.o.). In contrast to other copular elements in Eastern Bantu languages, e.g., copular *be*-verbs like *-li/-ba*, the invariant particle *ni* noticeably lacks a number of characteristic properties of inflectional verbs, e.g., it lacks subject noun class agreement and is incompatible with tense-aspect-mood morphology.

- (69) a. Magulú **ní** mw-íbi  
 Magulu be<sub>NI</sub> 1-thief  
 ‘Magulu is a thief’
- b. Magulú á-<sup>✓</sup>**byá/\*ní** mw-íbi  
 Magulu 1SM-be<sub>ba</sub>.PAST/be<sub>NI</sub> 1-thief  
 ‘Magulu was a thief’

c. \*Magulú á-byá                      ní    mw-íbi  
 Magulu 1SM-be<sub>ba</sub>.PAST be<sub>NI</sub> 1-thief  
 ‘Magulu was a thief’                      Kinande (Adapted from Schneider-Zioga 2018: 20)

By all accounts, Bantu *ni* patterns separately from verbal predicates and ostensibly verbal copulas, e.g., *-li/-ba*, in that it does not serve as a host for inflection. This is a significant problem for the inflectional support approach, if we assume it to be cross-linguistically generalizable: if copulas are just support items, then why would they ever be incompatible with inflection? In contrast, the definition in (81) cleverly avoids this issue by making no reference to inflection while still restricting the general distribution of the copula. This definition is therefore capable of accounting for the use of copulas as both inflectional support items, e.g., English *be*, and purely predicative elements, e.g., Kinande *ni*.

Ultimately there are many other approaches to copulas that have been advocated for in the literature (see Arche et al. 2019), some of which I will discuss in the remaining sections of this chapter. As we will see, there a range of theoretical issues that come to light as we try to narrow down a definition for the term *copula*. One issue that we have already seen relates to the function of the copula: what does the copula *do*? Is the copula merely a support item, or does it serve a predicative purpose? Another related issue pertains to how the copula — and (non-verbal) predication in general — is structurally represented. What head(s) do(es) the copula realize in the syntax? Is there a dedicated functional head that is realized by the copula, or can copulas realize multiple, distinct heads? This, in turn, feeds other questions related to the semantic profile of the copula. If copulas are truly meaningless, then how do we explain interpretive contrasts like that between Spanish *estar/ser*? If on the other hand copulas are associated with distinct meanings, then how do we represent those differences? Are they lexically encoded, or are they associated with distinct structures?

In the following sections I will explore these theoretical issues, focusing specifically on the meaning of copulas and copular clauses, and how they are structurally represented.

## 3.2 Semantics of copular clauses

From a semantic point of view, the copula is often taken to be a meaningless piece of functional material: it relates two elements in a predicative relation, but makes no semantic contribution itself. This is explicitly noted in Pustet's (2003) definition in (7), and taken as a general point of departure in many other analyses. At the same time however, there is considerable cross-linguistic evidence that copulas are sensitive to certain kinds of interpretive factors, e.g., the stage/individual contrast, copular clause type, etc. Broadly speaking, such sensitivities present a challenge to the assumption that copulas are semantically meaningless. If copulas are meaningless, why do different forms correspond with distinct interpretations and/or information structures?

In this section, I will explore this question through a review of previous work concerning the meaning of copular clauses and their different types, and the meaning of the copula itself.

### 3.2.1 How many copular clause types?

Though copulas — in the broadest descriptive sense — help relate a predicate and subject, they may do so in a variety of ways. As discussed in section 2.3.2, copular clauses are not limited to basic property-describing, subject-predicate structures, e.g., *John is tall*. In fact, copular clauses can take on different interpretive profiles depending on the nature of the subject and predicate — i.e., whether they are referential — and how they are configured in the syntax (Arche et al. 2019). It is these two variables that define the four-way classification of copular sentences presented in Higgins (1979). In a predicational clause (56a), the post-copular, non-referential predicate ascribes a property to the referential subject (56a). In a specificational clause (56b), the pre-copular (non-referential) element describes an individual that is unidentifiable in the context, and the post-copular (referential) element identifies its reference. In an identificational clause (56c), the pre-copular element is deictically referential, however the identity of the referent is provided by the post-copular element. Finally in an equational clause (56d), both the pre- and post-copular elements are asserted to have the same reference.



(56) *Types of copular clauses*

- a. *Predicational:* Susan is a doctor
- b. *Specificational:* The winner is Susan
- c. *Identificational:* That woman is Susan
- d. *Equational:* John's mother is Susan

(Adapted from Mikkelsen 2005: 48)

Though the four-way split among copular sentences presented in Higgins (1979) is still taken as the default assumption by many, there have been considerable challenges to this classification (see Arche et al. 2019 for further discussion). One commonly cited issue with the four-way split is that it is difficult to motivate identificational sentences (56c) as their own distinct category. Consider the analysis presented in Mikkelsen (2005), which defines three distinct types of copular sentences — predicational, specificational, and equational — based on the semantics of their elements (82).

(82) *Three types of copular clauses*

- a. Predicational:  $\langle e \rangle$  is  $\langle e, t \rangle$
- b. Specificational:  $\langle e, t \rangle$  is  $\langle e \rangle$
- c. Equational:  $\langle e \rangle$  is  $\langle e \rangle$

(Adapted from Mikkelsen 2005, Arche et al. 2019)

For Mikkelsen (2005), “identificational” sentences are indistinguishable from specificational or equational sentences. Like in specification and equation, identification involves a post-copular referential NP (type *e*). As such, the only difference between identification and specification/equation should be in the semantics of the pre-copular element. However, Mikkelsen (2005) argues that the class of sentences Higgins (1979) refers to as “identificational” is not actually a homogenous class, but rather two interpretively distinct sub-classes. Moreover, Mikkelsen argues that these two sub-classes — namely, truncated clefts and demonstrative equatives — pattern like specificational and equational sentences, respectively. So-called truncated cleft sentences (83a) feature a bare demonstrative in pre-copular position, and are interpretively identical to specificational sentences

(82b).<sup>1</sup> On the other hand, demonstrative equatives (83b) involve a referential pre-copular element consisting of a demonstrative and noun, and are interpretively akin to equational sentences; the sentence in (83b) asserts that *that woman* and *Susan* refer to the same individual.

- (83) a. *Truncated cleft:*                    That is Susan                    ⟨e,t⟩ is ⟨e⟩  
       b. *Demonstrative equative:*        That woman is Susan        ⟨e⟩ is ⟨e⟩

As discussed in Arche et al. (2019), some have tried to reduce copular clause distinctions even further than a three-way contrast. One relatively popular position is that there are two core types of copular clauses: predicational and specificational (Blom and Daalder 1977, Moro 1997, 2000, den Dikken 2006, a.o.). For many of these accounts, truly “equative” copular sentences are deemed impossible. Consider the argument made by Moro (1988) (as summarized by Arche et al. 2019: 8), who shows that ostensibly “equative” copular sentences like (84) differ from true identity statements with respect to binding.

- (84) [The morning star] is [the evening star]

As Moro shows in a similar example (85a), placing a possessive pronoun in the post-copular NP that is coreferential with the pre-copular NP yields an ungrammatical result — a well-known property of copular sentences in binding theory (Arche et al. 2019). A similar binding violation is incurred in examples like (85b), where a pronoun contained within a predicative nominal is bound by the subject (Arche et al. 2019).

- (85) a. \*[The morning star]<sub>i</sub> is [its<sub>i</sub> source of light]  
       b. \*[John]<sub>i</sub> is [his<sub>i</sub> cook]

---

<sup>1</sup>Mikkelsen (2005) suggests that demonstrative *that* is predicative (type ⟨e,t⟩) and not referential (type e) because it can be anaphoric to property-denoting expressions (1a), but cannot refer to humans (1b).

- (1) a. They said that Sheila was [beautiful] and she is **that**  
       b. I ran into [my cousin Audrey downtown]<sub>i</sub>. I saw **that**\*<sub>i</sub> downtown earlier today

(Adapted from Mikkelsen 2005: 68)

In contrast, true identity statements featuring referential — not predicative — NPs do not exhibit the same binding effects (86). Likewise, transitive verbs do not feature the same facts either (87), suggesting that copular sentences differ fundamentally from transitive, verbal sentences.

- (86) a. [The morning star]<sub>i</sub> is equal to [its<sub>i</sub> source of light]  
b. [The morning star]<sub>i</sub> is one and the same as [its<sub>i</sub> source of light]
- (87) [The morning star]<sub>i</sub> lost [its<sub>i</sub> source of light]

For Moro, the reason why the “equatives” in (85a) and (86) differ is because the former features a predicative NP in post-copular position; “equative” copular clauses like (85a) are simply predicational clauses.

That being said, there has been significant debate in the literature as to whether truly equative copular clauses are possible or not. For example, Heycock and Kroch (1998, 1999) argue that the existence of equative copular sentences is motivated by the fact that a post-copular free relative predicate, e.g., *what I want a man to be*, can occur with both a property-denoting subject (88a) and a referential subject (88b).

- (88) a. Honest is what I want a man to be  
b. John is what I want a man to be (i.e., he’s honest) (Heycock and Kroch 1998: 5)

Assuming that equational copular sentences do not exist, it is difficult to pin down the type of the free relative in (88). In (88a), the free relative would have to be a generalized quantifier of type  $\langle\langle e,t \rangle, t\rangle$  (it denotes a set of sets of individuals) since the subject *honest* is a predicate of type  $\langle e,t \rangle$ . At the same time, to capture the example in (88b) the same free relative must also be able to be a predicate of type  $\langle e,t \rangle$  since the subject *John* is interpreted as a constant, and is of type  $e$ .<sup>2</sup> This is what Heycock and Kroch (1998) call the “type identity problem”; the free relative must be

---

<sup>2</sup>This argument is predicated on the idea that *John* is of type  $e$  in (88b). One possible argument against this is that *John* could reasonably be argued to denote a property in this example. In my interpretation of (88b), the speaker identifies John as having certain properties that exemplify what they want a man to be (e.g., he’s honest, tall, nice, etc.). Crucially however, John is not literally what the speaker wants a man to be; the speaker does not want a man to be the individual *John*, but rather have some features that John possesses. In this sense John is not a constant, but rather an exemplar of a salient property (e.g., “honest”, “tall”, etc.).

type ambiguous to capture both examples in (88). That being said, if we assume that equational copular sentences are possible, then this ambiguity issue is irrelevant; the free relative is type  $\langle e,t \rangle$  in both (88a) and (88b), but the former example constitutes an equational sentence, and the latter a predicative sentence.

In summary, it is an open question as to how many types of copular clauses there are, at least from a semantic perspective. Following the four-way classification in Higgins (1979), various proposals have attempted to narrow down the number of copular clauses types in some way, with most settling on either a ternary contrast between predicational, specificational, and equational clauses, or a binary contrast between predicational and specificational clauses.

### 3.2.2 Meaning of the copula

As we have already seen, there are many languages that employ multiple distinct copulas in non-verbal predication (see section 2.3). In some languages, this corresponds to a difference in interpretation, e.g., Spanish *estar/ser* broadly distinguish between stage- and individual-level interpretations (Maienborn 2005, Roy 2013, a.o.). In other languages, copular contrasts are instead linked to morphosyntactic factors like predicate category, e.g., Kinyarwanda *-ri/ni* distinguish locative and non-locative predication (Jerro 2015). Moreover, we have also seen that there are different types of copular clauses, e.g., predicational, specificational, equational, identificational, etc. (Higgins 1979, Heycock and Kroch 1998, Mikkelsen 2005, a.o.; see section 3.2.1). These clause types crucially differ both in their configuration, i.e., the position of the two related elements with respect to the copula, and their interpretation, i.e., property ascription vs equative statement, etc.

Given the degree of variation and complexity in copular constructions across languages, it would be reasonably natural to surmise that different copulas may coincide with distinct structural and/or semantic representations. Indeed, one possible approach would be to assume that interpretive contrasts and clause type differences arise due to differences in the copula. Focusing specifically on clause type, one could posit that there are four (or fewer depending on how many clause types you assume) semantically distinct copulas (Mikkelsen 2011). These copulas may be realized by

the same morphosyntactic form, e.g., *be* in English, but each would be associated with a distinct semantics, e.g., an equational copula equates individuals, while a predicational copula ascribes properties to individuals.

The exact number of copulas and their respective meanings has been strongly debated in previous work on copular clauses, particularly in light of Higgins (1979) four-way taxonomy (see Mikkelsen 2011). Within this debate there are broadly three theoretical camps: those that posit three copulas, those that posit two copulas, and those that posit one sole copula. In what follows, I will provide an abbreviated summary of the extensive review of work on “multiple-be” presented in Mikkelsen (2011).

Overall, relatively few authors overtly propose a three-copula system, however some indirectly imply such a system, e.g., Schlenker (2003), Romero (2005), Comorovski (2007). Broadly speaking, these accounts take predicational, specificational, and equational sentences to involve fundamentally distinct copulas, though they tend to focus on the difference between specificational and equational copulas in particular (Mikkelsen 2011). The core of the analysis (as summarized in Mikkelsen 2011) is that the predicational copula attributes properties to individuals, the specificational copula equates propositions, and the equational copula equates individuals.

A relatively more popular position instead holds that there are two distinct copulas: an equative copula, and a predicative copula (Mikkelsen 2011). This approach broadly follows the ideas of Russell (1919: 174) — among many others — in distinguishing between an equative copula of identity (89a) and a meaningless copula of predication (89b).

(89) a.  $\llbracket \text{be}_{ident} \rrbracket = \lambda x \lambda y [y=x]$

b.  $\llbracket \text{be}_{pred} \rrbracket = \lambda x \lambda P [P(x)]$  (Mikkelsen 2011: 1814)

The idea here is that, in addition to a standard copula of predication that attributes properties to individuals, there is also an equative copula that equates the reference of two entities, regardless of type, e.g., in (89a), *y* could be of type *e* or  $\langle e, t \rangle$  (90) (Heller 2005) .

(90) a. John’s mother is Susan                       $\langle e \rangle$  is  $\langle e \rangle$                       (Mikkelsen 2005: 48)

b. Honest is what I want a man to be  $\langle e,t \rangle$  is  $\langle e,t \rangle$  (Heycock and Kroch 1998: 5)

This sort of type flexibility — or polymorphism (Partee 1986, Mikkelsen 2011) — is often assumed to be an implicit property of the copula, regardless of how many copulas one posits. For example, Partee (1986) defines the traditional, predicative copula as taking “arguments of types  $X$  and  $\langle X, t \rangle$ , for any type  $X$ ” (Partee 1986: 355), meaning it is of type  $\langle \langle X, t \rangle, \langle X, t \rangle \rangle$ . On the other hand, an equative copula like that in (90b) would naturally take two arguments of type  $X$ , making it type  $\langle X, \langle X, t \rangle \rangle$  (Mikkelsen 2011).

While two-copula accounts generally posit the same binary distinction as shown in (89), they do not all agree on the distribution of the two copulas (Mikkelsen 2011). For example, though both assume the binary distinction in (89), Mikkelsen (2005) assumes that the copula of identity is limited to equational sentences, while Heller (2005) assumes that it is found in both equational and specificational sentences.

Moving to the final, and perhaps most popular approach, one-copula accounts like Williams (1983) assume that all copular clauses involve a meaningless copula of predication like that in (89b). Under this view, the function of the copula in all copular sentences is to relate a predicative element to a referential one (Mikkelsen 2011).

To capture the full distribution of the copula across different clause types, Williams (1983) — among others — makes two claims about the nature of the copula. First, the copula must be able to combine with its arguments in either order (91). This is necessitated by the distinct configurations of predicational and specificational sentences, with the latter involving a “reversal” of arguments, e.g., a pre-copular predicate.

$$(91) \quad a. \quad \llbracket \text{be} \rrbracket = \lambda P \lambda x [P(x)]$$

$$b. \quad \llbracket \text{be} \rrbracket = \lambda x \lambda P [P(x)] \quad (\text{Mikkelsen 2011: 1815})$$

Second, the copula must be able to relate two referential individuals despite its predicational semantics, as illustrated by examples like (90). There are various ways that this particular issue has been approached in the literature. For many, this involves type-shifting either the post-copular

element or the copula itself. For Partee (1986), equational clauses involve a type-shifter IDENT that derives a predicate from the post-copular referential element, allowing the (predicational) copula to compose with it. Geist (2007) on the other hand proposes that the type-shifter IDENT instead composes with the predicational copula, yielding a derived copula of identity (see Geist 2007: 89 and Mikkelsen 2011).

Other single-copula approaches account for — or largely avoid — these issues through other means. For example, accounts like Heycock and Kroch (1999), Adger (2003), Baker (2003), Balusu (2014), and Myler (2016), a.o., assume that the “copula” doesn’t directly mediate the relationship between the subject and predicate. Instead, the subject and predicate compose — either directly or via a Pred head (see below) — as part of a small clause, which the “copula” then combines with. Under this view, the copula — or rather the head the copula realizes,  $v_{BE}$  — is merely a means of linking the small clause to higher functional heads in the extended verbal projection (Myler 2016). In this sense, the copula is nothing more than a type-neutral identity function (92). Importantly, this approach shifts the burden of accounting for argument ordering and type flexibility away from the copula; the relative ordering of elements and their composition are assumed to be properties of the small clause and/or other syntactic processes.<sup>3</sup>

$$(92) \quad \llbracket \text{be} \rrbracket = \lambda x.x \quad \text{(adapted from Myler 2016: 42)}$$

As suggested by the above discussion, the vast majority of “multi-copula” analyses focus on the distinctions between copular clause types in the sense of Higgins (1979). For many, there is an assumption that the existence of predicational, specificational, and equational clauses — or some subset thereof — motivates the existence of multiple distinct copulas. That being said, there are other typological phenomena that could reasonably be taken as evidence for multiple copulas. In fact, many languages employ distinct copulas to distinguish between things other than clause type (see section 2.3).

---

<sup>3</sup>Various explanations of the ordering of elements in predicational and specificational clauses have been proposed in the literature (see section 3.3.2). Some of these accounts propose that different clause types merge elements at different positions in the tree, resulting in different surface orders, e.g., predicate nominals merge higher in specificational clauses than in predicational clauses. Others instead suggest that such ordering differences are derived through independent syntactic processes that result in the inversion of the predicate nominal across the logical subject.

In Kinyarwanda (JD61) for example, the two copulas *ni* and *-ri* distinguish between standard predication and locative predication in certain morphosyntactic contexts (Jerro 2015). In present tense predicational clauses involving a third person subject, *-ri* is used with locative predicates, while *ni* is used with all other predicates.

- (67) a. Mukamana a-**ri** \*mu-nini/\*umw-alimu/√mu-rugo  
 Mukamana 3SM-be-<sub>RI</sub> 1-big/1-teacher/18LOC-house  
 ‘Mukamana is big/a teacher/at home’
- b. Kyle **ni** √mu-nini/√umw-alimu/\*mu-rugo  
 Kyle be<sub>NI</sub> 1-big/1-teacher/18LOC-house  
 ‘Kyle is big/a teacher/at home’
- Kinyarwanda (Jerro 2015: 93-96)

For Jerro (2015), this amounts to a semantic difference between *ni* and *-ri*; *ni* is a standard (polymorphic) copula of predication, while *-ri* is a copula specifically used for locative descriptions.

Now, one could ask whether Kinyarwanda *ni/-ri* are two semantically distinct copulas, or two suppletive allomorphs of a single predicative copula. For example, one could reasonably propose that *-ri* is a contextual allomorph that appears exclusively with morphologically locative predicates (at least in the presence of a third person subject). However, an allomorphy account of this type would have a relatively difficult time accounting for why *-ri* is limited to this environment, particularly given that it is used robustly in present tense sentences with first/second person subjects, and in all past tense copular sentences. While I will not overtly compare the allomorphy approach to the multi-copula approach here (see however chapter 5), the point I am trying to make is that there is a meaningful difference in interpretation between Kinyarwanda *ni* and *-ri*, and one way to account for that difference is to posit two semantically distinct copulas rather than two forms of a single copula.

Importantly there are many other languages that make interpretive distinctions in their copula systems. As with Kinyarwanda, such contrasts often lend themselves to the idea of multiple semantically distinct copulas. In Spanish for example, the copulas *estar/ser* are reported to distinguish stage- and individual-level predication, or something similar (Milsark 1974, Carlson 1977, Kratzer 1995, Diesing 1990, 1992b, Arche 2006, a.o.): *estar* is used with stage-level predicates

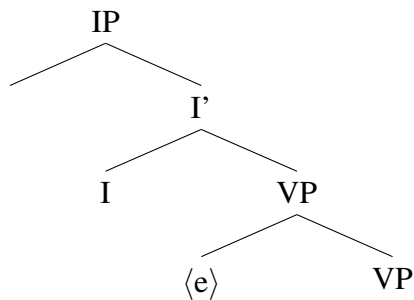


and yields "temporary" interpretations, while *ser* is used with individual-level predicates and yields "permanent" interpretations.

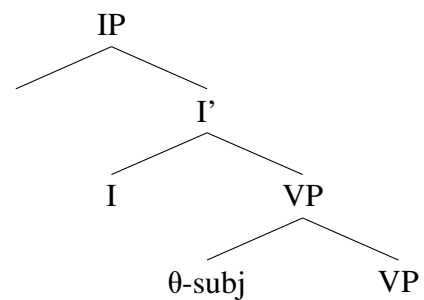
- (44) a. El artista  $\checkmark$ está/#es presente/ausente/lejos  
 The artist estar/ser.PRES.3SG present/absent/far away  
 ‘The artist is present/absent/far away’ Spanish (Maienborn 2005: 173)
- b. Maria #está/ $\checkmark$ es inteligente/altruista  
 Maria estar/ser.PRES.3SG intelligent/altruistic  
 ‘Maria is intelligent/altruistic’ Spanish (Deo et al. 2017: 8)

Given this interpretive difference, the question naturally arises as to where this distinction comes from: do *estar/ser* correspond to two semantically distinct copulas, or is this difference derived another way? As discussed in Arche (2012), the most popular conceptualization of the stage/individual contrast is arguably that presented in Kratzer (1995), which ties the interpretive contrast to the presence or absence of a davidsonian eventuality argument *e* (93).

(93) a. *Stage-level predicates*



b. *Individual-level predicates*



(Arche 2012: 112)

Under this approach, the stage/individual contrast is purely a property of the predicate: certain predicates (i.e., stage-level predicates) come pre-packaged with an eventuality argument, while others do not. In this sense, the contrast between *estar/ser* could be reduced to a contrast between stage/individual level predicates: the nature of the predicate determines which copula will be realized. From a morphosyntactic point of view, this analysis is compatible with a view of *estar/ser* as purely allomorphic variants of a single, meaningless copula, or as two distinct heads with different

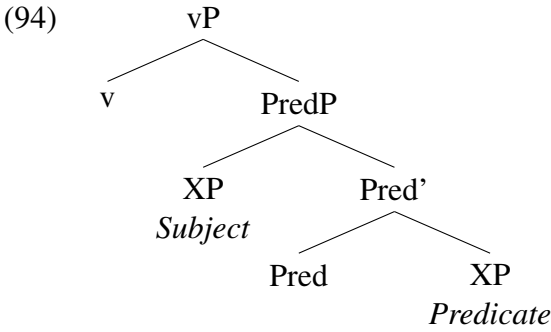
selectional properties. The allomorphy approach assumes that the form of the copula is determined by the predicate, with the copula being pronounced as *estar* in the presence of a stage-level predicate, and as *ser* in the presence of an individual-level predicate. In contrast, the selectional approach assumes that *estar* and *ser* realize distinct heads; the head realized as *estar* selects for predicates that include an eventuality argument, while the head realized as *ser* selects for predicates that lack an eventuality argument.

That being said, many have argued against a pure stage/individual analysis for Spanish *estar/ser*. One of the core observations within this camp is that certain predicates allow for either *estar* or *ser* to be used, albeit with different interpretations (47).

- (47) a. La chaqueta le **está** ajustada en los hombros  
 The jacket CL-IO estar.PRES.3SG tight.FEM on the shoulder.PL  
 ‘The jacket is tight on the shoulders’ (it doesn’t fit me)
- b. La chaqueta **es** ajustada y muy moderna  
 The jacket ser.PRES.3SG tight.FEM and very fashionable.FEM  
 ‘The jacket is tight and very fashionable’ (it is tight by design) Spanish (Deo et al. 2017)

The fact that both *estar* and *ser* can appear with the predicate *ajustada* ‘tight’ is unexpected under a rigid structural view of the stage/individual contrast, e.g., (93). Unless we posit two homophonous predicates — one stage-level version and one individual-level version — there is no reason to expect a single predicate to appear with both copulas. Moreover, there is ostensibly an interpretive difference between the examples in (47), suggesting that the copula is not as meaningless as previously assumed; the predicate doesn’t seem to dictate the interpretation alone.

One way of circumventing these issues is to assume that the stage/individual contrast is not derived by the presence or absence of an eventuality argument, but rather the presence of a particular functional head, i.e., Pred (Bowers 1993, Adger 2003, Baker 2003, Markman 2008a, Balusu 2014). The idea is that at the core of all copular clauses is a predicational head, Pred, that serves to relate the subject and predicate low. In this account, the copula realizes a meaningless head  $v_{BE}$  that takes the small clause headed by Pred as its complement.



(Adapted from Myler 2016)

Importantly, it is generally assumed that there exist two distinct Pred heads — Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub> — and that the choice of Pred head determines the interpretation of the predication relation (Adger 2003, Markman 2008a, Balusu 2014). In Spanish, these two heads coincide with distinct copulas: with Pred<sub>STAGE</sub> the copula is pronounced as *estar*, while with Pred<sub>INDIV</sub> it is pronounced as *ser*. By situating the stage/individual contrast on Pred, this account has a straightforward explanation of both the distributional flexibility of *estar/ser* and their interpretive differences; either variant of Pred can merge with a predicate (so long as they do not clash with the lexical semantics of the predicate), however each head yields a distinct interpretation (see section 3.3.3 for further discussion).

Another way of accounting for the distribution and interpretation of *estar/ser* is to assume that they are in fact two semantically distinct copulas. Building on Clements (1988) and Maienborn (2005), Deo et al. (2017: 27) characterize *estar/ser* as presuppositional variants that differ minimally in how they convey the truth of the proposition they embed. For them, *estar* presupposes that the proposition is boundedly true; the circumstance of evaluation is the maximal circumstance in which P(x) is true. As such, the use of *estar* in (47a) yields a "temporary" reading of *ajustada* 'tight'; the jacket is tight in this context, but not necessarily outside it. In contrast, Deo et al. (2017) suggest that *ser* conveys that the proposition is unboundedly true, i.e., there is no limiting context of evaluation. This is why *ser* yields a "permanent" or "individual-level" reading in (47b); the jacket is perceived as being tight independent of context.

There are other ways of accounting for meaning contrasts among copulas that I will discuss in section 3.3.3. For now, it suffices to say that there is not only considerable disagreement in how

many types of copulas there are, but also what those copulas should express, and what they realize in the syntax.

### 3.3 Structure of copular clauses

#### 3.3.1 Realizing the copula

As summarized by Arche et al. (2019), most analyses agree that the copula is, in a broad sense, a support element used to establish a (non-verbal) predication relation. That being said, analyses differ quite radically in terms of how they treat the copula syntactically, i.e., what syntactic category the copula instantiates. In an effort to summarize this variation, Arche et al. (2019) distinguish three general approaches (95).

- (95) i. Copulas are the spell out of T/Infl
- ii. Copulas are the spell out of a type of v/V
- iii. Copulas are the spell out of predication heads (Arche et al. 2019: 16)

Accounts falling under the first category in (95) broadly claim that the copula serves primarily as inflectional support. For these accounts, the copula is not a verb, but rather a morphological means of hosting inflectional information associated with T/Infl, i.e., the copula realizes the head T/Infl (Hengeveld 1990, Carnie 1995, Doherty 1996, Baker 2003, Bjorkman 2011, Roy 2013, O'Neill 2019, a.o.). The copula appears in contexts where inflectional features have been "stranded" (Bjorkman 2011). When a predicate cannot host inflectional features on its own, the copula provides a morphological host for the features associated with T/Infl. In this sense the copula is little more than a last resort operation; the copula is inserted as a means of satisfying the Stray Affix Filter (Lasnik 1981, 1995).

As noted by Arche et al. (2019), the "inflectional support" approach has merit in that it can account for some key typological facts about copulas. First, it provides a straightforward explanation for why copulas sometimes appear with both categorically non-verbal predicates and non-finite

verbs in inflectional languages; if the predicate cannot bear inflectional features, the copula spells out T/Infl (Arche et al. 2019). Second, assuming that copulas do not contribute any semantic meaning, this approach also has a straightforward explanation for why copulas are semantically vacuous. Since the copula is merely a means of rescuing stranded inflectional features, it does not contribute to the semantics.<sup>4</sup> Finally, this approach also provides a feasible means of accounting for tense-/polarity-conditioned alternations between zero copulas and overt copulas in some languages. Arche et al. (2019) illustrate this in Arabic, which features a zero copula in the present tense but an overt copula in the past tense. As shown in section 2.3.3, Russian exhibits a nearly identical alternation; though a zero copula is used in the present tense (68a), an inflected form of the copula is obligatory in the past tense (68b).

- (68) a. Segodnja reka spokojna  
today river.FEM calm.FEM.SF  
‘Today the river is calm’
- b. Ivan byl goloden  
Ivan be.PAST.MASC hungry.MASC.SF  
‘Ivan was hungry’
- Russian (Adapted from Roy 2013: 119)

As pointed out by Arche et al. (2019) however, the “inflectional support” view of the copula faces considerable theoretical challenges. The first and perhaps most serious of these is that it cannot account for the use of copulas in non-inflectional environments. For example, Mandarin — an analytic language that does not morphologically mark tense or subject agreement — employs a copula *shi* (19) (Hengeveld 1990, Arche et al. 2019). If the copula is purely an inflectional support item, one would predict that an analytic language like Mandarin would not feature a copula at all, since there would presumably be no “stranded” features to save. The same issue could reasonably be extended to small clauses in inflectional languages as well; if copulas appear in reduced clauses that lack tense, can they universally be introduced in T/Infl? There are a few other salient issues faced by the inflectional support approach, e.g., co-occurrence with modals, sensitivity to predicate,

<sup>4</sup>Accounts differ as to whether the copula is syntactically represented at T/Infl, or rather inserted as a last resort operation at PF (Bjorkman 2011). I will not discuss this distinction here, though see Bjorkman (2011) and Arche et al. (2019) for considerable discussion.

etc., however I will direct the reader to Arche et al. (2019) for further discussion of these issues.

Skipping to the third broad category in (95), another popular proposal posits that the copula realizes a functional head, e.g., Pred (or Pr), that mediates the relationship between a (non-verbal) predicate and its subject (Arche et al. 2019). This approach builds heavily on earlier work on small clauses that assume an asymmetric relationship between the subject and predicate of the small clause (Williams 1975, 1983, Kayne 1985, Bowers 1993, a.o.).

Going back as far as Williams (1975), it has been observed that noun-adjective or noun-noun combinations like *John crazy/a fool* can sometimes form a constituent that exhibits clause-like properties. Like main clauses, so-called “small clause” constituents express full, independent propositions, however, unlike main clauses (96b-c), they do not contain any inflectional material (Stowell 1982, Bruening 2018).

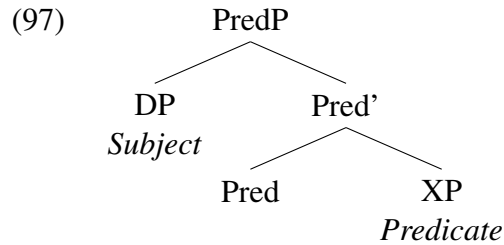
(96) a. I consider [<sub>SC</sub> John crazy/a fool ]

b. John **is** crazy/a fool

c. John **will** laugh

(Adapted from Bowers 1993)

Under the assumption that the relation between the elements of the small clause is predicational, i.e., that *crazy/a fool* are predicates that attribute a property to *John* (96a), accounts like Williams (1983), Bowers (1993, 2001), and Mikkelsen (2005) attempt to provide a unifying structural analysis of predication that accounts for both small clauses and traditional (verbal) predication. Following research suggesting that the external argument is generated in a position below T — the VP-internal subject hypothesis (Fukui 1986, Kuroda 1988, Pollock 1989, Koopman and Sportiche 1991) — Bowers (1993, 2001) proposes a functional head, Pred (also Pr; for “predication”), between T and V. Broadly, the Pred head serves two key functions: syntactically, Pred introduces the external argument and establishes a hierarchical relation between it and the predicate, while semantically it serves to establish a predication relation between the predicate and external argument. In both main clauses and small clauses, the Pred head selects a predicate XP as its complement, and introduces the external argument in Spec, PredP (97).



Under this analysis, English main clauses and small clauses differ only in where PredP is situated, and what material — if any — realizes the Pred head. In main clauses featuring a verbal predicate, PredP is selected by T/Infl, and the Pred head is phonologically null, e.g., Pred is realized as e “empty” (98). In small clauses, PredP is instead selected by the matrix verb, and the Pred head is either not realized at PF (99), or realized by *as* (100).

- (98) a. John will laugh  
 b. [IP will [PredP John [Pred' e [VP laugh ]]]]

- (99) a. I consider John crazy  
 b. [PredP I [Pr' e [VP consider [PredP John [Pred' e [AP crazy ]]]]]]

- (100) a. They regard John as crazy  
 b. [PredP they [Pred' e [VP regard [PredP John [Pred' as [AP crazy ]]]]]]

(Adapted from Bowers 1993)

Importantly, non-verbal predication in main clauses is taken to be structurally identical to verbal predication under this approach. The sole difference between main clauses involving a verbal predicate (98) and copular clauses (101) is that the Pred head is realized by the copula in the latter.

- (101) a. John is crazy.  
 b. [IP PRES [PredP John [Pred' *be* [AP crazy ]]]]

(Adapted from Bowers 1993)

The idea that copulas realize a Pred head has been proposed for a variety of languages in addition to English, e.g., Modern Irish (Bowers 1993), Japanese (Nishiyama 1999), Edo, Mandarin, Somali, and Berber (Baker 2003), and Kinande (Schneider-Zioga and Mutaka 2015b,a).

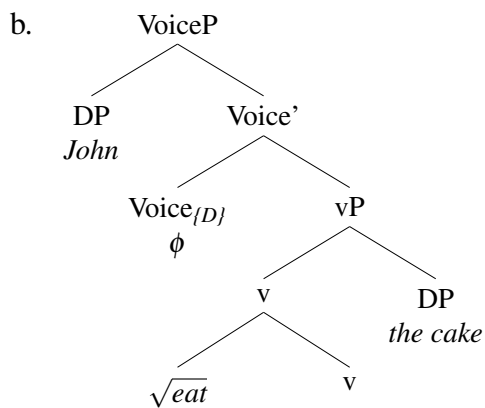
As an alternative to proposals which take the copula to realize either a verbal inflectional head (e.g., T/Infl) or a predicational head (e.g., Pred), some posit that copulas — or at least, copular verbs — are light verbs (Halle and Marantz 1993, Harley 1995, Moro 1997, Mikkelsen 2005, a.o.). This is the second approach in (95): copulas are the spell out of a type of  $v/V$ . Assuming the definition of a light verb in (102), one set of accounts takes copulas to specifically realize the head  $v$  (or  $v_{BE}$ ).

(102) *Definition of a light-verb construction*

A light-verb construction is one that contains a  $v$  but no root. (Myler 2016: 24)

Broadly speaking, constructivist approaches of this type assume that argument/event structures are derived via the combination of distinct heads in the syntax, as opposed to being projected by verbs themselves (Kratzer 1996, Hale and Keyser 2002, Baker 2003, Borer 2005, a.o.). Such accounts generally take the thematic domain to consist of two key functional projections: one to introduce the external argument (e.g., VoiceP, in the Kratzerian tradition), and one to introduce an eventuality variable (e.g.,  $v$  à la Pesetsky 1995, Marantz 1997). Importantly, traditional “verbal roots” are treated as acategorical roots that modify some aspect of an eventuality. Consider a simple transitive verb like *eat* in (103). Here, Voice serves to introduce the external argument, *John*, while  $v$  introduces an eventuality variable and selects a DP complement. The acategorical root  $\sqrt{eat}$ , adjoins to  $v$  and serves as an event modifier — it derives an eventuality of eating.

(103) a. John ate the cake

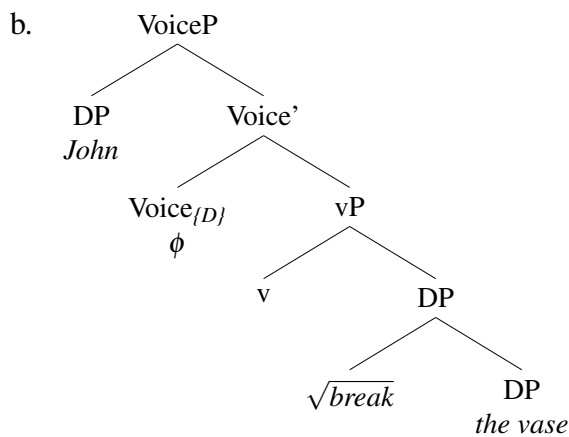


(Myler 2016: 20)



With a change-of-state predicate like *break*, things are largely similar. Voice again introduces the external argument (if present), and *v* introduces an eventuality variable and selects for a DP complement. However, to capture the fact that the “action” eventuality in (104) is distinct from the stative eventuality in (103), the acategorial root  $\sqrt{break}$  adjoins to the DP object rather than *v*. In examples like these, the function of the acategorial root is simply to identify the state of the DP complement of *v* (Marantz 2009a, Myler 2016).

(104) a. John broke the vase



(Myler 2016: 22)

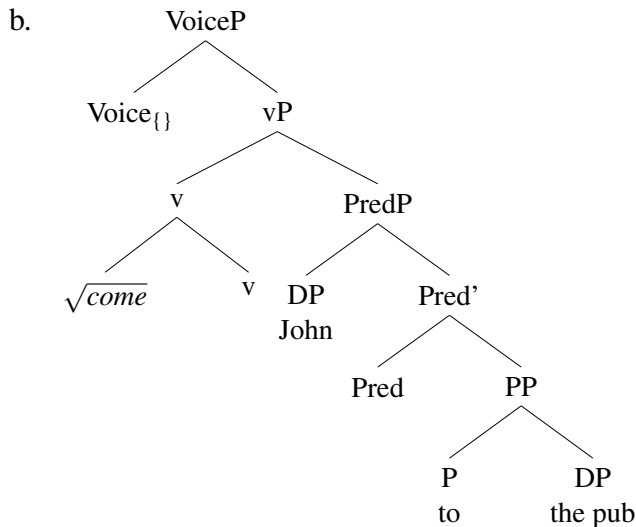
Things are slightly different in contexts where the complement of *v* involves a predication relation itself — a so-called small clause (Adger and Ramchand 2003). Here, accounts differ in how they handle the small clause. For accounts like Adger and Ramchand (2003), Baker (2003), Balusu (2014), a.o., the small clause represents the “predicational core” of the clause where thematic relations are established. As we have seen, this smaller thematic domain is assumed to be encased within a functional projection PredP, whose head, Pred, serves to relate the predicate and subject of the small clause. When this small clause is present, Voice no longer serves to introduce the external argument. Instead, Voice is taken to be expletive in the sense that it “passes the denotation of its complement up the tree” (Myler 2016: 43).<sup>5</sup>

<sup>5</sup>Following Wood (2015), Myler (2016) defines the following three interpretations of Voice:

- (1) *Rules for the interpretation of Voice* (Myler 2016: 43)
  - a.  $\llbracket \text{Voice} \rrbracket = \lambda x_e . \lambda e_s . \text{Agent}(x, e) / \text{_____}$  (agentive, dynamic event)

In (105)  $v$  takes a small clause complement [*John to the pub*]. Since the logical subject is introduced in the small clause, Voice is taken to be expletive. Just as in a transitive configuration, the acategorial root, here  $\sqrt{\text{come}}$ , adjoins to  $v$  as an event modifier.

(105) a. John came to the pub



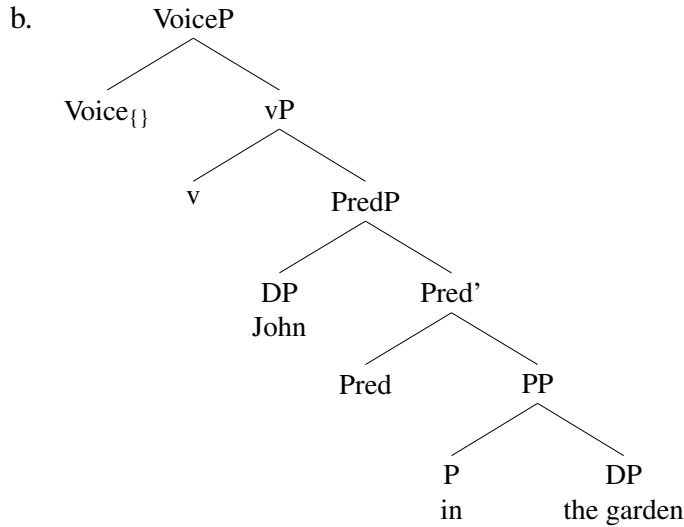
(Myler 2016: 23)

Under the PredP analysis, non-verbal predication is geometrically identical to examples like (105), with the sole difference being the absence of an acategorial root. In a copular sentence involving a locative predicate like (106), the Pred head serves to attribute a location (*in the garden*) to the subject (*John*). Given the absence of an acategorial root, the copula *be* realizes  $v$  and provides a morphological host for tense.

(106) a. John is in the garden

- 
- b.  $\llbracket \text{Voice} \rrbracket = \lambda x_e. \lambda e_s. \text{Holder}(x,e) / \text{_____}$  (stative eventuality)  
 c.  $\llbracket \text{Voice} \rrbracket = \lambda x.x / \text{_____}$  (elsewhere)

The first interpretation occurs when  $vP$  denotes a set of dynamic events. In this case, Voice serves to integrate the external argument as an agent. The second interpretation occurs when  $vP$  instead denotes a set of states. In this case, Voice integrates the external argument as the holder of that state. The third interpretation occurs in all other cases. Here, Voice does not introduce an external argument, but rather functions as a type-neutral identity function.



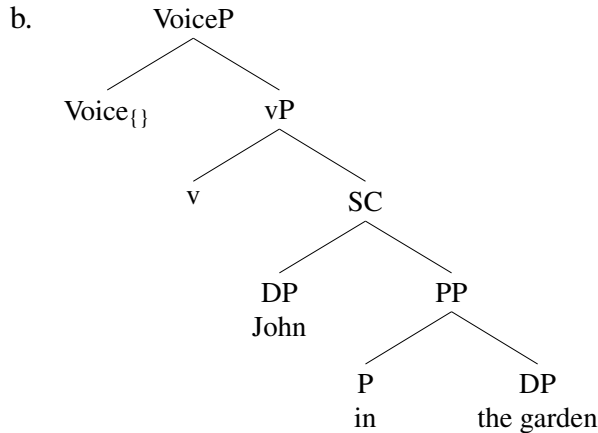
(Myler 2016: 27)

In contrast, accounts like Stowell (1995) and Moro (1995, 1997) assume a different structure (or at least, a second possible structure) for the small clause “core” of copular sentences. In lieu of a Pred head, these accounts suggest that small clauses relate the subject and predicate symmetrically; a small clause contains two elements X-Y such that X and Y are sisters, and X is always the subject. In so doing, the “bare” small clause approach is argued to account for the apparent reversibility of elements in equational copular clauses, e.g., *John is the professor/The professor is John* (Roy 2013; see discussion in section 3.3.2 for more on inversion). Structurally, Moro (1997) defines the bare (or “flat”) small clause as a case of adjunction in which the subject DP is adjoined to the predicate DP (or presumably XP in the case of other predicates).

(107)  $[_{DP=SC} [_{DP_S}] [_{DP_P}]]$  (Moro 1997: 57)

Labelling issues aside (see Lohndal 2006), this analysis makes the same general assumption about the realization of the copula as the PredP approach: the copula realizes a verbal head that selects a small clause as its complement. Assuming that this verbal head is *v*, Moro’s (1997: 97) analysis would look something like (108).

(108) a. John is in the garden



Regardless of whether Pred is present in the small clause or not, the function of  $v$  in copular sentences is assumed to be different than its function in sentences featuring a finite verb, which crucially feature an acategorial root that adjoins to  $v$  or its complement DP (104)-(105). Synthesizing previous work by Bach (1967), Lyons (1968), Roy (2013) and Wood (2015), a.o., Myler (2016, 2018) posits two structurally distinct varieties of  $v$ : substantive  $v$  and copula  $v$ .

Substantive  $v$  is the standard categorizing head that introduces an eventuality variable in verb constructions. For Wood (2012, 2015), this head comes in three flavors. The first two alloemes of  $v$  denote an activity (109a), and a state (109b), respectively. In general either alloeme may enter into the derivation so long as it is compatible with the lexical semantics of the predicate (Myler 2016). Following Marantz (2009a,b), Wood defines the third alloeme as a causative variant of  $v$  that appears in cases where its complement denotes an eventuality (109c).

(109) *Substantive v* (Wood 2012: 37)

- a.  $\llbracket v \rrbracket = \lambda e_s. \text{activity}(e)$
- b.  $\llbracket v \rrbracket = \lambda e_s. \text{state}(e)$
- c.  $\llbracket v \rrbracket = \lambda P_{\langle s, t \rangle}. \lambda e_s. \exists e'_s. \text{activity}(e) \wedge \text{CAUSE}(e, e') \wedge P(e') / \text{_____}(\text{eventuality})$

In contrast to substantive  $v$ , copula  $v$  — alternatively,  $v_{BE}$  — makes no thematic contribution in that it does not introduce an eventuality variable (Myler 2016, 2018). Instead, copula  $v$  serves a purely syntactic function; it serves to relate its non-verbal complement — e.g., PredP (Myler 2016, 2018) — to higher functional heads in the verbal extended projection that would otherwise be

inaccessible. By presenting copula  $v$  as a type-neutral identity function (110), Myler provides a thematically vacuous means of integrating non-verbal elements into the verbal extended projection.

(110) *Copula v* (Myler 2016: 42)

$$\llbracket v \rrbracket = \lambda x.x$$

To summarize this section, we have seen that there are broadly three ways of thinking about the syntactic realization of the copula. One approach assumes that the copula is inflectional support; it realizes an inflectional head (e.g., T/Infl) in the verbal extended projection. Another approach takes the copula to serve a purely relational function; it realizes a lower predicational head (e.g., Pred) that relates the subject and predicate. Finally, the third approach treats the copula as a light-verb; it realizes  $v/V$ . While these three approaches exist in opposition, that is not to say that they are mutually exclusive. By this I mean that it is unclear whether any of these approaches can account for copular predication data across all the world's languages. Indeed, many authors suggest that certain analyses that work in one language may not work in another (see Baker 2003 and Mikkelsen 2011 for discussion). As such, it is possible that there is no single unifying account for copular predication, at least based on our current theoretical models.

Before moving on there is one additional approach that I would like to mention, as it doesn't quite fit into the three-way dichotomy outlined above. Similar to accounts that posit a predicational head Pred, den Dikken (2006) assumes that predication (both verbal and non-verbal) fundamentally involves a functional head that serves to relate two elements. In contrast to the PredP approach however, the analysis presented in den Dikken (2006) doesn't assume the presence of a highly specialized functional head (e.g., Pred) in all types of predication. Instead, den Dikken (2006) suggests that most predication relations may be expressed using existing functional material.

In this analysis, any functional head within the thematic/inflectional domain can (in theory) relate a predicate to its subject, however the particular functional head used to express this relation will vary by predication type. For den Dikken (2006), this is the crucial difference between verbal and non-verbal predication; while the tense head (T) relates a verbal predicate to its subject, a specialized relational head (R) relates a non-verbal predicate to its subject. In either case, the head

that establishes the predication relation functions as the RELATOR, i.e., the functional head that relates the subject and predicate.

Under this approach, a verbal predicate is related to the subject via an existing functional head — either T/Infl, or  $v$  depending on whether it selects for an external argument or not (Chomsky 1986, den Dikken 2006). Assuming that the selectional properties of functional heads like T/Infl and  $v$  stipulate that they select a verbal complement, said heads cannot also serve as the RELATOR with non-verbal predicates; non-verbal predicates are simply incompatible with the higher functional heads that relate verbal predicates to their subjects (Myler 2016, 2018). Instead, non-verbal predicates require an additional projection, RP, that satisfies the selectional requirements of T/Infl, and mediates the relation between the predicate and subject. In many languages, including English, the head of this abstract projection is realized by the copula (111c).

- (111) a. Verbal Predication (Unaccusatives): [TP [NP SUBJECT<sub>*i*</sub>] [T' T=RELATOR [VP PREDICATE *t<sub>i</sub>* ]]]
- b. Verbal Predication (Transitives): [VP [NP SUBJECT] [<sub>V</sub>'  $v$ =RELATOR [VP PREDICATE ]]]
- c. Non-verbal Predication: [RP [NP SUBJECT] [<sub>R</sub>' *be*=RELATOR [<sub>AdjP</sub> PREDICATE ]]]

The key feature of den Dikken's approach is that there are no category restrictions on the identity of the RELATOR. By allowing any functional projection in the thematic/inflectional domain, e.g., T/Infl,  $v$ , or R (*be*), to mediate the relation between a predicate and its subject, den Dikken (2006) provides a unifying analysis for verbal and non-verbal predication; though different functional heads may serve as the RELATOR in a particular construction, verbal (111a-b) and non-verbal predication (111c) are geometrically identical. In this sense, the relator approach makes the same claim about the copula as accounts assuming that the copula realizes Pred, however it doesn't assume that a dedicated predicational head is present in all types of predication, c.f., Bowers (1993, 2001).

### 3.3.2 Clause types and structure

In the previous section, I showed that accounts differ in terms of which head the copula realizes in the syntax, and how universally this holds for copulas across languages. Adding to this complexity is the fact that different types of copular clauses involve different alignments of the subject and predicate relative to the copula (112).

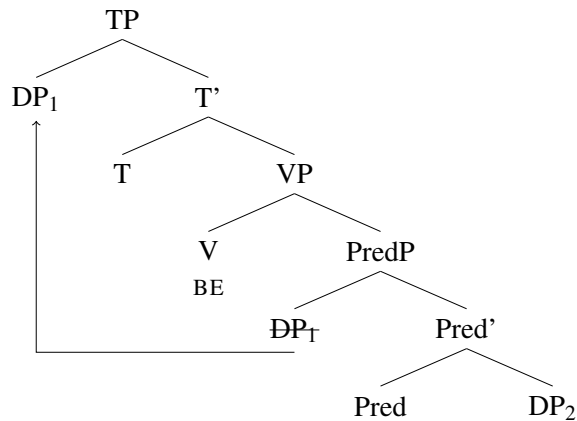
(112) *Types of copular clauses*

- |                             |                        |                   |
|-----------------------------|------------------------|-------------------|
| a. <i>Predicational:</i>    | Susan is a doctor      | SUBJECT-PREDICATE |
| b. <i>Specificational:</i>  | The winner is Susan    | PREDICATE-SUBJECT |
| c. <i>Identificational:</i> | That woman is Susan    | PREDICATE-SUBJECT |
| d. <i>Equational:</i>       | John's mother is Susan | -                 |

Setting aside equational clauses, there are two general syntactic configurations across the different types of copular clauses. Predicational clauses involve the canonical syntactic configuration of a copular clause, namely a SUBJECT-PREDICATE configuration. In contrast, specificational and identificational clauses feature a reversal of these elements, namely a PREDICATE-SUBJECT configuration. Given this structural asymmetry, one prominent question is how each structure is derived, and whether the two are derivationally related. Broadly speaking, there are two theoretical positions on this issue. On the first position, the two configurations are derivationally related (Williams 1983, Heggie 1988, Moro 1990, 1997, 2000, Heycock 1994, den Dikken 1998, 2006). As in previous models, this approach assumes the presence of a small clause headed by Pred (or something similar) that serves as the thematic core of all copular clauses; Pred relates the predicate (its complement) to the logical subject (housed in Spec, PredP). Given this uniform structure, the sole difference between predicational and specificational clauses (or any PREDICATE-SUBJECT copular configuration) lies in which element raises to Spec, TP. In predicational clauses (113), the higher nominal (i.e., the logical subject) raises as expected, however in specificational clauses (114), the lower nominal raises across the higher nominal (Schneider-Zioga 2021).<sup>6</sup>

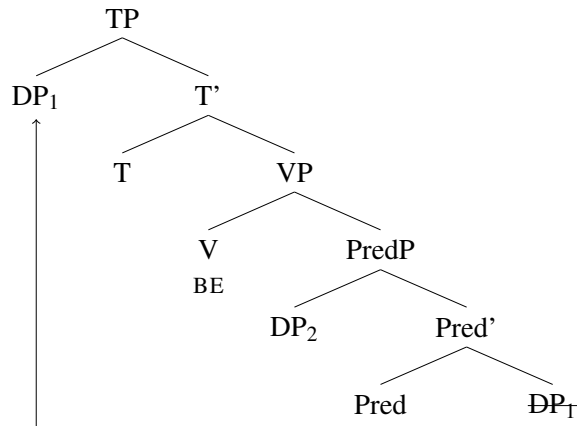
<sup>6</sup>An immediate question this analysis raises is why the lower DP is allowed to invert across the higher DP in Spec, PredP in specificational clauses, as this would presumably violate the Minimal Link Condition (Chomsky 1995). There

(113) a. *Predicational clause*



b. [<sub>DP1</sub> Susan] is [<sub>DP2</sub> a doctor]

(114) a. *Specificational clause*



b. [<sub>DP1</sub> The doctor] is [<sub>DP2</sub> Susan]

On the second account, predicational and specificational clauses do not share a uniform base structure (Partee 1986). While both still include a small clause, the logical subject is not uniformly introduced in the higher position (e.g., Spec, PredP) in both: in predicational clauses, the higher nominal is the logical subject, however in specificational clauses, it is the predicate. As such, this account posits that “inverse” order of elements arises because the predicate is base-generated in the subject position of the small clause in specificational sentences (Partee 1986).

Though it remains an open question as to which of these approaches is ultimately the better one, are various solutions that have been proposed for this issue, for example, assuming that the higher DP (i.e., the logical subject) is syntactically “frozen” in some way (den Dikken 2006), or that there is another functional head above PredP that enters into an Agree relation with the higher DP, rendering it inaccessible to a higher probe on T (Schneider-Zioga 2021). I will explore these analyses later in this section.



there are considerable arguments in favor of the derivational approach. Perhaps the most compelling argument in favor of a derivational approach — or rather, against a base-generated approach — is that presented in den Dikken (2006). For den Dikken (2006), a base-generated approach predicts a structure that corresponds to what he calls PREDICATE-SPECIFIER constructions, i.e., constructions in which the predicate is base-generated in the specifier position of the relator phrase. Consider the examples in (115), where *for* and *as* realize the relator head.

(115) *Predicate-specifier constructions*

- a. Brian is [RP [AP clever] [RELATOR=*for* [DP a five-year-old]]
- b. Brian is [RP [DP a real disgrace] [RELATOR=*as/for* [DP a heavyweight champion]]

(den Dikken 2006: 84)

Despite their similarities, den Dikken (2006) argues that PREDICATE-SPECIFIER constructions like those above fundamentally differ from so-called Copular (or more broadly, Predicate) Inversion constructions like those in (116).

(116) *Copular Inversion constructions*

- a. *Uninverted:* Brian is the best candidate
- b. *Inverted:* The best candidate is Brian

(den Dikken 2006: 81)

Specifically, he argues that unlike PREDICATE-SPECIFIER constructions, the logical subject in Copular Inversion constructions is syntactically “frozen”. One way that this manifests is that post-copular elements in Copular Inversion constructions are unable to project focus up the tree through focus projection (Selkirk 1995). In PREDICATE-SPECIFIER constructions, sentence-final focal stress (indicated by small caps) can project to higher nodes (indicated by brackets), resulting in a larger focused constituent (117).

(117) Imogen said Brian is clever for a FIVE-YEAR-OLD

- a. She did not say he’s clever for [A NINE-YEAR-OLD]
- b. She did not say he’s [ugly for A CENTERFOLD]

- c. She did not say he [smelled like A MARIGOLD]
- d. She did not say [the United States has the world in A STRANGLEHOLD]

(den Dikken 2006: 85)

Similarly, *uninverted* copular constructions also permit focus projection: sentence-final focal stress can percolate up the tree to place larger constituents in focus (118).

(118) Imogen considers Brian to be the best CANDIDATE

- a. She does not consider him to be the best [MAGISTRATE]
- b. She does not consider him to be the [DEVIL INCARNATE]
- c. She does not consider him to [have the brains to GRADUATE]
- d. She does not consider [the country to need another WATERGATE]
- e. She does not [think that the test results would be easy to REPLICATE]

(den Dikken 2006: 83)

However, Predicate Inversion constructions — including Copular Inversion — do not allow focus projection in the same way. Instead, inversion obligatorily yields narrow focus on the post-copular element alone (119). No element larger than the post-copular (or post-relator) element can be in focus.

(119) Imogen considers the best candidate to be BRIAN

- a. She does not consider the best candidate to be [RYAN]
- b. \*She does not consider the best candidate to [have been LYING]
- c. \*She does not consider [this pilot to be particularly good at FLYING]
- d. \*She does not [think those allegations about her sex life are worth DENYING]

(den Dikken 2006: 83)

As such, there is a clear asymmetry between uninverted copular clauses /PREDICATE-SPECIFIER constructions on the one hand, and inverted copular clauses on the other. Importantly for our

discussion, this asymmetry provides strong evidence against a base-generation approach to specificational/inverted copular clauses. The rationale here is that if the predicate is base-generated in the subject position of the small clause in inverted copular clauses, we would expect such constructions to pattern with uninverted copular clauses and PREDICATE-SPECIFIER constructions since they would feature the same basic A-movements: the element base-generated in Spec, PredP (or Spec, RP for den Dikken 2006) moves to Spec, TP. However inverted copular clauses differ from uninverted clauses and PREDICATE-SPECIFIER constructions in that they do not support focus projection, suggesting that they may not be structurally identical constructions. Den Dikken (2006) further shows that this dichotomy is supported by similarly asymmetric A'-extraction facts: while inverse copular clauses ban extraction of the post-copular element, PREDICATE-SPECIFIER constructions allow extraction of post-relator elements. On the basis of these observations, den Dikken (2006) concludes that copular inversion must be derived via movement.

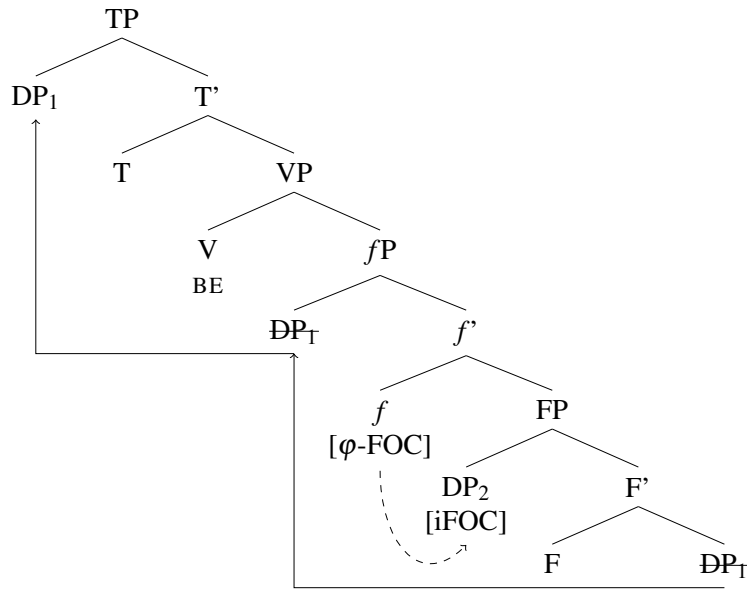
Of course, even if we assume a derivational approach to specificational clauses, there are still many questions to be addressed. For example, one salient question is why inversion happens at all; why does the predicate move across the subject in specificational clauses? One possible approach to this question is to assume that this is a product of Agree: a  $\varphi$ -probe on T can “find” two different goals (Hartmann and Heycock 2016, 2018, Béjar and Kahnemuyipour 2017, Schneider-Zioga 2021). In a standard configuration, the  $\varphi$ -probe on T finds the logical subject DP (e.g., DP<sub>S</sub>) in Spec, PredP, as this is structurally superior to the predicate DP (120a). However, it is also possible that the predicate DP (e.g., DP<sub>P</sub>) could adjoin to PredP, making it the highest DP and the first to be found by the  $\varphi$ -probe (120b).

- (120) a. [ $\varphi$ -T ... [<sub>PredP</sub> DP<sub>S</sub> [ Pred [ DP<sub>P</sub> ] ] ]  
           └──────────────────────────┘  
       b. [ $\varphi$ -T ... DP<sub>P</sub> [<sub>PredP</sub> DP<sub>S</sub> [ Pred [ DP<sub>P</sub> ] ] ]  
           └──────────────────┘

As an alternative to this approach, Schneider-Zioga (2021) proposes that specificational copular clauses are structurally distinct from predicational copular clauses in one specific way: in specificational clauses, the complement of the copula is a focus projection *fP* that contains the predicational

small clause, FP. Following Heycock (2012), Schneider-Zioga (2021) assumes that specificational clauses involve a functional projection F as the head of the small clause, which encodes a unique specificational semantics.<sup>7</sup>

(121) *Specificational clause*



(Adapted from Schneider-Zioga 2021)

The reason the structurally lower DP ( $DP_1$ ) inverts across the logical subject is because of the presence of the focus head  $f$ . The focus head is endowed with an uninterpretable  $\phi$ -FOC feature that must be valued by entering into an Agree relation with an element featuring an interpretable focus feature and interpretable  $\phi$ -features. For Schneider-Zioga (2021), this results in an application of downward Agree; the  $\phi$ -FOC probes downward to find the closest goal that will value it. If  $fP$  encases the small clause FP,  $\phi$ -FOC will necessarily find  $DP_2$  (i.e., the logical subject) first, and therefore place it in focus. Since  $DP_2$  has entered into an Agree relation with  $\phi$ , this leaves only the predicate nominal  $DP_1$  to value the higher probe on T, leading to inversion.<sup>8</sup>

<sup>7</sup>Following Romero (2005), Heycock (2012), and Arregi et al. (2021), Schneider-Zioga (2021) takes the pre-copular DP in specificational clauses to be an individual concept rather than a predicate. Specifically, Schneider-Zioga (2021) adopts the syntax presented in Heycock (2012), where specificational clauses are assumed to involve a relational head F (rather than Pred) which takes an individual concept complement.

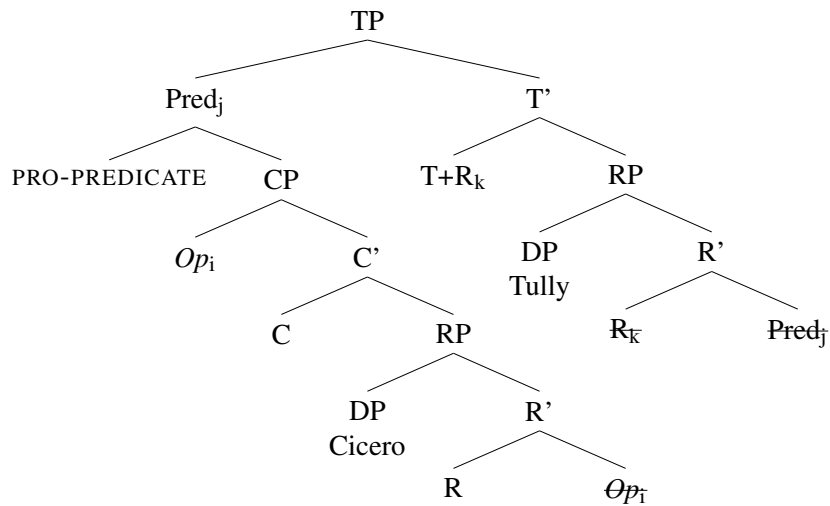
<sup>8</sup>To account for two different patterns of agreement in Kinande, Schneider-Zioga (2021) argues that  $DP_1$  moves through Spec,  $fP$  first before moving to Spec, TP. An alternate analysis where  $DP_1$  moves directly to Spec, TP is also explored, but ultimately rejected (see Schneider-Zioga 2021 for more).

Den Dikken (2006) presents a different analysis of inversion that is driven by the licensing requirements of the post-copular element. He argues that equational clauses like that in (122) obligatorily involve inversion, as they pattern like inverse copular sentences with respect to A'-extraction restrictions and other inversion diagnostics.

(122) Cicero is Tully (den Dikken 2006: 92)

Assuming that equational clauses necessarily involve inversion, den Dikken (2006) presents an analysis in which the surface subject, e.g., *Cicero*, inverts across the logical subject because it is contained within a reduced free-relative that raises to Spec, TP to be licensed (123). This free-relative, e.g., [<sub>Pred</sub> PRO-PREDICATE [<sub>CP</sub> *op* [<sub>C</sub> [<sub>RP</sub> *Cicero* [RELATOR *t<sub>i</sub>* ]]]]] meaning “what Cicero is”, raises to Spec, TP to be licensed in the same way as *pro* in Rizzi (1986).

(123)



(den Dikken 2006: 73)

In addition to providing an account of inversion, den Dikken’s analysis also accounts for an issue briefly mentioned earlier in this section: equational clauses seem to not distinguish a subject and predicate. As briefly mentioned in section 3.3.1, the apparent reversibility of arguments in equational clauses, e.g., *John is the professor/The professor is John* (Roy 2013), is the central argument for the existence of “bare” small clauses that relate the subject and predicate symmetrically (Moro 1995, 1997). For den Dikken (2006) however, equational clauses are assumed to have the same hierarchical structure that serves to distinguish the subject and predicate as other copular clauses. The only

difference between equational clauses and other types of copular clauses in his analysis is that equational clauses involve a more complex predicate (i.e., they involve a reduced free-relative). This complexity results in the movement of the predicate for licensing reasons, meaning that equational clauses exhibit a similar pattern of inversion to specificational (and identificational) clauses.

To summarize this section, there is significant debate as to the structural similarity between different types of copular clauses. The central debate in the literature focuses on the relationship — or lack thereof — between predicational and specificational copular clauses. One approach takes specificational copular clauses to be directly derived from predicational copular clauses; predicate nominals invert across the logical subject in specificational clauses. Though analyses of this type differ as to how exactly they account for this inversion, most suggest that the post-copular nominal is in some way “frozen” or inaccessible to an active probe on T, resulting in the structurally lower nominal raising — or rather “inverting” — across the higher (logical subject) nominal. In contrast, an alternative approach takes specificational copular clauses to be structurally distinct from their predicational counterparts. Under this approach, predicate nominals are base-generated higher than the logical subject in specificational clauses, and therefore do not undergo “inversion”.

### **3.3.3 Interpretation and structure**

Transitioning from our discussion of structure as it pertains to copular clause type, there are also structural accounts of other interpretive contrasts in copular systems, e.g., stage vs individual/temporary vs permanent. Consider for example the constructivist view of predication advocated for in accounts like Adger and Ramchand (2003), Baker (2003), Balusu (2014), Myler (2016, 2018). As detailed in section 3.3.1, these accounts posit that the subject and predicate of copular clauses are initially related to each other in a small clause (i.e., PredP) that is then selected by  $v_{BE}$  — the head realized by the copula. Like alternative accounts which suggest that the copula realizes Pred (rather than  $v_{BE}$ ), the constructivist approach has a built in means of accounting for interpretive contrasts in copular systems; interpretive contrasts arise due to the existence of multiple Pred heads, e.g., Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub> (Adger and Ramchand 2003, Markman 2008a, Balusu 2014). Consider the

treatment of Spanish *estar/ser* in Myler (2016, 2018). Instead of positing two distinct copular verbs, Myler characterizes *estar/ser* as allomorphic variants:  $v_{BE}$  is realized as *estar* in the presence of  $\text{Pred}_{\text{STAGE}}$ , and as *ser* in the presence of  $\text{Pred}_{\text{INDIV}}$ .

(124) *Spanish copula VI rules* (Myler 2018: 9)

a.  $v_{BE} \leftrightarrow \text{estar} / \text{ \_\_\_\_\_\_ } \text{Pred}_{\text{STAGE}}$

b.  $v_{BE} \leftrightarrow \text{ser} / \text{ \_\_\_\_\_\_ } \text{Pred}_{\text{INDIV}}$

Importantly, by splitting the copula — which realizes  $v_{BE}$  — from  $\text{Pred}$ , this account enjoys the benefits of both the copula-as- $\text{Pred}$  approach and the copula-as- $v$  approach; with both  $v_{BE}$  and  $\text{Pred}$  present, there is a straightforward explanation of interpretive contrasts that also is compatible with treating the copula as a light-verb. This analysis also provides a nice account of the relationship between BE and HAVE, something I will discuss more in section 3.3.4.

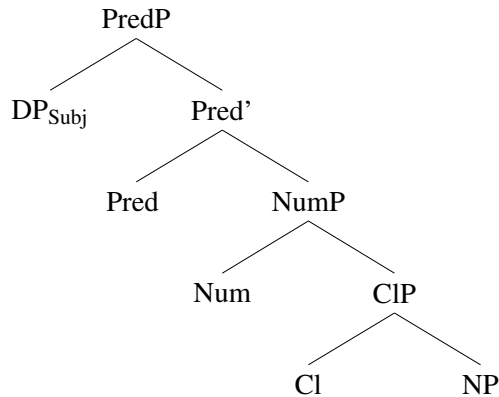
Unlike other competing analyses that rely on lexical semantics or discrete functional heads to derive interpretive contrasts, Roy (2013) attributes this to structural differences among predicates. Based on interpretive and distributional contrasts in French, Spanish and Russian, Roy specifically posits three syntactically distinct types of non-verbal predicates that each correspond with a unique interpretation (and sometimes a unique morphosyntactic realization).

So-called *situation-descriptive* predicates are structurally defined as bare predicate XPs, and are associated with *dense* interpretations; they describe properties that hold of a particular situation. On the other hand, *characterizing* predicates involve an additional projection, a Classifier Phrase (CIP), and correspond to *non-dense* interpretations; they describe properties that hold of an entire eventuality, but not necessarily of all its sub-parts. Finally, *maximal* predicates are associated with the presence of yet a higher functional projection, the Number Phrase (NumP), and are associated with *defining* interpretations; they describe a property that holds of a maximal eventuality.

(125) *Predicate types*

- a. [XP] dense (situation-descriptive)
- b. [CIP[ ... ]] non-dense (characterizing)
- c. [NumP[ ... ]] maximal (defining)

(126) *Structure of nominal predicates*



(Adapted from Roy 2013)

As discussed in section 2.3.1, this three-way contrast is overtly expressed in French via the presence or absence of the indefinite article *un(e)*, which Roy (2013) argues heads the NumP projection. Given the structural distinctions in (198), only maximal predicates surface with an indefinite article (127). The two structurally smaller predicates differ minimally in whether or not they project a classifier phrase (CIP); of the two, only non-dense predicates include this additional functional projection. While unpronounced in French, the function of the classifier head is to divide mass terms into countable units — it turns dense predicates into non-dense ones by creating divisions of a mass term (Roy 2013).

The presence of the classifier head is what derives the interpretive contrast between non-dense and dense predicates. Crucially, non-dense predicates allow for ‘interpretive gaps’ in an eventuality. Unlike the maximal predicate *un ivrogne* in (127), which describes a defining property of *Paul*, the property denoted by the non-dense predicate *ivrogne* in (128) does not entail his being a drunkard *continually* — the utterance in (128) would be felicitous even if Paul were in a prolonged period of sobriety, but the utterance in (127) would not be. In contrast, the property denoted by the dense predicate *ivre* in (128) holds strictly of a particular situation, e.g., the situation surrounding the



utterance time. In this sense, the absence of the classifier phrase with dense predicates entails that the eventuality described by the predicate is interpreted as a mass eventuality — it doesn't allow for interpretive gaps.

(127) *Maximal (defining) — associated with presence of NumP*

Paul est un ivrogne  
Paul is a drunkard  
'Paul is a drunkard'

(128) *Non-dense (characterizing) — associated with presence of ClP*

Paul est ivrogne  
Paul is drunkard  
'Paul is a drunkard'

(129) *Dense (situation-descriptive) — no additional projections*

Paul est ivre  
Paul is drunk  
'Paul is drunk'

French (adapted from Roy 2013)

Below is a rough sketch of the semantics of the three predicate types as they are described in Roy (2013).<sup>9</sup>

(130) *Maximal (defining) predicates*

- a.  $\exists e[\text{MAX}(e) \ \& \ P(e) \ \& \ \text{SUBJ}(e,x)]$
- b. There is an event of x being P, and for all e P(e), there is no e' such that e is a proper part of e' and P(e').

(131) *Non-dense (characterizing) predicates*

- a.  $\exists e[P(e) \ \& \ \text{SUBJ}(e,x) \ \& \ P \text{ is nondivisible}]$
- b. There is an event of x being P where P is nondivisible (P need not hold of all subparts of e).

---

<sup>9</sup>Note that the SUBJ function is a cover for any thematic relation. As such, it may be interchangeable with a specific thematic function, e.g., PATIENT, in a particular context.

(132) *Dense (situation-descriptive) predicates*

- a.  $\exists e[\text{P}(e) \ \& \ \text{SUBJ}(e,x)]$
- b. There is an event of x being P.

As with many of the other issues discussed in this chapter, it is an open debate as to whether interpretive contrasts in copular systems can be attributed to structure, and if so, where those interpretive contrasts are situated. For example, accounts like Adger (2003), Markman (2008a) and Balusu (2014) take interpretive contrasts to reflect the existence of two distinct Pred heads. However for Roy (2013), interpretive contrasts are situated lower than Pred, and directly relate to the structure of non-verbal predicates.

### 3.3.4 Deriving HAVE from BE

Another topic related to the structure of copular clauses and non-verbal predication is the structure of clausal possession. As discussed in section 2.2.7, non-verbal predication and clausal possession are sometimes expressed in similar ways in some languages. In some cases, this manifests in the form of a possessive *have*-verb being used to establish a predication relation, e.g., French *avoir* (15). In other cases, this instead manifests as a copula being used to express a possessive relation, usually alongside a locative element, e.g., Rangi *-ri + na* (37). In either case, the crucial observation is that non-verbal predication and possession — or alternatively, abstract BE and HAVE — appear to be related in some way.

- (15) Pierre **a**                      faim  
      Pierre have.PRES.3SG hunger  
      ‘Pierre is hungry’

French (Francez and Koontz-Garboden 2017: 22)

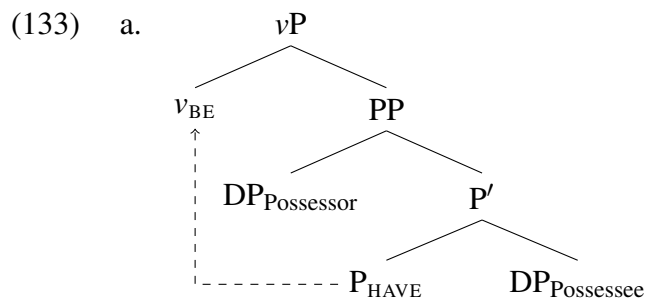
- (37) **A-ri**    **na**    vijombulo vikuulu  
      1SM-be COM 8.calves   8.AGR.big  
      ‘S/he has big calves’

Rangi (Gibson 2012: 101)

In light of this overlap, many have attempted to identify a clearly-defined “link” between BE and HAVE, particularly with respect to the expression of clausal possession. That is, in an effort

to account for the crosslinguistic tendency for languages to either encode possession via an overt lexical *have*-verb, or the combination of copular BE + locative, many accounts have advocated for a compositional relationship between abstract HAVE and BE whereby HAVE is directly “built” from BE (Freeze 1992, Kayne 1993, Ritter and Rosen 1997, Harley 1998, Beavers et al. 2009, Myler 2016, 2018 a.o.). Such decompositional approaches to possessive HAVE generally fall into two theoretical camps in terms of how they relate HAVE and BE. The first approach (P-incorporation) assumes that HAVE is derived via the incorporation of a locative element by BE (Freeze 1992, Kayne 1993, Harley 1998). In contrast, the second approach assumes that HAVE is derived via contextual allomorphy; HAVE is the transitive allomorph of BE (Myler 2016, 2018). Let us briefly walk through each of these analyses in turn.

Accounts like Freeze (1992), Kayne (1993), and Harley (1998) suggest that HAVE is underlyingly composed of two verbal heads:  $v_{BE}$  (which realizes abstract BE) and a functional, locative element  $P_{HAVE}$  that serves to relate two DPs in a possessive relationship. The core idea of these accounts is that HAVE is realized through the incorporation of a lower functional element  $P_{HAVE}$  by a stative verbal head  $v_{BE}$  (133).



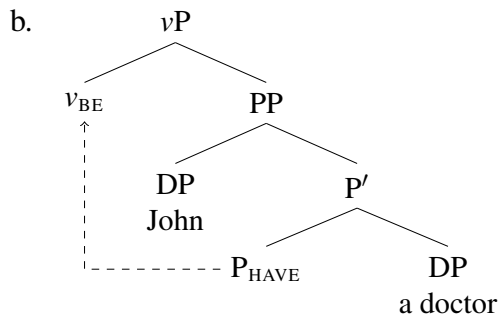
b. [  $v_{BE}$   $P_{HAVE}$  ]  $\rightarrow$  /have/

Under this genre of analysis, the sole difference between non-verbal, copular predication and clausal possession is the functional projection that merges with  $v_{BE}$ . In a possessive structure,  $v_{BE}$  takes a PP complement headed by  $P_{HAVE}$ , and then incorporates  $P_{HAVE}$  (134). Assuming a PredP analysis, (Adger and Ramchand 2003, Markman 2008a, Myler 2016, 2018, a.o.),  $v_{BE}$  instead takes a PredP complement in non-verbal predications like (135). Like  $P_{HAVE}$ , the Pred head — which

houses interpretive contrasts (section 3.3.3) — is incorporated by  $v_{BE}$ .<sup>10</sup>

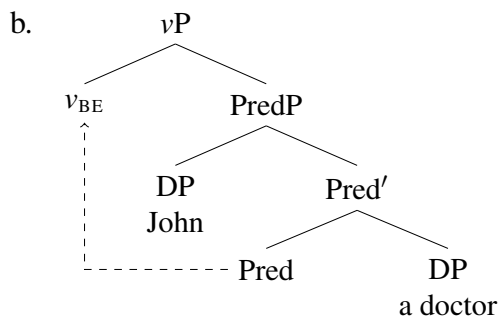
(134) *Possessive Configuration*

a. John has a doctor



(135) *Non-verbal Predication*

a. John is a doctor



An alternative approach to the parallelism between non-verbal predication and possession is to take HAVE as a contextually conditioned allomorph of  $v_{BE}$  (Myler 2016, 2018). Specifically, Myler (2016, 2018) describes HAVE as an allomorph of  $v_{BE}$  that appears in the context of a transitive Voice head.<sup>11</sup>

(136) a. [  $v_{BE}$  ]  $\rightarrow$  HAVE / Voice<sub>D</sub> \_\_\_\_\_

b. [  $v_{BE}$  ]  $\rightarrow$  BE

(Myler 2018: 11)

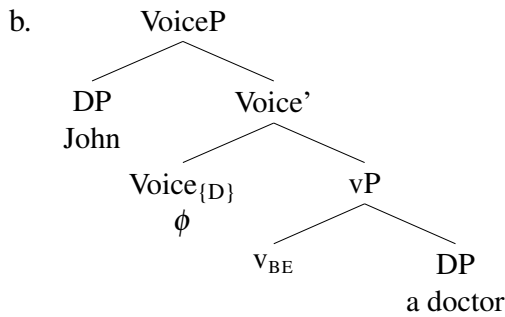
<sup>10</sup>To account for the similar interpretive contrasts that hold of the *be*- and *have*-verbs in Mashi (JD53), Finholt (2022, 2024) argues that a PredP projection is also present in possessive structures. In this case, HAVE is taken to be the realization of  $v_{BE} + \text{Pred} + P_{\text{HAVE}}$ .

<sup>11</sup>See Myler (2016) for discussion of HAVE as a transitive verb.

Under this approach, clausal possession differs structurally from non-verbal predication. If possessive HAVE is transitive, then the possessor DP (the subject) must be introduced in the specifier of a transitive Voice head (137) — a position where it can be assigned a theta role (Myler 2016). In non-verbal predication however, the subject DP is not introduced in Spec, VoiceP, but rather Spec, PredP (138).

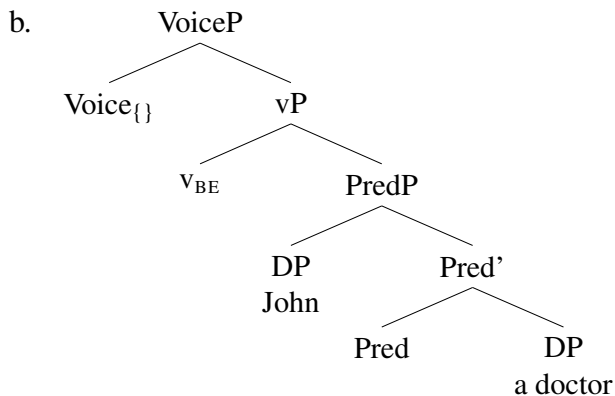
(137) *Possessive Configuration*

a. John has a doctor



(138) *Non-verbal Predication*

a. John is a doctor



(adapted from Myler 2016, 2018)

In contrast, the P-incorporation approach discussed previously assumes that the subject is base-generated low (i.e. below  $v_{BE}$ ) in clausal possession and non-verbal predication, e.g., in Spec, PP or Spec, PredP.

### **3.4 Summary**

In this chapter, I outlined some of the prominent theoretical approaches to non-verbal predication discussed in the literature. As I have shown, there has been significant theoretical debate about the meaning and structure of copulas and copular clauses. With respect to meaning, accounts differ as to whether the copula is truly semantically vacuous; for some, the copula is always a meaningless predicational element, while for others it may encode some “meaning”, e.g., copulas may serve an equative function or lexically encode interpretive contrasts. As for the structural representation of copulas and copular clauses, accounts again differ quite strongly. With respect to the realization of the copula, there is disagreement as to what syntactic head(s) are realized by the copula, and how copular contrasts are derived. With respect to the structure of copular clauses more broadly, accounts differ as to whether predicational, specificational, and equational clauses are derivationally related, and if they are, how this should be represented structurally.

## Chapter 4

### Interpretation

#### 4.1 Introduction

One topic of considerable debate in the literature on copular clauses is the semantic contribution of the copula. As briefly explored in section 3.2, there is a stark divide between those who treat the copula as a meaningless predicative element, e.g., Pustet (2003), and those who take the copula to be more complex, e.g., Clements (1988), Maienborn (2005). Much of the work on this topic has focused on the meaning of the copula across the various types of copular clauses presented in Higgins (1979), i.e., whether a single meaningless copula can account for predicational, specificational, equational, and identificational clauses. However, there has also been considerable work focused on the meaning of the copula in standard predicational clauses, particularly in languages where multiple copular forms are employed. In many of these languages, it has been reported that those distinct forms correspond with distinct interpretations of the predication relation, e.g., the Spanish copulas *estar/ser* encode a distinction between stage/individual-predication (Milsark 1974, Carlson 1977, Kratzer 1995, a.o.). For many, these contrasts present a credible challenge to the idea that the copula is universally meaningless, though such an analysis is not fully incompatible with these facts (see for example Adger 2003). In any case, copular contrasts like that in Spanish provide a window into the overall interpretive profile of the copula, allowing us to begin to identify the kinds of meanings copulas can and cannot express cross-linguistically.

In this chapter, I explore the question of copular meaning through a narrow investigation of copular contrasts in a small group of Great Lakes Bantu languages, namely Kihavu (JD52), Mashi

(JD53), Kinyamulenge (JD61a), Kirundi (JD62), and Kifuliiru (JD63).<sup>1</sup> These five languages are ideally suited for an investigation of copular meaning since they each employ multiple copular forms (a set of 2-3) in predicational clauses that are interpretively distinct. In the sections to come, I illustrate the general interpretive contrasts observed in each of the languages, and compare them with previous analyses of copular contrasts in other languages. Given the close genealogical relationship between the languages, and the similarity of their copular forms and contrasts, I present a generalizable semantic analysis for the copular forms attested across the six languages. At its core, the proposed analysis ascribes a uniquely Austinian semantics to the copula; the copula “links” the truth of the proposition to a particular situation or circumstance à la Maienborn (2005), Deo et al. (2017). The difference between the relevant copular forms in the Bantu languages under discussion is roughly the “size” of that situation/circumstance, i.e., limited vs non-limited situations/circumstances.

## 4.2 Utterances and situations

The idea that utterances are about particular situations has long been attributed to Austin (1950). To illustrate, consider the story of the butler and the judge from Kratzer (1998).

The judge was in financial trouble. He told his butler that he had been ready to commit suicide, when a wealthy man, who chose to remain anonymous, offered to pay off his debts. The butler suspected that Milford was the man who saved his master’s life by protecting him from financial ruin and suicide. While the butler was away on a short vacation, the judge fell into a ditch, drunk. Unconscious and close to death, he was pulled out by a stranger and taken to the local hospital, where he recovered. When the butler returned to the village, he ran into a group of women who were speculating about

---

<sup>1</sup>In my discussion of Kinyamulenge and Kirundi, I also discuss data from closely-related Kinyarwanda (JD61), which exhibits an almost identical copular system to these languages. Importantly, the majority of Kinyarwanda data is taken from Jerro (2015). I do not list Kinyarwanda as one of the primary languages of focus in this study since I do not present novel data from the language. I will nonetheless discuss the Kinyarwanda data alongside the Kinyamulenge and Kirundi data for illustrative purposes.



the identity of the stranger who saved the judge's life by taking him to the hospital. One of the women said she thought that Milford saved the judge's life. The butler, who hadn't yet heard about the accident and thought the women were talking about the judge's financial traumas, reacted with (139) (Kratzer 2023)

(139) I agree. I, too, suspect that Milford saved the judge's life.

Given the situation under discussion, the group of women in this story would logically interpret the butler's suspicion to be about the identity of the stranger who pulled the judge from the ditch. However, since the butler was at this point unaware that the judge had fallen into a ditch in the first place, the utterance in (139) could not possibly be about the situation in which the judge was rescued from the ditch. Instead we can safely assume that this utterance is about the only "life-saving" situation the butler is aware of: the situation in which the judge's debts were paid off by an anonymous man. In this sense, the butler and the group of women have different situations in mind when speculating about the identify of the "stranger" who saved the judge; for the butler, the relevant situation is that in which the judge was saved from financial ruin, while for the group of women, it is that in which the judge was rescued from the ditch.

This mismatch becomes particularly apparent if we consider the belief ascription in (140), which is infelicitous if uttered by someone in regard to the judge's hospitalization.

(140) The butler suspects that Milford saved the judge's life.

As reported by Kratzer (2023), there is a strong sense that (140) is infelicitous if uttered in discussion of the judge's hospitalization, but felicitous if uttered in discussion of the judge's debts being paid off. What this shows is that the butler's utterance in (139) and the belief ascription in (140) — like all utterances — are about particular situations. As discussed above, the butler's utterance in (139) is ostensibly about the situation in which the judge was saved from debt, since he is unaware that the judge was also saved from a ditch. It follows that (140) is infelicitous when uttered in discussion of the judge's hospitalization because it makes a claim about a different situation than (139).

### 4.2.1 Situation semantics

Situation semantics is the formal approach that was developed from the Austinian mold as an alternative to possible world semantics (Kratzer 2023). It analyzes utterances in terms of three types of situations: topic situations, utterance situations, and resource situations (Devlin 2006).

The topic or “focal” situation  $s_t$  is the particular part of the world that the utterance is about (Devlin 2006, Kratzer 2023). Often referred to as Austinian topic situations, these situations directly correspond to the “situations” of Austin (1950). In the story of the butler and the judge discussed above, there are two topic situations at play: one where an anonymous man saved the judge from debt, and another where an anonymous man saved the judge from a ditch. Unaware of the latter situation, the butler speculates about the situation in which the judge was saved from debt — this is the topic situation of the utterance in (139). Likewise, the topic situation in (141) is the situation of a man being at some contextually-specified door, e.g., the speaker’s front door in a context where they look out the window and see a man at the front door.

(141) A man is at the door (Devlin 2006: 610)

The utterance situation  $s_u$  is the immediate situation in which an utterance is made (Devlin 2006). The utterance situation necessarily includes the speaker and addressee, as well as other information needed to interpret the utterance and the objects it refers to. In general, the utterance situation contains all of the information needed to identify the objects referred to in the utterance. For example, the object referred to by *the door* in (141) could be identified by the location and/or environment in which the utterance was made, e.g., there is a single door in the room in which the utterance was made. Alternatively, the reference of *the door* could be identified by previously established discourse in  $s_u$ , e.g., there is a salient door under discussion.

In some cases, a speaker may make use of other situations aside from the utterance situation in order to identify objects included in the utterance. In (142), the reference of *the man* is not determined by the utterance situation, but rather a separate situation that the speaker witnessed a day earlier, e.g., a situation of a particular man running. This kind of situation is defined as a resource

situation  $s_r$ .

(142) The man I saw running yesterday is at the door (Devlin 2006: 611)

#### 4.2.2 Pragmatic approaches to copular contrasts

One particularly relevant application of this idea that utterances are about particular situations can be found in work on Spanish *estar/ser* contrast, particularly Clements (1988, 2006), Maienborn (2005), and Deo et al. (2017). The traditional view of the *estar/ser* contrast assumes a binary distinction between stage- and individual-level predication; *estar* co-occurs with stage-level predicates/readings, and *ser* co-occurs with individual-level predicates/readings (Kratzer 1995, Diesing 1992a, Arche 2006, a.o.). That being said however, it has long been observed that the distribution of *estar* and *ser* does not perfectly align with the stage/individual distinction. Going back as far as Ramsey (1894), it has been reported that certain predicates, e.g., *sucio* ‘dirty’, (143), may co-occur with either *estar* or *ser*. In these cases the choice of copula restricts the interpretation of the predicate; *estar* yields a transient reading of ‘dirty’ in (143a), while *ser* yields a more characteristic, individual-like reading (143b).

(143) a. El reportero **está** sucio  
The reporter *estar*.PRS.3SG dirty  
‘The reporter is dirty’ (they need a shower)

b. El reportero **es** sucio  
The reporter *ser*.PRS.3SG dirty  
‘The reporter is dirty-minded/corrupt’

Spanish (Ramsey 1894: 313)

While the interpretative difference between the two copulas is similar to a stage/individual contrast, the distributional difference between them is not; if *estar/ser* distinguish stage- vs individual-level predication, they should not co-occur with the same predicates.<sup>2</sup> In an effort to account for this divide between the interpretation and distribution of the two copulas, accounts like Clements

---

<sup>2</sup>This assumes that predicates are categorized as either stage- or individual-level predicates. Assuming this divide, there is no explanation for why a given predicate would allow both a stage- and individual-level copula. Unless we permit two homophonous predicates, e.g., one stage-level and one individual-level, there is no way to account for the distribution of *estar/ser*.

(1988, 2006) and Maienborn (2005) propose a pragmatic solution to *estar/ser* contrast; unlike *ser*, *estar* presupposes a “link” to a state or particular situation.

For Clements (1988), the central difference between *estar* and *ser* is that the former describes properties that hold of the subject in a particular, limited situation, e.g., the interpretation of ‘dirty’ in (143a). Clements models this difference presuppositionally; *estar* presupposes a connection to a previous situation where  $P(x) = 0$ , while *ser* does not. Formally, Clements represents this presupposition using the feature [Nexus]: *estar* is [+Nexus] while *ser* is [-Nexus]. To illustrate, consider the use of each copula with the gradable predicate *alto* ‘tall’ below (as discussed in Deo et al. 2017: 18). With *estar* ([+Nexus]) in (144a), the speaker ascribes “tallness” to the subject Pepe by comparing his current height with a preconceived expectation or standard for his height; compared to his usual/previous height, Pepe is quite tall. In contrast, the use of *ser* ([-Nexus]) in (144b) makes no reference to any alternative state of the subject, but rather defines “tallness” according to some general contextual-standard that applies to individuals in Pepe’s demographic.

- (144) a. Pepe, ¡qué alto **estás!**  
 Pepe that tall.MASC estar.PRS.2SG  
 ‘Pepe, how tall you are!’
- b. Para su edad, Pepe **es** bien alto  
 For his age Pepe ser.PRS.3SG quite tall.MASC  
 ‘For his age, Pepe is quite tall’
- Spanish (Clements 1988: 788)

Building on Clements’ idea that *estar* establishes a “link” to a previous state, Maienborn (2005) proposes that *estar* presupposes that the relevant property attribution is limited to a specific discourse situation. In Maienborn’s analysis, both *estar* and *ser* introduce a stative argument  $z$  — which she refers to as a “Kimian state” — that represents the attribution of a property  $P$  to the subject. The crucial difference with *estar* is that it involves an additional presupposition, e.g.,  $[s_i | R(z, s_i)]$ , that links a specific discourse situation  $s_i$  to the stative argument  $z$  through a contextually determined relation  $R$ .<sup>3</sup>

<sup>3</sup>Formally, the stative argument  $z$  — a “Kimian state” — corresponds to the ascription of a property  $P$  to the subject  $x$  (hence  $z \approx [P(x)]$ ). The relational function  $R$  included in the presupposition in (145b) serves to “link” the stative argument to a particular situation  $s_i$ , i.e., it expresses that the stative eventuality  $z$  holds specifically of  $s_i$ .

- (145) a.  $[[\text{ser}]] = \lambda P \lambda x \lambda z. [z \approx [P(x)]]$   
 b.  $[[\text{estar}]] = \lambda P \lambda x \lambda z. [z \approx [P(x)]/[s_i | R(z, s_i)]]$  (Maienborn 2005: 168)

One of the key aspects of this analysis is that the additional presupposition gives *estar* a strong pragmatic effect. When a speaker uses *estar* instead of *ser*, they signal that the topic situation (i.e., the Kimian state  $z$ , which constitutes the attribution of some property to the subject) is limited to a particular discourse situation  $s_i$ . This in turn generates an implicature that there are alternative topic situations in which the Kimian state does not hold, i.e., situations where  $P(x) = 0$ .

Maienborn (2005: 172) defines three dimensions along which such alternatives can be distinguished. Alternatives can be defined *temporally*, in which case the current topic situation  $s_i$  contrasts with a past/future topic situation  $s_i'$  where the predication relation need not hold. This gives rise to a temporary/transient interpretations. Alternatives can also be defined *spatially*, in which case  $s_i$  contrasts with a spatially distinct topic situation  $s_i'$  where the predication relation need not hold. This gives rise to localized interpretations. Finally, alternatives can also be defined *epistemically*, in which case  $s_i$  contrasts with a distinct topic situation  $s_i'$  where the predication relation cannot be evaluated as true, e.g., there is not evidence to suggest that  $P(x) = 1$ .

To illustrate, consider the use of *estar* in (146). In this example (as presented in Deo et al. 2017: 19), *estar* signals that the predication relation holds strictly of a spatially-defined topic situation; the road is wide *here*, but not necessarily elsewhere. By using *estar*, there is an implicature that there are other spatially-defined (alternative) topic situations where the road is not wide.

(146) *Context: A journalist is reporting on the Pan-American highway and she is now near Lima.*

La carretera **está** ancha  
 The road estar.PRS.3SG wide  
 'The road is wide'

Spanish (Maienborn 2005: 171)

Yet another approach that attributes the *estar/ser* to pragmatic is that presented in Deo et al. (2017). Following Clements (1988, 2006) and Maienborn (2005), Deo et al. similarly propose a presuppositional account of the *estar/ser* contrast that captures the intuition that *estar* sentences are restricted to specific discourse contexts. In contrast to previous analyses however, they represent this

via a “boundedness” presupposition; *estar* presupposes that the predication relation is “boundedly” true at a particular circumstance of evaluation.

As a basis for their proposal, Deo et al. (2017) identify a number of inadequacies in Maienborn’s (2005) analysis. The two most important in my view are the following. First, Maienborn does not offer a clear formalization of “specificity” as it pertains to “specific” discourse situations, nor explain how their alternatives are accessed; if the use of *estar* is dependent on the speaker’s ability to access alternative discourse situations, this should be formalized in its lexical entry. Second, the analysis provides no clear explanation of the function and/or limitations of the element R that relates the topic situation, i.e., the Kimian state, to a specific situation. As explicitly mentioned by Deo et al. (2017), there is no notion of how R relates these two items, e.g., do these situations overlap, does one contain the other, etc., nor whether R restricts the interpretation of the various elements involved in the predication relation.

In an effort to address these issues, Deo et al. propose that the context-specific interpretation of *estar* arises from a comparison of possible circumstances of evaluation. For them, *estar* presupposes that the predication relation, i.e., “the prejacent”, is “boundedly” true with respect to some circumstance of evaluation  $i$ . By using *estar* the speaker signals that there is both an accessible alternative circumstance  $i'$  where the prejacent is false, i.e.,  $P(x)(i') = 0$ , and no larger circumstance  $i''$  where the prejacent is true, i.e.,  $i < i'' \wedge P(x)(i') = 0$ .

Deo et al.’s analysis hinges on two pieces of semantic hardware: the Circ function, and the boundedness presupposition (notated **Bound**( $P(x),c,i$ )). The Circ function takes as its argument a discourse context  $c$  and returns a set of relevant alternative circumstances at which the truth of a proposition can be evaluated (Deo et al. 2017: 23). These alternatives are generated by varying the feature values of certain contextual parameters, e.g., worlds ( $w$ ), times ( $t$ ), locations ( $l$ ), agents ( $a$ ), and delineations ( $d_c$ ).<sup>4</sup> The alternative circumstances relevant to a given proposition are identical in all ways except for the truth of the prejacent and the values of their contextual parameters. Circumstances are comparatively-ranked according to strength, i.e, how close they are to the actual

---

<sup>4</sup>The delineations parameter maps gradable properties to a determined contextual standard.

world aside from the truth of the prejacent, relative to a single parameter setting.

The boundedness presupposition (e.g., **Bound**(P(x),c,i)) has two conditions:

(147) A proposition P(x) that is true at circumstance  $i \in \text{Circ}(c)$  is boundedly true at c iff:

- i.  $\exists i' \in \text{Circ}(c) : i' \not\prec_{p_c} i \wedge P(x)(i') = 0$
- ii.  $\forall i'' \in \text{Circ}(c) : P(x)(i'') = 1 \rightarrow i \succeq_{p_c} i''$

The first condition encodes that there is an alternative circumstance  $i'$  in the set generated by  $\text{Circ}(c)$  where  $P(x) = 0$ , and where  $i'$  is no weaker ( $\not\prec$ ) than  $i$  on some contextually determined parameter  $p_c$ . That is, there must be an alternative circumstance  $i'$  that is as similar to the actual world as  $i$  where the prejacent is false, and where  $i'$  differs in the value of a single parameter from  $i$ , e.g.,  $i$  and  $i'$  describe different time periods, locations, worlds, etc. To illustrate, consider the example in (148), where the set of circumstances generated by  $\text{Circ}(c)$  would contain those that differ with respect to the time parameter.

(148) Context: Pedro went on a diet for six months. I just saw him yesterday:

El	está	delgado	
He	estar.PRS.3SG	skinny	
'He is	skinny'	(now)	Spanish (Deo et al. 2017: 34)

In this example, the relevant parameter is time. The time parameter of  $i$ , the circumstance at which (148) is asserted, is  $t_0$  (the utterance time). In this case, circumstances  $i'$  involve temporal intervals that precede or include  $t_0$ , making them no weaker than  $i$  along the time parameter;  $t_{i'} \leq t_i$ . In effect, this ensures that there is an accessible alternative context where  $P(x) = 0$ ; this establishes a “link” to a previous state much like the [Nexus] feature in Clements (1988).

The second condition of the boundedness presupposition states that no circumstance of evaluation in  $\text{Circ}(c)$  that verifies  $P(x)$  is strictly stronger than  $i$  with respect to a parameter  $p_c$ . In (148), this means that there is no  $i''$  with a time frame  $t_{i''}$  such that  $t_i$  ( $t_0$ ) is a proper part of  $t_{i''}$  and  $P(x)(i'') = 1$ . This ensures that the property ascription does not extend beyond a particular circumstance; the property does not hold of a larger context of evaluation.

The lexical entries for *estar/ser* presented in Deo et al. (2017) can be seen in (149) below.

(149) a.  $\llbracket \text{ser} \rrbracket = \lambda P_{\langle s, \text{et} \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s. i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$

b.  $\llbracket \text{estar} \rrbracket = \lambda P_{\langle s, \text{et} \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s : \mathbf{Bound}(P(x), c_0, i). i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$

(Deo et al. 2017: 27)

The general idea in this analysis is that *estar* presupposes that the proposition is true in a bounded way; *estar* describes properties that hold of a circumstance that is limited by a single parameter, e.g., a specific time frame, a particular location, a contextually-defined expectation, etc.. Important to their analysis is the observation that, in addition to its temporary readings, *estar* can also describe properties that are spatially bound (150), surprising/novel (151), or that exceed contextually-determined expectations (152).

(150) **Spatially bound (location parameter)**

*Context: A journalist is reporting on the Pan-American highway and she is now near Lima.*

La carretera **está** ancha  
The road estar.PRS.3SG wide  
'The road is wide'

Spanish (Deo et al. 2017: 40)

(151) **Surprising (world parameter)**

*Context: Louisa rents an expensive farmhouse on Airbnb for a weekend and when she gets into the house, she realizes that the house is a badly maintained place from the 19th century*

¡La casa **está** vieja!  
The house estar.PRS.3SG old  
'The house is old!'

Spanish (Deo et al. 2017: 35)

(152) **Exceeds expectations (delineation parameter)**

*Context: A baker is critiquing the tiramisú just baked by her assistant, which did not turn out as it was supposed to.*

La capa del tiramisú **está** muy gruesa  
The layer of.the tiramisu estar.PRS.3SG very thick  
'The layer of the tiramisu is too thick'

Spanish (Deo et al. 2017: 36)



### 4.3 Interpretive contrasts among Great Lakes Bantu copulas

In each of the Great Lakes Bantu languages investigated in this work, there is one context where multiple copular forms exist in variation, namely present tense predicational sentences (in the sense of Higgins 1979). In this environment, each of the copular forms in a given language are associated with distinct interpretations of the predication relation. In other words, language exhibits an interpretive contrasts among their respective copulas in present tense predicational sentences.

In total, there are five morphologically distinct copular forms attested in present tense predicational sentences in the languages discussed. These include four verbal copulas *-li*, *-ri*, *-ba*, and *-tula*, and an invariant copula *ni*. The first three of these forms are widely attested among Bantu languages (Gibson et al. 2019), and are taken to be reflexes of the proto-Bantu copulas *\*-de* and *\*-bà* (Meeussen 1967, Guthrie 1967 1971, Wald 1973); *-li* and *-ri* reconstruct as *\*-de*, and *-ba* reconstructs as *\*-bà*. The invariant copular form *ni* is also commonly reported in Bantu languages, although it is reported to have a source as a focus marker rather than a copula (McWhorter 1994). The final form *-tula* has a less clear source, though it is assumed to be derived from a lexical verb.

Languages differ in terms of the particular set of forms they employ, however no single language employs more than three distinct forms. Interpretively, copular forms tend to share a core semantic profile across languages, however related forms may take on slightly different interpretations in across languages depending on the context they are used in and the forms they contrast with. I define three broad ‘types’ of interpretive contrasts that are distinguished by their component forms, e.g., the *-li* vs *-ba* contrast, the *-ri* vs *ni* contrast, and the *-li* vs *-ba* vs *-tula* contrast. Though these contrasts vary slightly in meaning, at their core is a distinction between property ascriptions that are true of limited contexts (e.g., *-li/-ri*), and those that are true of more general contexts. These contrasts are summarized below in Table 4.1.

In the following subsections, I illustrate the three types of copular contrasts by discussing data from the individual languages that belong to each category. In so doing, I show that — despite their differences — cognate forms share similar interpretive profiles across languages, and broadly reflect a contrast between limited and non-limited properties. I further discuss these contrasts relative to

Language(s)	Contrast Type	Environment(s)	Forms (Proto-Bantu root)	Interpretation
Kihavu (JD52)	<i>-li</i> vs <i>-ba</i>	Unrestricted	<i>-li</i> (*- <i>de</i> )	Limited
Mashi (JD53)			<i>-ba</i> (*- <i>bà</i> )	Non-limited
Kinyamulenge (JD61a)	<i>-ri</i> vs <i>ni</i>	Locative pred., 3rd person subj.	<i>-ri</i> (*- <i>de</i> )	Limited
Kirundi (JD62)			<i>ni</i> (?)	Known
Kifuliiru (JD63)	<i>-li</i> vs <i>-ba</i> vs <i>-tula</i>	Unrestricted	<i>-li</i> (*- <i>de</i> )	Limited
			<i>-ba</i> (*- <i>bà</i> )	Non-limited
			<i>-tula</i> (?)	Maximal

Table 4.1: Summary of copular forms in present tense predicational sentences

previous analyses of copular contrasts in other languages, particularly analyses of Spanish *estar/ser* (Carlson 1977, Diesing 1990, Kratzer 1995, Arche 2006, Deo et al. 2017, a.o.). Ultimately I show that no one previous analysis is capable of deriving the gamut of copular contrasts in the language sample, and provide an alternative pragmatic-oriented analysis based on the approach presented in Deo et al. (2017) that draws on aspects of multiple previous analyses to capture the data. In particular, I argue for a competition-based approach that distinguishes four (rather than five) interpretively distinct copulas, e.g., *-li* (or variant *-ri*), *-ba*, *-tula*, and *ni*. The core meaning difference between them pertains to the circumstances at which they assert the proposition to be true; *-li* expresses that the proposition is true of a limited circumstance, *-tula* expresses that the proposition is true of a maximal circumstance, *ni* expresses that the proposition is true of the utterance circumstance, and *-ba* expresses that the proposition is true of a more general circumstance.

### 4.3.1 *-li* vs *-ba*

One of the more salient copular contrasts among the Eastern Bantu languages surveyed in this project is the contrast between *-li* (\*-*de*) and *-ba* (\*-*bà*). Reflexes of the proto-Bantu copulas \*-*de* ‘be’ and \*-*bà* ‘become, be’ (Meeussen 1967, Guthrie 1967 1971, Wald 1973) are regularly attested in Bantu languages, and are often associated with distinct interpretive effects and distributional restrictions (see chapter 2). With respect to interpretation, reflexes of \*-*de*, e.g., *-le* in Mongo (C61; Hulstaert 1965), *-li* in Lamba (M54; Doke 1922), *-ri* in Kinyarwanda (JD61; Jerro 2015), etc., have been reported to coincide with stative or locative interpretations in some languages (Schneider-Zioga 2018, Gibson et al. 2019), while reflexes of \*-*bà*, e.g., *-ba* in Zulu (S42; Posthumus 2000),

*-wa* in Lamba (M54; Doke 1922), etc., are sometimes reported to yield inchoative “becoming” interpretations (Schneider-Zioga 2018). Distributionally, derivatives of *\*-de* and *\*-bà* often differ in terms of the Tense-Aspect-Mood (TAM) environments they may appear in. For example, reflexes of *\*-de* are often reported to be the sole copula employed in past tense or perfective environments, e.g., *-li* in Lamba (M54; Doke 1922) and Chichewa (N31b; Kiso 2012), and *-ri* in Kinyarwanda (JD61; Jerro 2015). In contrast, reflexes of *\*-bà* tend to have much wider distributions, though they often appear in irrealis environments like the future tense and infinitival mood, e.g., *-ba* in Kinyarwanda (JD61; Jerro 2015). In some languages, reflexes of both proto-Bantu copulas are present in a single environment, in which case they sometimes exhibit distinct interpretations.

In the languages that feature this contrast, *-li* and *-ba* generally distinguish between “temporary” and “permanent” predication; *-li* describes a temporary, stage-like property, while *-ba* describes a more permanent, individual-like property. The temporary/permanent *-li/-ba* contrast is particularly salient in the two JD50 languages considered in this study, namely Kihavu (JD52) and Mashi (JD53), though traces of the contrast can also be found in the other languages under investigation. For this reason, the following discussion will focus on elicited data from Kihavu (JD52) and Mashi (JD53).

In present tense “pure” predicational clauses, both Kihavu and Mashi exhibit two distinct copular verbs: *-li* and *-ba*. Importantly, the distribution and interpretation of these copular forms is largely the same in both languages. In general, either copula (in both languages) may be used with non-verbal predicates of any category, e.g., nominal, adjectival, or locative predicates.<sup>5</sup>

- (153) a. Johne a-<sup>✓</sup>li/<sup>✓</sup>ba                      ci-shambo  
           John 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 7NC-thief  
           ‘John is a thief’
- b. O-ma-lunga ga-<sup>✓</sup>li/<sup>✓</sup>ba                      ga-bulee  
           AUG-6-sky 6SM-be-<sub>LI</sub>/be-<sub>BA</sub> 6AGR-blue  
           ‘The sky is blue’
- c. Johne a-<sup>✓</sup>li/<sup>✓</sup>ba                      e      Kinshasa  
           John 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> LOC 9.Kinshasa  
           ‘John is in Kinshasa’

Mashi

<sup>5</sup>The present tense in Kihavu and Mashi is marked by the absence of tense morphology (i.e., T<sub>[+PRS]</sub> is unpronounced at PF). In other tenses, overt tense morphology surfaces between the subject marker and the verb root.

- (154) a. Murhula a-<sup>✓</sup>li/<sup>✓</sup>ba gi-shambo  
 Murhula 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 7NC-thief  
 ‘Murhula is a thief’
- b. O-mu-shangi gu-ani gwa-<sup>✓</sup>li/<sup>✓</sup>ba gu-kére  
 AUG-3NC-clothes 3AGR-1sg.poss 3SM-be-<sub>LI</sub>/be-<sub>BA</sub> 3AGR-clean  
 ‘My clothes are clean’
- c. Mugisha a-<sup>✓</sup>li/<sup>✓</sup>ba i Bujumbura  
 Mugisha 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> LOC 9.Bujumbura  
 ‘Mugisha is in Bujumbura’
- Kihavu

Though both copulas are generally available with most predicates, they are not freely interchangeable in either Kihavu or Mashi. Instead, each copula corresponds to a unique interpretation of the predication relation. At first glance, the distribution and interpretation of the two copulas closely aligns with the differences observed between the Spanish copulas *estar* and *ser* — which are reported to encode a stage/individual-level contrast (Carlson 1977, Diesing 1990, Kratzer 1995, Diesing 1992b, Arche 2006 a.o.). In such descriptions of the Spanish copular system, *ser* is treated as an individual-level copula in that it is used to predicate an inherent/permanent property of the subject, while *estar* is treated as a stage-level copula, as it is instead used to predicate a transient/temporary property of the subject. As such, the two copulas are reported to exhibit distributional differences in non-verbal predication, with *estar* co-occurring with stage-level predicates, e.g., *presente* ‘present’ (155), and *ser* with individual-level predicates, e.g., *inteligente* ‘intelligent’ (156).

- (155) El artista <sup>✓</sup>está/#es presente/ausente/lejos  
 The artist estar/ser.PRS.3SG present/absent/far away  
 ‘The artist is present/absent/far away’ Spanish (Maienborn 2005: 173)
- (156) Maria #está/<sup>✓</sup>es inteligente/altruista  
 Maria estar/ser.PRS.3SG intelligent/altruistic  
 ‘Maria is intelligent/altruistic’ Spanish (Deo et al. 2017: 8)

Likewise, we find that the two copular *be*-verbs in Kihavu and Mashi exhibit some sensitivity to the stage/individual-level predicate distinction. In Mashi for example, *-li* is preferred with canonically stage-level predicates, e.g., ‘sick’, (157), while *-ba* is instead preferred with individual-level predicates e.g., ‘tall’ (158).

(157) Maria a-<sup>✓</sup>li/#ba mu-lwala  
 Maria 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 1AGR-sick  
 ‘Maria is sick’ (She has a cold)

(158) Johne a-#li/<sup>✓</sup>ba mu-liri  
 John 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 1AGR-tall  
 ‘John is tall’ (A fact about John)

Mashi

The same general observation is true in Kihavu as well; *-li* is generally used to describe stage-level properties, e.g., ‘sick’ (159), while *-ba* is used to describe individual-level properties, e.g., Mugisha’s ethnicity (160).

(159) Murhula a-<sup>✓</sup>li/#ba mu-lwala  
 Murhula 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 1AGR-sick  
 ‘Murhula is sick’ (She has a cold)

(160) Mugisha a-#li/<sup>✓</sup>ba Mu-havu  
 Mugisha 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 1AGR-havu  
 ‘Mugisha is Muhavu’ (His ethnicity)

Kihavu

Although these examples point toward a lexicalized stage-individual contrast in both Spanish and Kihavu/Mashi, the true nature of this distinction turns out to be much more complicated. As discussed in subsection 4.2.2, the distribution of Spanish *estar/ser* is not determined by the type of predicate (i.e., stage/individual) involved in the predication relation, as expected in a true stage/individual contrast. In principle either copula may be used with a given predicate, they simply yield distinct interpretations of the predication relation; *estar* describes properties that hold of particular, limited situations, while *ser* describes more general properties (Clements 1988, 2006, Maienborn 2005, Deo et al. 2017).

Like what is reported in Spanish, we find that both *-li* and *-ba* in Kihavu/Mashi may be used with canonically stage-level predicates or canonically individual-level predicates, depending on the context. For example, while *-ba* is generally infelicitous with the stage-level predicate ‘sick’ in both Mashi (157) and Kihavu (159), it may be used if the property ‘sick’ holds indefinitely of the subject. In this sense, the use of *-ba* does not yield an individual-level property per se, but rather a “permanent” state; the property ‘sick’ holds permanently of Maria/Mugisha (161)-(162).

(161) *Context: Maria has had a chronic illness for her entire life.*

Maria a-<sup>✓</sup>ba mu-lwala  
Maria 1SM-be-<sub>BA</sub> 1AGR-sick  
'Maria is sick'

Mashi

(162) *Context: Mugisha has had a chronic illness for his entire life.*

Mugisha a-<sup>✓</sup>ba mu-lwala  
Mugisha 1SM-be-<sub>BA</sub> 1AGR-sick  
'Mugisha is sick'

Kihavu

Similarly, while *-li* is broadly dispreferred with individual-level descriptions uttered out of the blue in Mashi (158) and Kihavu (160), it can be used with canonically individual-level predicates to yield a context-specific, “temporary” state reading. For example, we find that Mashi *-li* may be used with the predicate ‘tall’ in (163), but it crucially yields a reading in which the denoted property holds only of a specific (limited) context of evaluation (e.g., John is taller than Maria *only* in this situation). Such a reading is naturally unavailable with *-ba*, which yields a permanent interpretation (164).

(163) *Context: John is normally shorter than Maria, but right now he’s wearing shoes that make him taller than her.*

Johne a-<sup>✓</sup>li/#ba mu-liri kulisha Maria  
John 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 1AGR-tall than Maria  
'John is taller than Maria'

(164) *Context: You are telling someone about John, and you know that he is taller than Maria.*

Johne a-#li/<sup>✓</sup>ba mu-liri kulisha Maria  
John 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 1AGR-tall than Maria  
'John is taller than Maria'

Mashi

In the same vein, Kihavu *-li* may be used with the ostensibly individual-level predicate ‘Muhavu’ in (165), however, it yields a slightly different reading than *-ba*. While *-ba* is used to describe known facts about Mugisha in (166), *-li* is used to describe a property that is relevant to (or even surprising in) a particular situation.

(165) *Context: You see Mugisha with other Bahavu, and you realize that he is Muhavu.*

Mugisha a-<sup>✓</sup>li/#ba Mu-havu  
 Mugisha 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 1AGR-havu  
 ‘Mugisha is Muhavu’

(166) *Context: You are telling someone about Mugisha, and you know that he is Muhavu.*

Mugisha a-#li/<sup>✓</sup>ba Mu-havu  
 Mugisha 1SM-be-<sub>LI</sub>/be-<sub>BA</sub> 1AGR-havu  
 ‘Mugisha is Muhavu’ Kihavu

Predictably, this sort of context sensitivity is also relevant in our original examples (153) and (154) where both *-li* and *-ba* are “available”. In the Mashi example (153b), *-li* and *-ba* yield distinct readings of the predication relation. With *-li*, the speaker describes a property that temporarily holds of ‘sky’; the sky is blue right now, but it may not be later. With copula *-ba* however, the speaker instead states a general fact about — or a permanent property of — the sky; it is known that the sky is blue (even if it is dark and stormy at the utterance time).

(167) *Context: It is noon, and you look outside and see that the sky is clear.*

O-ma-lunga ga-<sup>✓</sup>li/#ba ga-bulee  
 AUG-6-sky 6SM-be-<sub>LI</sub>/be-<sub>BA</sub> 6AGR-blue  
 ‘The sky is blue’

(168) *Context: It is overcast, but you are teaching a child everything you know about the sky.*

O-ma-lunga ga-#li/<sup>✓</sup>ba ga-bulee  
 AUG-6-sky 6SM-be-<sub>LI</sub>/be-<sub>BA</sub> 6AGR-blue  
 ‘The sky is blue’ Mashi

#### 4.3.1.1 Beyond temporary vs permanent

Given the more flexible use of *-li/-ba* than what is traditionally described of the Spanish copulas *estar/ser*, I use the terms “temporary state” (*-li*) and “permanent state” (*-ba*), as opposed to “stage-level” and “individual-level”, to describe the meanings of the two copulas in Kihavu and Mashi. Upon further inspection however, this description of the *-li/-ba* contrast is also insufficient.

One of the key observations in work on the Spanish *estar/ser* contrast is that *estar* is used to describe properties that are delimited in some way, e.g., temporally, locationally, etc. (Maienborn 2005, Deo et al. 2017). For Maienborn (2005), this restriction is formalized as a link to a particular discourse situation; *estar* describes a property that holds of a specific situation. Deo et al. (2017) instead suggest that *estar* describes properties that are true of a limited circumstance of evaluation that is defined according to particular contextual parameter values, e.g., world, time, location, etc. (see subsection 4.2.2). For them, *estar* “restricts” the property ascription to a specific discourse context in the sense that it describes properties that are limited in some way, e.g., properties that are true relative to a particular location (150), particular world, (151), or particular contextual standard (152).

In a similar way, the “temporary state” copula *-li* in Kihavu and Mashi can also be used to describe properties that are limited by things other than time. Consider again the use of Mashi *-li* in the sentence “the sky is blue” (167). In the original context in (167), *-li* yields a temporary interpretation; the sky is blue in this specific time interval (e.g., right now). However, the use of *-li* in this sentence is also compatible with a context like that in (169), which changes the relevant context to one that is locationally defined; the sky is blue in this specific location, but not (necessarily) elsewhere. In this context, *-li* describes a property that is spatially — not temporally — limited.

(167) *Context: It is noon, and you look outside and see that the sky is clear.*

O-ma-lunga ga-<sup>✓</sup>li/#ba ga-bulee  
 AUG-6-sky 6SM-be-<sub>LI</sub>/be-<sub>BA</sub> 6AGR-blue  
 ‘The sky is blue’

(169) *Context: John’s village is on the other side of the mountain from yours and he calls to tell you it is storming, however you look outside your home and see that the sky is clear.*

O-ma-lunga ga-<sup>✓</sup>li/#ba ga-bulee  
 AUG-6-sky 6SM-be-<sub>LI</sub>/be-<sub>BA</sub> 6AGR-blue  
 ‘The sky is blue’

Mashi

Like the use of *estar* in (151), *-li* can also be used to signal that a property is novel or surprising, i.e., that the relevant property does not hold of the subject in most (if not all) of the speaker’s other



possible worlds. An example of this can be seen in the Kihavu example in (165), where the use of *-li* communicates that the speaker was previously unaware that Mugisha is Muhavu. The property described by *-li* doesn't hold of a restricted time interval, but rather a limited set of worlds. In this sense, the predication relation receives a surprise interpretation because the majority of the speaker's possible worlds are not worlds where Mugisha is Muhavu.

(165) *Context: You see Mugisha with other Bahavu, and you realize that he is Muhavu.*

Mugisha	a- <b>li/#ba</b>	Mu-havu	
Mugisha	1SM-be- <sub>LI</sub> /be- <sub>BA</sub>	1AGR-havu	
'Mugisha is Muhavu'			Kihavu

Finally, *-li* can also be used to describe properties that exceed contextual expectations, just like Spanish *estar* (152). With a gradable predicate like 'tall', *-li* can be used to signal that the subject exceeds some pre-determined contextual standard. In (170) for example, the use of Mashi *-li* signals that the children's height exceeds the speaker's expectations of "tallness" in that context; compared to the heights of other similarly-aged children, the two children in question are quite tall.

(170) *Context: You are at the market, and you see two children who are much taller than the other children of the same age.*

a-ba-ana	ba-biri	ba- <b>li/#ba</b>	ba-liri	
AUG-2-child	2AGR-two	2SM-be- <sub>LI</sub> /be- <sub>BA</sub>	2AGR-tall	
'(The) two children are tall'				Mashi

In each of these cases, the function of *-li* is to restrict the property ascription to a context or situation that is limited in some way. If this limitation is temporally defined, *-li* describes a property that holds of a particular time frame, e.g., the sky is blue (now, but not earlier). If it is instead spatially defined, *-li* describes a property that holds of a particular location, e.g., the sky is blue (here, but not over there). Likewise, if this limitation is defined according to possible worlds or contextual standards, then *-li* describes properties that are true in a limited set of possible worlds, e.g., the sky is blue (but nobody predicted it would be), or properties that are true relative to a

particular contextual standard that differs from a more general standard, e.g., the sky is blue (relative to recent cloudy weather, but not summer weather).

As such, the general distinction between *-li* and *-ba* does not truly reflect a temporary/permanent contrast, but rather a contrast between properties that hold of smaller contexts/situations, and those that hold of larger contexts/situations. With *-li*, the speaker restricts the property ascription to a context/situation that is delimited by a single parameter, e.g., world, time, location, etc.. The effect of this is that the property ascription is, in a sense, “bound” by the limitations imposed by the relevant parameter; it does not extend beyond this context/situation. With *-ba* however, there is no similar restriction, and therefore no limited interpretation; *-ba* describes properties that are not anchored to a particular context/situation.

In summary, we observe that the copulas *-li* and *-ba* exhibit distinct interpretations in two JD50 languages, namely Kihavu (JD52) and Mashi (JD53). At first glance, the two copulas appear to distinguish something similar to the stage/individual-level contrast; *-li* describes temporary, transient properties, while *-ba* describes permanent properties. However, like what Deo et al. (2017) report of the Spanish copula *estar*, we find that the “temporary state” copula *-li* not only describes properties that are temporally limited, but also properties that are limited in other ways, e.g., properties that hold of particular locations, particular worlds, or particular contextual standards.

### 4.3.2 *-ri* vs *ni*

Another copular contrast observed in Eastern Bantu languages is that between the verbal copula *-ri* and the invariant copula *ni*. This contrast is primarily observed in languages of the JD60 Ruanda-Rundi subfamily, including Kinyarwanda (JD61; Jerro 2015), Kinyamulenge (JD61a), and Kirundi (JD62; Zorc and Nibagwire 2007). The verbal copula *-ri* is a reflex of proto-Bantu *\*-de*, just like *-li* in Kihavu (JD51) and Mashi (JD52). Languages that feature *-ri* tend to be languages that lack a phonemic /l/ vs /r/ contrast, e.g., the Ruanda-Rundi (JD60) languages. Invariant *ni* is observed in a variety of Bantu languages (Gibson et al. 2019), and is generally assumed to be related to focus given its use as a focus marker in many Bantu languages (McWhorter 1994, Schwarz 2003, Güldemann

2003). Although the distribution and interpretation of *-ri/ni* differ slightly across the JD60 languages surveyed in this project, the general contrast is the same: in present tense predicational sentences featuring a third person subject, *-ri* specifically describes a location, while *ni* describes a more general property.

Subject Person	Nominal Predicates	Adjectival Predicates	Locative Predicates
1st	<i>-ri</i>	<i>-ri</i>	<i>-ri</i>
2nd	<i>-ri</i>	<i>-ri</i>	<i>-ri</i>
3rd	<i>ni</i>	<i>ni</i>	<i>-ri/ni (-ba)</i>

Table 4.2: Copular forms in present predicational sentences - Kinyamulenge and Kirundi

Among the JD60 languages surveyed, there are three copular forms that appear in predicational sentences, namely invariant *ni*, and the two inflectional verbs *-ri* and *-ba*. The distribution of these three copulas is highly constrained by certain morphosyntactic factors, most notably Tense-Aspect-Mood environment. Of the three copular forms, *ni* has the most restrictive distribution, appearing only in present tense predicational sentences involving a third-person subject. Like other reflexes of proto-Bantu *\*-de*, *-ri* has a somewhat more variable distribution, appearing in both the present and past tense predicational sentences. In the present tense, *-ri* shows an additional sensitivity to the person feature of the subject; it is used in all predicational sentences involving a first- or second-person subject, but only occurs with locative predicates in sentences with a third-person subject. As in other Bantu languages, *-ba* (from proto-Bantu *\*-bà*) tends to appear in irrealis environments, e.g., future tense, infinitival mood, etc., however it is also observed in certain present tense sentences, particularly in Kinyamulenge and Kirundi. The distribution of the three copular forms is summarized in the table below.

Forms	Past	Present	Future	Infinitive/Subjunctive
<i>-ri</i>	✓	✓	✗	✗
<i>ni</i>	✗	✓ (3rd person)	✗	✗
<i>-ba</i>	✗	%	✓	✓

Table 4.3: Distribution of copulas by tense/mood environment - Kinyamulenge and Kirundi

Like the JD50 copulas *-li* and *-ba* discussed in subsection 4.3.1, the distributional overlap

between the three JD60 copulas is limited to present tense predicational sentences. In general, *-ri* and *ni* are the most commonly observed copular forms in this environment, with *-ba* appearing in only a subset of sentences (see subsection 4.3.2.1). Given the restrictive distribution of *ni*, the contrast between *-ri* and *ni* is narrowly restricted to sentences involving a third person subject.

In each of the JD60 languages under discussion, the primary contrast in present tense predicational sentences is that between *-ri* and *ni*. As previously mentioned, the distribution of *-ri* and *ni* is sensitive to the person feature value of the subject. With first and second person subjects, only *-ri* (or its allomorph *-di*) is available (171)-(172). In the presence of a third person subject however, either *-ri* or *ni* may be used depending on the nature of the predicate; *-ri* appears only with locative predicates (173b), while *ni* is used with all other types of predicates, e.g., nominals (173a). To illustrate this pattern, consider the following data from Kinyarwanda (Jerro 2015).

- (171) a. N-**di**            mu-nini  
           1SG.SM-be-<sub>RI</sub> 1AGR-big  
           ‘I am big’
- b. N-**di**            mu    rugo  
           2SG.SM-be-<sub>RI</sub> 18LOC house  
           ‘I am in the house’
- Kinyarwanda (Jerro 2015: 100)
- (172) a. U-**ri**            mu-nini  
           2SG.SM-be-<sub>RI</sub> 1AGR-big  
           ‘You are big’
- b. U-**ri**            mu    rugo  
           2SG.SM-be-<sub>RI</sub> 18LOC house  
           ‘You are in the house’
- Kinyarwanda (Jerro 2015: 100)
- (173) a. Kyle **ni**    mu-nini  
           Kyle be-<sub>NI</sub> 1AGR-big  
           ‘Kyle is big’
- b. Umw-aana a-**ri**        i    Kigali  
           1-child    1SM-be-<sub>RI</sub> LOC Kigali  
           ‘The child is in Kigali’
- Kinyarwanda (Jerro 2015: 94, 102)

Though the examples above suggest a close association between the form of the copula and its morphosyntactic environment, some have argued that *-ri* and *ni* are also *interpretively* distinct.

For example, Jerro (2015) shows that the distribution of Kinyarwanda *-ri* and *ni* cannot always be morphosyntactically defined; although *-ri* is generally used with locative predicates, *ni* is sometimes also available with subjects that denote locations, e.g., the city ‘Kigali’ (174b). In this context, Jerro notes that a subtle interpretive contrast arises: speakers judge *-ri* sentences as being “more about a location” than their *ni* counterparts.

- (174) a. Kigali **i-ri** mu Rwanda  
 Kigali 9SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’
- b. Kigali **ni** mu Rwanda  
 Kigali be-<sub>NI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’
- Kinyarwanda (Jerro 2015: 103)

In light of this contrast and the correlation between *-ri* and locative predicates, Jerro (2015) proposes that *-ri* involves an inherently locative semantics; *-ri* ascribes a physical location to the subject of the predication relation. For Jerro, the distribution of *-ri* in Kinyarwanda is not morphosyntactically determined, but rather a product of its locative semantics.<sup>6</sup>

The general idea proposed by Jerro (2015) — that *-ri* and *ni* in Kinyarwanda are interpretively distinct — is corroborated by similar interpretive effects in closely-related Kinyamulenge (JD61a) and Kirundi (JD62). Similar to the Kinyarwanda example in (174), both Kinyamulenge and Kirundi permit the use of either *-ri* or *ni* with a locative predicate depending on the context. In this environment, the two copulas yield interpretations that are similar to what is described by Jerro (2015): *-ri* describes “pure” location relations, while *ni* describes more general properties. Consider the interpretive contrast between *-ri/ni* in the Kinyamulenge examples below. In (175), the speaker makes an explicit statement about the spatial relationship between two things on a map; ‘Kigali’ is located within the borders of ‘Rwanda’. In this context, only *-ri* is available. In (176) however, the speaker describes a location as a known property of the subject; one thing that is true about Kigali is that it is in Rwanda.

---

<sup>6</sup>Jerro (2015) demonstrates that Kinyarwanda *-ri* may also co-occur with locative adverbial predicates such as *hano* ‘here’ and *aho* ‘there’. On the basis of this observation, Jerro suggests that *-ri* is not categorized to appear with prepositional phrases, but rather only appears with predicates that denote locations.

(175) *Context: You have never heard of Kigali or Rwanda before, but you see an official-looking map of East Africa and there is a city named ‘Kigali’ marked inside the borders of ‘Rwanda’.*

Kigali #ni/√i-ri mu Rwanda  
 Kigali be<sub>NI</sub>/9SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’

(176) *Context: You just got back from visiting Kigali, and you are listing all of the things you know about the city.*

Kigali √ni/#i-ri mu Rwanda  
 Kigali be<sub>NI</sub>/9SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’

Kinyamulenge

The same judgements are also observed in Kirundi (JD62): *-ri* is used to describe an explicit spatial relation (177), while *ni* is used to describe a known fact about the subject (178).

(177) *Context: You have never heard of Kigali or Rwanda before, but you see an official-looking map of East Africa and there is a city named ‘Kigali’ marked inside the borders of ‘Rwanda’.*

Kigali #ni/√i-ri mu Rwanda  
 Kigali be<sub>NI</sub>/9SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’

(178) *Context: You just got back from visiting Kigali, and you are listing all of the things you know about the city.*

Kigali √ni/#i-ri mu Rwanda  
 Kigali be<sub>NI</sub>/9SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’

Kirundi

#### 4.3.2.1 Additional complications

Just as Jerro (2015) reports in Kinyarwanda, both Kinyamulenge and Kirundi exhibit an interpretive contrast between *-ri* and *ni*. In third person predicational clauses featuring a locative predicate, *-ri* describes overt spatial relations, while *ni* describes more general properties. Unlike Kinyarwanda

however, the interpretive contrasts in the Kinyamulenge and Kirundi copular systems extend beyond a simple distinction between locative and non-locative predication.

The central claim in Jerro (2015) is that Kinyarwanda *-ri* is inherently locative. In present tense predicational sentences featuring a third person subject, *-ri* appears exclusively with locative predicates and yields a purely spatial interpretation. As illustrated above, *-ri* behaves similarly in closely-related Kinyamulenge, where it again correlates with locative predicates and spatial interpretations. In contrast to its cognate in Kinyarwanda however, Kinyamulenge *-ri* is not interpretively restricted to statements about locations. In (179)-(180), *-ri* is used to describe a taxonomic relation, e.g., class membership.

(179) Inka i-ri mu nyamanswa  
 9.cow 9SM-be-RI 18LOC 10.animal  
 ‘The cow is an animal’ (by classification)

(180) Kigali i-ri mu mi-ji my-iiza  
 Kigali 9SM-be-RI 18LOC 4-city 4AGR-nice  
 ‘Kigali is one of the nice cities’ Kinyamulenge

Though both feature a locative predicate, the above sentences do not express overt spatial relations. In both cases, the combination of the copula *-ri* and the locative predicate yields a classificational reading, e.g., the species ‘cow’ belongs to the class of ‘animals’ (179), and ‘Kigali’ belongs to the class of ‘nice cities’ (180). Correspondingly, such examples do not permit the *-ri/ni* (locative vs non-locative) alternation observed in other predicational sentences featuring a locative predicate, as they do not encode a true spatial relation. With a locative class-denoting predicate like that in (181), only *-ri* is available. The copula *ni* may only be used if a restrictive relative clause is adjoined to the predicate (182).<sup>7</sup>

(181) Imbwa #ni/<sup>✓</sup>i-ri mu nyamanswa  
 9.dog be-RI/9SM-be-RI 18LOC 10.animal  
 ‘The dog is an animal’ (by classification)

---

<sup>7</sup>It is not immediately clear why the addition of a restrictive relative clause results in the acceptability of *ni*. I will leave any discussion of this effect to future work.

- (182) Imbwa <sup>✓</sup>**ni**/<sup>✓</sup>**i-ri** mu nyamanswa zi-ruka cyane  
 9.dog be<sub>NI</sub>/9SM-be-<sub>RI</sub> 18LOC 10.animal 10SM-run fast  
 ‘The dog is one of the animals that runs fast’ Kinyamulenge

Another difference between the Kinyarwanda copular system described by Jerro (2015) and the copular systems of Kinyamulenge and Kirundi is that a third copula, *-ba*, is attested in present tense predicational sentences in the latter two languages. In all three JD60 languages under discussion, there is a general restriction on the use of *ni* with locative predicates; *ni* cannot co-occur with a locative predicate in the presence of a definite subject. Jerro (2015) briefly mentions this restriction in Kinyarwanda, as he reports that the interpretive contrast in examples like (174b) is restricted to locative statements about immovable locations, e.g., the location of a city. With a non-location subject, the *-ri/ni* contrast collapses (183)

- (183) Umw-aana #**ni**/<sup>✓</sup>**a-ri** i Kigali  
 1-child be<sub>NI</sub>/9SM-be-<sub>RI</sub> LOC Kigali  
 ‘The child is in Kigali’ Kinyarwanda (Adapted from Jerro 2015: 102)

A similar neutralization is observed in Kinyamulenge and Kirundi, albeit with an additional complication. Although *ni* cannot be used to ascribe a location to the subject in (184)-(185), a third copula, *-ba*, can be used to ascribe a “permanent” location to the subject, e.g., where John lives.

- (184) a. John **a-ri** i Kigali  
 John 1SM-be-<sub>RI</sub> LOC Kigali  
 ‘John is in Kigali’ (John’s current location)
- b. John **a-ba** i Kigali  
 John 1SM-be-<sub>BA</sub> LOC Kigali  
 ‘John is in Kigali’ (John lives in Kigali)
- c. #John **ni** i Kigali  
 John 1SM-be-<sub>NI</sub> LOC Kigali  
 ‘John is in Kigali’ Kinyamulenge
- (185) a. John **a-ri** mu Rwanda  
 John 1SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘John is in Rwanda’ (John’s current location)



b. John a-**ba** mu Rwanda  
 John 1SM-be<sub>BA</sub> 18LOC Rwanda  
 ‘John is in Rwanda’ (John lives in Rwanda)

c. #John **ni** mu Rwanda  
 John 1SM-be<sub>NI</sub> 18LOC Rwanda  
 ‘John is in Rwanda’

Kirundi

Interestingly, the contrast between *-ri* and *-ba* in these examples directly mirrors the contrast between their respective cognates *-li* and *-ba* in the JD50 languages Kihavu (JD52) and Mashi (JD53) (see subsection 4.3.1.1); *-ri/-li* (\*-*de*) describes a “limited” spatial relation (e.g., one that holds right now), while *-ba* (\*-*bà*) describes a more permanent spatial relation. One possible conclusion that could be taken from this overlap is that *-ri* is not truly locative in the sense of Jerro (2015). Instead, *-ri* — like JD50 *-li* — could describe properties that are in some way delimited (Maienborn 2005, Deo et al. 2017), albeit in a much more constrained way.

Following the discussion of *-li/-ba* in subsection 4.3.1.1, one way to capture the distribution and interpretation of JD60 *-ri* is to assume that *-ri* describes properties that are true of limited circumstances of evaluation (Deo et al. 2017), and that *-ri* is only licensed in the presence of a morphosyntactically locative predicate (in present tense predicational sentences featuring a third person subject). Like Spanish *estar* and JD50 *-li*, *-ri* is used to describe properties that hold of a circumstance limited by some contextual parameter, e.g., world, time, location, etc.. For example, in (175) Kinyamulenge *-ri* describes a spatial relation that is limited to a locationally-defined context, e.g., a map.<sup>8</sup>

(175) *Context: You have never heard of Kigali or Rwanda before, but you see an official-looking map of East Africa and there is a city named ‘Kigali’ marked inside the borders of ‘Rwanda’.*

Kigali i-**ri** mu Rwanda  
 Kigali 9SM-be<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’

Kinyamulenge

---

<sup>8</sup>Even the referents of the relative objects involved in this sentence may be defined relative to this circumstance, e.g., ‘Kigali’ could simply refer to the location/marker indicated on the map, not the actual city in the real world.

Similarly, *-ri* may be used to describe a spatial relation that holds of a temporally-limited circumstance as well, e.g., a spatial relation that is true today, but not necessarily another day (186).

(186) *Context: John is visiting Kigali while on vacation today.*

John a-**ri** i Kigali  
 John 1SM-be-<sub>RI</sub> LOC Kigali  
 ‘John is in Kigali’

Kinyamulenge

From an interpretive perspective, these uses of *-ri* are almost identical to the locationally-bound and temporally-bound interpretations of its JD50 cognate *-li* (see subsection 4.3.1.1); in both cases, the two copulas describe properties that hold of a limited context. One place where we see this similarity is when *-ri* and *ni* are both available with a locative predicate, e.g., (175)-(178). In these examples, *-ri* is used to describe a property that is restricted to a particular context, e.g., the spatial relationship between two things on a map, while *ni* is used to describe more general, known properties, e.g., things that are true about ‘Kigali’. That said, the similarity between JD60 *-ri* and JD50 *-li* is even more apparent in examples where *ni* is unavailable (e.g., sentences with definite subjects), and *-ri* is in contrast with *-ba* (185). Here, *-ri* and *-ba* quite clearly distinguish between a temporally-constrained property (185a) and a more general property (185b), just like the JD50 copulas *-li/-ba*.

(185) a. John a-**ri** mu Rwanda  
 John 1SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘John is in Rwanda’ (John’s current location)

b. John a-**ba** mu Rwanda  
 John 1SM-be-<sub>BA</sub> 18LOC Rwanda  
 ‘John is in Rwanda’ (John lives in Rwanda)

Kirundi

Despite this overlap in interpretation, there is still a salient difference in the distribution of the JD60 copula *-ri* and its JD50 counterpart *-li*; in predicational sentences involving a third person subject, *-ri* may only occur with locative predicates. What’s more, the two cognates participate in distinct contrasts; JD60 *-ri* generally contrasts with a copula, e.g., *ni* that is not observed in the relevant JD50 languages.

With respect to its restricted distribution, it is not immediately clear why *-ri* is limited to locative predicates in the third person in these three JD60 languages, nor why two reflexes of proto-Bantu *\*-de* (e.g., JD50 *-li* and JD60 *-ri*) would exhibit such distinct distributions. That being said, many Bantu languages exhibit similar interactions between person and predicate category. As discussed in Schneider-Zioga (2018) and Gibson et al. (2019), Bantu languages often employ a dedicated copula for locative predication.<sup>9</sup> In many languages, this involves the use of an invariant copula, e.g., Kinande (JD42) *ni*, with nominal and adjectival predicates, but a verbal copula, e.g., Kinande (JD42) *-li*, with locative predicates (Schneider-Zioga and Mutaka 2015a). The use of a verbal copula — specifically reflexes of proto-Bantu *\*-de*, e.g., *-li/-ri* — with locative predicates is particularly common among Bantu languages, and is observed in languages such as Herero (R31), Mbundu (H21), Luvale (K14), and Kinande (JD42), among others (Lanham 1953, Schneider-Zioga 2018).

In a similar way, many Bantu languages also exhibit a sensitivity to subject person in predicational sentences, specifically the distinction between first/second person subjects and third person subjects. Interestingly, this generally involves a similar pattern across languages: first/second person subjects co-occur with a verbal copula, e.g., *-li/-ri*, while third person subjects co-occur with an invariant copula, e.g., *ni* (Lanham 1953, Schneider-Zioga 2018). As such, many Bantu languages exhibit similar systems to that observed in the JD60 languages under discussion. For example, both Lamba (M54; Doke 1922) and Kinande (JD42; Schneider-Zioga and Mutaka 2015a) employ a verbal copula *-li* (*\*-de*) in all first/second person sentences, and an invariant copula (*ni* in Kinande) or tonal marking (Lamba) in third person sentences featuring a nominal or adjectival predicate. Both languages employ the verbal copula *-li* with locative predicates regardless of person (Schneider-Zioga 2018). This pattern is almost identical to what we have observed in Kinyarwanda, Kinyamulenge, and Kirundi; the contrast between *-ri* and *ni* is limited to the third person, in which case *-ri* occurs with locative predicates.

As such, there is strong typological evidence to suggest that reflexes of proto-Bantu *\*-de* are associated with locative predication in some way, and that cognates of the invariant copula *ni* are

---

<sup>9</sup>Outside Bantu, this can also be seen in languages like Thai, which utilizes a dedicated copula *jù:* with locative predicates (Wongwattana 2015).

sometimes limited to the third person, though the exact nature of these associations remain unclear. Of course, the first of these points is one of the central motivations for the locative analysis of Kinyarwanda *-ri* in Jerro (2015): *-ri* co-occurs with locative predicates because it has a locative semantics. As previously mentioned however, a purely locative analysis of JD60 *-ri* would fail to account for the fact that Kinaymulenge *-ri* can be used to signal a taxonomic relation in examples like (179)-(180).

(179) Inka i-ri mu nyamanswa  
 9.cow 9SM-be-RI 18LOC 10.animal  
 ‘The cow is an animal’ (by classification)

(180) Kigali i-ri mu mi-ji my-iiza  
 Kigali 9SM-be-RI 18LOC 4-city 4AGR-nice  
 ‘Kigali is one of the nice cities’ Kinyamulenge

That being said, these readings could feasibly be accounted for if *-ri* instead describes a limited context/situation in the sense of Maienborn (2005) and Deo et al. (2017). Assuming that property ascriptions can be restricted to a particular physical context (e.g., a map, here, etc.), it is plausible that they may also be limited to a particular non-physical context as well (e.g., a taxonomic set). The taxonomic reading of *-ri* arises from the speaker describing the metaphorical “spatial” relationship between the subject and a larger set. Just as *-ri* can be used to describe the location of a city in a defined space, e.g., a map, it can also be used to describe an object as a member of a taxonomic set. In this sense, *-ri* describes a relation between two sets, e.g., ‘cow’ and ‘animal’, specifically as it pertains to taxonomy; just as a locational relation can be interpreted relative to a map, a set relation can be interpreted relative to a general taxonomic tree.

If the JD60 copula *-ri* is not inherently locative as I argue here, then there must be an alternative explanation for its restrictive distribution in the third person. One possibility is that this restriction is in some way conditioned by the distribution of the invariant (focus) copula *ni*, which is often restricted to the third person and rarely occurs with locative predicates across Bantu languages (Lanham 1953, Schneider-Zioga and Mutaka 2015a, Schneider-Zioga 2018). In any case, I will not attempt to fully resolve this issue here. For the time being, I will simply assume that the *-ri/ni*

contrast is — for whatever reason — restricted to certain morphosyntactic environments.

### 4.3.3 *-li vs -ba vs -tula*

The final copular contrast observed in our language sample builds on the *-li/-ba* contrast observed in other Great Lakes Bantu languages (see sections 4.3.1-4.3.2), and is only found in the JD60 language Kifuliiru (JD63).<sup>10</sup> Like other JD50 and JD60 languages, Kifuliiru exhibits a broad interpretive contrast between reflexes of proto-Bantu *\*-de* and *\*-bà* in present tense predicational sentences; *-li* (*\*-de*) generally describes temporary or otherwise limited properties, while *-ba* (*\*-bà*) describes more permanent, individual-like properties. Unlike the other languages discussed in this chapter however, Kifuliiru also exhibits a third copula, *-tula*, in present tense predicational sentences that directly contrasts with *-li* and *-ba*. Interpretively, this third copula yields a specific kind of “permanent” interpretation that is subtly distinct from *-ba*; *-tula* describes properties that are true of maximal situations/circumstances.<sup>11</sup> This is, to my knowledge, the only three-way copular contrast of this type to be reported in Bantu languages.<sup>12</sup>

Like the other JD50/JD60 languages discussed previously, the distribution of the three Kifuliiru copulas *-li*, *-ba*, and *-tula* is heavily constrained by Tense-Aspect-Mood environment. Just as before, the general contrast between the three copulas is restricted to the present tense. In most cases, the present tense is null-marked in copular clauses in Kifuliiru, however it is obligatorily expounded as the present tense marker *-mu* when *-ba* is used (187).<sup>13</sup>

<sup>10</sup>The majority of data in this section were first presented in Finholt (2023). For this reason, there are at times substantial overlap between the two.

<sup>11</sup>As mentioned previously, the historical source of *-tula* is unknown, however it appears to be related to Kinyarwanda (JD61) *-tura* ‘dwell, inhabit, live’.

<sup>12</sup>The other JD60 languages in this work, Kinyamulenge (JD61a) and Kirundi (JD62) also involve three copular forms, e.g., *-ri*, *-ba*, and *ni*, however they are not mutually contrastive (i.e., there is no context where they are simultaneously in competition with one another). These languages also differ from Kifuliiru in that they involve two verbal copulas (*-ri*, *-ba*) and one invariant copula (*ni*). Lamba (M54; (Doke 1922) is the only other Bantu language I know of that features three copular *be*-verbs. That said, the interpretive contrast between the *be*-verbs in Lamba is slightly different from those in Kifuliiru (see Schneider-Zioga 2018 for further discussion).

<sup>13</sup>The same marker can be seen in non-copular clauses, where it is sometimes used as a present tense/present progressive marker.

- (1) a-mu-soma  
ISM-PRS-read  
‘S/he is reading’

- (187) a. Safari a-**li** mu-nganga  
 Safari 1SM-be-<sub>LI</sub> 1NC-doctor  
 ‘Safari is a doctor’
- b. Safari a-*mu*-**ba** mu-nganga  
 Safari 1SM-PRS-be-<sub>BA</sub> 1NC-doctor  
 ‘Safari is a doctor’
- c. Safari a-**tula** mu-nganga  
 Safari 1SM-be-<sub>LI</sub> 1NC-doctor  
 ‘Safari is a doctor’

Kifuliiru

In other Tense-Aspect-Mood environments, the picture is more complicated. Unlike the other Bantu languages discussed here, Tense-Aspect-Mood is often expressed periphrastically through the use of an auxiliary verb in Kifuliiru, including in copular constructions (Otterloo 2011a,b). Such auxiliary verbs are transparently related to the copulas *-li/-ba*, as is reported in other JD60 languages, e.g., Kinyarwanda (Botne 1986).

That said, a full description of the Tense-Aspect-Mood system in Kifuliiru is well beyond the scope of this chapter. What is important for our discussion is the fact that, to my knowledge, no two copulas appear in the same environment outside the present tense.

With respect to the interpretative differences between the three Kifuliiru copulas, the primary contrast is a familiar one. Just as we see in Spanish and many of the other JD50/JD60 languages, the Kifuliiru copular system makes a strong distinction between transient, limited properties on the one hand and more durative, individual-level properties on the other. In general, this manifests as a binary distinction between the copulas *-li* and *-tula*; with the locative predicate “in Kinshasa”, *-li* is used to describe a temporally-limited property (188), while *-tula* is used to describe a “permanent” property (189).<sup>14</sup>

---

<sup>14</sup>It is unclear why there is no overt locative morphology in these examples, as locatives are generally associated with a locative noun class marker as in many other Bantu languages. This may be related to a similar phenomenon in other JD60 languages where a “reduced” locative form *i* is used with certain locations like cities (Jerro 2015).

- (1) Umw-aana a-**ri** i Kigali  
 1-child 1SM-be-<sub>RI</sub> LOC Kigali  
 ‘The child is in Kigali’

Kinyarwanda (Jerro 2015: 102)

(188) *Context: Maneno is on vacation in Kinshasa today.*

Maneno a-<sup>✓</sup>**li/#tula** Kinshasa  
Maneno 1SM-be-<sub>LI</sub>/be-<sub>TULA</sub> Kinshasa  
'Maneno is in Kinshasa'

(189) *Context: Maneno lives in Kinshasa, and everyone knows this about him.*

Maneno a-**#li/**<sup>✓</sup>**tula** Kinshasa  
Maneno 1SM-be-<sub>LI</sub>/be-<sub>TULA</sub> Kinshasa  
'Maneno is in Kinshasa' (he lives there) Kifuliiru

Though similar to a distinction between stage- and individual-level predication (Carlson 1977, Kratzer 1995, Diesing 1990, Arche 2006) a.o., the two copulas do not show a distributional sensitivity to the identity of the predicate. Both copulas are generally available with any predicate, including adjectival and nominal predicates. For example, either copula may be used with the individual-level predicate 'tall' depending on the context. In (190), the description 'short' is limited to a particular context (or rather, a particular contextual standard Deo et al. 2017); when standing next to this very tall person, Safari is short. In contrast, the context in (191) licenses a more general interpretation of the same predicate; 'short' is something that Safari is by nature.

(190) *Context: Safari is considered tall, but now he's standing next to someone much taller than him.*

Safari a-<sup>✓</sup>**li/#tula** mu-fi  
Safari 1SM-<sub>LI</sub>/be-<sub>TULA</sub> 1AGR-short  
'Safari is short'

(191) *Context: Safari is shorter than everyone in town, and everyone knows this fact about him.*

Safari a-**#li/**<sup>✓</sup>**tula** mu-fi  
Safari 1SM-be-<sub>LI</sub>/be-<sub>TULA</sub> 1AGR-short  
'Safari is short' Kifuliiru

Considering the other copular contrasts discussed in this chapter — which similarly distinguish properties that hold of a limited context and those that do not — the distinction between *-li* and *-tula* is relatively unsurprising. What is surprising however are the copular forms that Kifuliiru employs

to make this contrast; unlike Kihavu (JD52), Mashi (JD53), and some other JD60 languages, this is expressed as a distinction between *-li/-tula* rather than *-li/-ba*. Despite this difference, the copula *-ba* does appear in present tense predicational clauses in Kifuliiru, albeit with the addition of the present tense marker *-mu*. Interpretively, *-ba* yields individual-like interpretations that are quite similar to those associated with *-tula*. In examples like (192)-(193) below, the two copulas are essentially interchangeable in that they both yield a “permanent” reading.

(192) *Context: Maneno lives in Kinshasa, and everyone knows this about him.*

- a. Maneno a-mu-<sup>✓</sup>**ba** Kinshasa  
 Maneno 1SM-PRS-be-<sub>BA</sub> Kinshasa  
 ‘Maneno is in Kinshasa’
- b. Maneno a-<sup>✓</sup>**tula** Kinshasa  
 Maneno 1SM-be-<sub>TULA</sub> Kinshasa  
 ‘Maneno is in Kinshasa’

(193) *Context: Safari is a doctor by trade, and he has been his whole working life.*

- a. Safari a-mu-<sup>✓</sup>**ba** mu-ganga  
 Safari 1SM-PRS-be-<sub>BA</sub> 1NC-doctor  
 ‘Safari is a doctor’
- b. Safari a-<sup>✓</sup>**tula** mu-ganga  
 Safari 1SM-be-<sub>TULA</sub> 1-doctor  
 ‘Safari is a doctor’

Kifuliiru

In other environments however, there is a salient difference between *-ba* and *-tula*. For example, *-ba* cannot co-occur with an individual-level predicate and a definite human (class 1) subject, regardless of context; *-ba* is unavailable in both a limited context where *-li* is available (194), and a more general context where *-tula* is available (195).

(194) *Context: Safari is considered short, but now he’s standing next to someone much shorter than him.*

- a. Safari a-<sup>✓</sup>**li** mu-la  
 Safari 1SM-be-<sub>LI</sub> 1AGR-tall  
 ‘Safari is tall’



- b. Safari a-mu-**#ba** mu-la  
 Safari 1SM-PRS-be-<sub>BA</sub> 1AGR-tall  
 ‘Safari is tall’
- c. Safari a-**#tula** mu-la  
 Safari 1SM-be-<sub>TULA</sub> 1AGR-tall  
 ‘Safari is tall’

(195) *Context: Safari is taller than everyone in town, and everyone knows this fact about him.*

- a. Safari a-**#li** mu-la  
 Safari 1SM-be-<sub>LI</sub> 1AGR-tall  
 ‘Safari is tall’
- b. Safari a-mu-**#ba** mu-la  
 Safari 1SM-PRS-be-<sub>BA</sub> 1AGR-tall  
 ‘Safari is tall’
- c. Safari a-<sup>✓</sup>**tula** mu-la  
 Safari 1SM-be-<sub>TULA</sub> 1AGR-tall  
 ‘Safari is tall’

Kifuliiru

There are however other contexts where *-ba* may co-occur with an individual-level predicate. Specifically, *-ba* can be used with an individual-level predicate in the presence of a generic subject, in which case it yields a kind-reading. When used with predicates like ‘tall’ or ‘short’ (196), *-ba* characterizes kinds of trees; trees of this kind are generally tall/short.<sup>15</sup>

- (196) Kino ki-ti ki-mu-**ba** ki-refu/ofi  
 7.DEM 7-tree 7SM-PRS-be-<sub>BA</sub> 7AGR-tall/short  
 ‘This (kind of) tree is tall/short’

In contrast, the use of either *-li* or *-tula* in the same sentence frame will yield non-kind readings; with *-li*, a particular tree is interpreted to be tall/short in a particular context (197a), while with *-tula*, a particular tree is assumed to have always been tall/short (197b).

- (197) a. Kino ki-ti ki-**li** ki-refu/ofi  
 7.DEM 7-tree 7SM-be-<sub>LI</sub> 7AGR-tall/short  
 ‘This (particular) tree is tall/short’ (as compared to the trees around it)

<sup>15</sup>The predicate *-refu* ‘tall’ in (196) is different from the predicate *-la* ‘tall’ seen previously, e.g., (194)-(195). That said, these predicates are interchangeable; *-refu* could be used in (194)-(195) with no change in meaning. The relative distribution of the two forms is likely highly variable across speakers, as *-refu* is likely a borrowing of the Swahili predicate *-refu* ‘tall’.

- b. Kino    ki-ti    ki-**tula**    ki-refu/ofi  
       7.DEM 7-tree 7SM-be-<sub>LI</sub> 7AGR-tall/short  
       ‘This (particular) tree is tall/short’ (it has been since it first began to grow)    Kifuliiru

In summary, Kifuliiru utilizes three different copular *be*-verbs, e.g., *-li*, *-muba*, and *-tula*, in present tense pure predicational clauses. The first copula, *-li*, broadly coincides with temporary, stage-like interpretations; it describes a property that holds of a limited context of evaluation. In contrast, *-tula* is associated with permanent, individual-like interpretations; it describes a property that holds of a much broader context. Finally, *-muba* similarly yields permanent interpretations, but it can only co-occur with individual-level predicates in the presence of a generic subject, in which case it yields a kind-reading.

#### 4.3.3.1 Three-way contrast

From a theoretical standpoint, the three-way contrast in Kifuliiru represents a significant obstacle to accounts that attribute copular contrasts to a binary distinction between stage- and individual-level predication, or something similar, e.g., Milsark (1974), Carlson (1977), Diesing (1992a), and Kratzer (1995), a.o.. If we assume that copular systems only distinguish properties that hold temporarily from properties that don’t, then there is no way to account for a third, intermediate copula like Kifuliiru *-ba*, particularly given its unique interpretive effects.

One of the few accounts to propose a three-way distinction in copular predication is that presented in (Roy 2013). For Roy, copular contrasts like those discussed here are derived from syntactic differences between predicates. Assuming contra Kratzer (1995) that all predicates are predicates of eventualities, the stage/individual split — among other interpretive distinctions — must arise due to differences in their structure.

On the basis of predicational contrasts in French, Spanish, and Russian, Roy (2013) posits three types of non-verbal predicate that differ in both meaning and structure (198). Beginning with the least structurally complex of the three, *dense* (situation-descriptive) predicates are said to range over mass (non-atomic) eventualities, yield temporary/limited interpretations, and consist of a bare predicate XP. On the other hand, *non-dense* (characterizing) predicates are said to range over

atomic eventualities, allow for interpretive “gaps”, and are associated with an additional projection above XP, the classifier phrase (CIP). Finally, *maximal* (defining) predicates range over maximal eventualities, yield “permanent” state interpretations, and are uniquely associated with the presence of a number phrase (NumP) above CIP and XP.

(198) *Predicate types in Roy (2013)*

- a.                    [XP]            dense    (situation-descriptive)
- b.                    [CIP[ ... ]]    non-dense            (characterizing)
- c.                    [NumP[ ... ]]    maximal                    (defining)

Structural differences aside, the important part of this analysis for our discussion of the Kifuliiru copular system is the fact that there are three possible interpretations of predicational sentences. Situation-descriptive readings involve properties that are in some way limited, e.g., being sick with a cold (temporally limited), being tall in a crowd (contextually limited), etc. Characterizing readings involve properties that hold of the subject more generally, but need not be true in all contexts, e.g., being a sad person (where a sad person does not need to be “sad” all the time to qualify as a “sad person”). Finally, defining readings involve properties that are true of maximal events/situations, e.g., being human.

The distinction between situation-descriptive and characterizing interpretations can be seen in a sentence like (199), which Roy (2013) argues has two possible readings.

(199) Paul est malade  
        Paul is sick  
        ‘Paul is sick’ French (Roy 2013: 72)

The first reading is one where Paul is understood to be sick right now; he is currently experiencing sickness (Roy 2013: 72). In this reading, the property ‘sick’ is tied to the situation/context being described. This is the situation-descriptive reading; the property holds of the subject in a particular, limited situation. In contrast, the second reading is one where Paul is understood to be someone who is chronically or regularly ill. In this context, ‘sick’ is a property that is generally true of Paul even if he is not feeling sick at any particular moment; Paul is still ‘sick’ even if he is feeling well

right now. This is the characterizing reading; the property holds of a broader situation, but not necessarily all of its subparts.

Like characterizing sentences, defining sentences involve properties that hold of a more general context/situation, however they do not allow for “interpretive gaps” in the same way. Roy (2013) illustrates this difference by comparing their ability to answer two kinds of subject-oriented Wh-questions, namely *what* and *who* questions. In French, felicitous answers involving a nominal predicate differ for each question type; answers to *what* questions involve a bare nominal predicate, e.g., *acteur* ‘actor’ (200), while answers to *who* questions involve an indefinite nominal, e.g., *un acteur* ‘an actor’ (201). Roy (2013) argues that this difference directly corresponds to the characterizing/defining contrast; characterizing sentences involve a bare nominal predicate, while defining sentences involve an indefinite nominal predicate.

(200) Qu’est Raymond?

what.is Raymond

‘What does/is Raymond?’

a. #Raymond est un acteur

Raymond is an actor

‘Raymond is an actor’

b. Raymond est acteur

Raymond is actor

‘Raymond is an actor’

French (Roy 2013: 38)

(201) Qui est Raymond?

who is Raymond

‘Who is Raymond?’

a. Raymond est un acteur

Raymond is an actor

‘Raymond is an actor’

b. #Raymond est acteur

Raymond is actor

‘Raymond is an actor’

French (Roy 2013: 38)

To felicitously answer a *what* question, the speaker must provide a characterizing description of the subject. In (200), this corresponds to a job/career description, e.g., Raymond is an actor by

training. Importantly, this kind of description does not require Raymond to be an actor continuously; Raymond may not be currently involved in any acting projects, but he still works as an actor. To felicitously answer a *who* question however, the speaker must instead provide a defining description of the subject. In (201), this amounts to describing a core aspect of who Raymond is, e.g., identifying him as an actor. In this sense, the speaker describes Raymond being an actor in the same vein as Raymond being a person; being an actor is part of who Raymond is, not just something he does.

Returning to our discussion of the Kifuliiru copular system, we observe a similar three-way contrast that distinguishes temporary/limited properties from two types of “permanent” properties (subsection 4.3.3). Like dense (situation-descriptive) predicates, the copula *-li* describes properties that hold of limited contexts. In (197a) for example, the use of *-li* yields a reading in which the tree under discussion to is perceived to be tall/short in a specific context; compared to the other trees nearby, this tree is short (even if it is generally quite tall for a tree).

(197a) Kino ki-ti ki-**li** ki-refu/ofi  
 7.DEM 7-tree 7SM-be-<sub>LI</sub> 7AGR-tall/short  
 ‘This (particular) tree is tall/short’ (as compared to the trees around it) Kifuliiru

On the other hand, *-ba* and *-tula* describe properties that are not limited in the same way. Like non-dense (characterizing) predicates, *-ba* describes properties that hold of broader situations, but not necessarily all of its subparts. In (196), this manifests as a kind reading; though this (kind) of tree is generally tall/short, not all individuals in this kind are.

(196) Kino ki-ti ki-mu-**ba** ki-refu/ofi  
 7.DEM 7-tree 7SM-PRS-be-<sub>BA</sub> 7AGR-tall/short  
 ‘This (kind of) tree is tall/short’

(202) Safari a-mu-**ba** mu-ganga  
 Safari 1SM-PRS-be-<sub>BA</sub>  
 ‘Safari is a doctor’ Kifuliiru

Likewise, *-tula* is similar to maximal (defining) predicates in the sense that it describes properties that are true of maximal situations. In (197b), this yields a reading in which a particular tree has

always been short/tall; since it was first planted, this tree has always been short/tall relative to other trees/our expectations.

- (197b) Kino ki-ti ki-**tula** ki-refu/ofi  
 7.DEM 7-tree 7SM-be-<sub>LI</sub> 7AGR-tall/short  
 ‘This (particular) tree is tall/short’ (it has been since it first began to grow) Kifuliiru

As for examples where the two copulas appear to be interchangeable (193), there is actually a subtle interpretive distinction that closely aligns with the interpretive distinction between bare nominal predicates and indefinites in French (200)-(201). With *-ba*, the speaker simply describes *what* Safari is/does, e.g., he works as a doctor, but he sometimes also works as a teacher (193). However with *-tula*, there is a sense that the speaker is instead describing *who* Safari is on a more fundamental level, e.g., Safari is and will always be a doctor.

(193) *Context: Safari is a doctor by trade, and he has been his whole working life.*

- a. Safari a-mu-<sup>✓</sup>**ba** mu-ganga  
 Safari 1SM-PRS-be-<sub>BA</sub>  
 ‘Safari is a doctor’
- b. Safari a-<sup>✓</sup>**tula** mu-ganga  
 Safari 1SM-be-<sub>TULA</sub> 1-doctor  
 ‘Safari is a doctor’ Kifuliiru

## 4.4 Analysis

In this chapter, I have described three types of copular contrasts found in Great Lakes Bantu languages. The first contrast is observed in two JD50 languages, e.g., Kihavu (JD52) and Mashi (JD53), and involves the copular verbs *-li* and *-ba*. Interpretively, the two copulas encode a distinction that is not unlike a stage/individual or transient/permanent contrast; *-li* describes properties that hold of limited contexts, while *-ba* describes more general properties. The second type of contrast is attested in a subset of JD60 languages, e.g., Kinyarwanda (JD61), Kinyamulenge (JD61a), and Kirundi (JD62), and primarily involves a distinction between the verbal copula *-ri* and the invariant copula *ni*. Unlike the other contrasts discussed in this chapter, the *-ri/ni* contrast is strictly limited to

predicational sentences involving a third person subject and a locative predicate. In this context, the two generally distinguish between context-specific locative descriptions (*-ri*) and more general, known properties (*ni*). The final type of copular contrast is found in a single JD60 language, e.g., Kifuliiru (JD63), and involves a three-way contrast between the copulas *-li*, *-ba*, and *-tula*. In many ways, this contrast is similar to the limited/general (*-li/-ba*) contrast observed in the JD50 languages, but with an additional third category; while *-li* again describes properties that are limited in some way, *-tula* describes maximal properties, and *-ba* describes other more general and non-restricted properties.

From a descriptive point of view, distinguishing these three types of contrasts follows directly from their observable differences. Languages of each type employ (partially) distinct sets of copulas, e.g., *-li/-ba* vs *-li/-ba/-tula*, to make different kinds of interpretive contrasts, e.g., limited/general vs limited/general/maximal. Despite these differences however, there are also clear similarities across the three language types, particularly with respect to the form and meaning of their component copulas. In fact, all three types exhibit a nearly identical interpretive contrast between reflexes of the proto-Bantu copulas *\*-de* and *\*-bà*; reflexes of *\*-de*, *-ri/-li* express properties that hold of limited contexts, while reflexes of *\*-bá*, e.g., *-ba*, express properties that are less restricted.

In all three language types, there is one form of the copula that is used to describe properties that are in some way limited, and this is always a reflex of proto-Bantu *\*-de*. In the *-li/-ba* languages, e.g., Kihavu (JD52), Mashi (JD53), and the sole *-li/-ba/-tula* language, e.g., Kifuliiru (JD63), the copular verb *-li* (*\*-de*) is used to express temporary, localized, or context-specific properties. In the three *-ri/ni* languages, the copula *-ri* (*\*-de*) serves a similar function, albeit with locative descriptions; *-ri* expresses a spatial relation that holds of a particular context, e.g., right now, on a map, etc.. Likewise, all three languages types exhibit a reflex of *\*-bà*, e.g., *-ba*, that is used to describe properties that hold of more general circumstances. When in direct competition with *-li* — as in the *-li/-ba* languages and some *-ri/ni* languages — *-ba* often yields a permanent, individual-like interpretation. However, in Kifuliiru, which exhibits a dedicated maximal copula *-tula*, *-ba* instead yields a more vague reading; it describes properties that are non-limited but also non-maximal, e.g.,

kind-level properties.

Given the similar interpretive profile of these cognate copulas, I will propose a unified copula semantics for the Great Lakes Bantu languages under discussion. Specifically, I will argue that the three types of copular contrasts we observe in these languages can all be derived from the same four copular forms: *-ri/-li*, *-ba*, *-tula*, and *ni*. Though the exact interpretation of a single form may vary according to the language it appears in and the other copular forms it is in competition with, the general idea is that cognate copulas share a core semantic profile across languages.

#### 4.4.1 Approach

Though complicated by the fact that there are four distinct forms to account for, the general interpretive difference between the copulas in question relates to the “size” of the context or situation in which the property ascription holds, e.g., restricted vs non-restricted. As we have seen throughout this chapter, there are many ways to conceptualize an interpretive contrast of this type. Accounts of similar copular systems in languages like Spanish have attributed such contrasts to a distinction between stage/individual-level predication (Kratzer 1995, Diesing 1992a, Arche 2006, a.o.), structural differences amongst predicates (Roy 2013), and even the presuppositional content of the copula, i.e., whether it presupposes a link to a particular situation (Clements 1988, 2006, Maienborn 2005, Deo et al. 2017). In terms of their ability to account for the Bantu copular contrasts in question however, many of these approaches fall flat.

Like what is reported of the Spanish copulas *estar/ser* (Ramsey 1894, Clements 1988), a stage/individual approach makes the wrong prediction for the Bantu copulas. Consider for example the contrast between *-li* and *-ba*, which broadly encodes a distinction between stage- and individual-level predication; *-li* describes limited or stage-level properties, while *-ba* describes general, individual-level properties. Assuming a traditional stage/individual distinction in the sense of Kratzer (1995), we would expect these two copulas to exhibit distinct distributions, with the “stage-level” copula occurring with stage-level predicates, and the “individual-level” copula with individual-level predicates. This however is not the case, as either *-li* or *-ba* may co-occur with any



predicate. A further complication for a binary stage/individual analysis is the fact that one language in our sample, e.g., Kifuliiru, has a productive three-way copular contrast. If we assume that the interpretive differences amongst copulas are derived from a binary stage/individual split, there is no way account for the three-way contrast between *-li*, *-ba*, and *-tula* in Kifuliiru.

If the Bantu copulas cannot be distinguished according to the stage/individual contrast, another possible approach is to assume the three-way predicate contrast presented in Roy (2013). At first glance, this system maps quite well to the interpretive differences we observe among Bantu copulas, particularly in Kifuliiru. The distinction between the three copular forms in the language, e.g., *-li*, *-ba*, and *-tula*, broadly coincides with the interpretive distinction between dense (situation-descriptive), non-dense (characterizing), and maximal (defining) predicates (see subsection 4.3.3). Importantly, this three-way system is also amenable to binary copular contrasts like the *-li/-ba* distinction in Kihavu and Mashi. To account for binary copular contrasts in languages like Spanish (e.g., *estar* vs *ser*) and Modern Irish (e.g., *bí* vs *is*), Roy (2013) suggests that a single copular form may appear in two different kinds of sentences; Spanish *ser* and Modern Irish *is* may appear in both characterizing and defining sentences.

	Situation-descriptive	Characterizing	Defining
Spanish	<i>estar</i>	<i>ser</i>	<i>ser</i>
Modern Irish	<i>bí</i>	<i>is</i>	<i>is</i>

Table 4.4: Copular forms and interpretations in Spanish and Modern Irish (Roy 2013: 189)

In the *-li/-ba* languages, it is possible that one copular form may similarly serve two interpretive functions. In this case, *-ba* — which yields general, unrestricted, or even permanent interpretations — behaves like *ser* and *is* in that it appears in both characterizing and defining sentences. The effect of this is that characterizing and defining interpretations are not distinguished in the copular systems of these languages. Instead of a three-way interpretive contrast, there is distinction between situation-descriptive and non-situation-descriptive interpretations, i.e., a limited vs non-limited contrast.

Despite this flexibility, fully adopting the analysis presented in Roy (2013) for the Bantu copula

	Situation-descriptive	Characterizing	Defining
Kihavu, Mashi	<i>-li</i>	<i>-ba</i>	<i>-ba</i>

Table 4.5: Copular forms and interpretations in Kihavu and Mashi à la Roy (2013)

data poses significant issues. First, it is unclear whether the relevant interpretive contrasts are completely captured by the situation-descriptive/characterizing/defining contrast. For Roy, situation-descriptive interpretations are associated with dense eventualities — eventualities where the relevant property holds of all perceptible subparts of the eventuality. Characterizing interpretations on the other hand are associated with non-dense eventualities — eventualities where the relevant property need not hold of all perceptible subparts of the eventuality. As such, the distinction between situation-descriptive and characterizing interpretations is a simple question of density; if the property holds of all subparts of an eventuality — however large it may be — it yields a situation-descriptive interpretation. Assuming that the Bantu copula *-li* corresponds to Roy’s situation-descriptive/dense predicate, we would predict that *-li* should be used to describe dense eventualities regardless of size. This however is not what is observed. While *-li* does describe dense eventualities, it is also sensitive to the size of the eventuality; *-li* describes limited eventualities, e.g., short time frames, particular locations, context-specific standards, etc.. One possible way to circumvent this issue is to suggest that these “limited” interpretations are a byproduct of the dense/non-dense split; if dense eventualities tend to be smaller than non-dense eventualities, then the use of a situation-descriptive (dense) copula like *-li* may yield an implicature that the eventuality is “limited”. In either case, this analysis still provides no means of identifying or defining the relevant eventuality. As it stands, there is no way to account for the fact that *-li* describes a particular eventuality (or context/circumstance/situation, etc.) that is limited according to a single contextual parameter, e.g., time, world, location, etc. (Deo et al. 2017).

A more fundamental issue with the analysis in Roy (2013) however is that it attributes the three-way interpretive contrast to syntactic differences amongst predicates. Since different interpretations are assumed to arise due to structural differences in the predicate, the copulas they are associated with tend to exhibit distinct distributions. For example, Roy takes maximal predicates to be uniquely

defined by the presence of a NumP projection (198); if NumP is present, the predicate must denote a defining property.

(198) *Predicate types in Roy (2013)*

- a. [XP] dense (situation-descriptive)
- b. [CIP[ ... ]] non-dense (characterizing)
- c. [NumP[ ... ]] maximal (defining)

If this were the case, we would expect that there to be a clear distributional distinction in the Bantu data; if a NumP predicate, e.g., ‘two teachers’, ‘three cities’, etc., is present, only a maximal (defining) copula, e.g., *-ba/-tula* should be available. Once again however, this is not what we observe. In general, there is no correlation between the presence of a NumP predicate and the form of the copula. Even in Kifuliiru all three copulas may be used with a NumP predicate, e.g., ‘two teachers’, depending on the context.

- (203) a. Johana na Maria ba-**li** ba-alimu ba-biri  
 John and Maria 2SM-be-<sub>LI</sub> 2AGR-teacher 2AGR-two  
 ‘John and Maria are two teachers’
- b. Johana na Maria ba-mu-**ba** ba-alimu ba-biri  
 John and Maria 2SM-PRS-be-<sub>BA</sub> 2AGR-teacher 2AGR-two  
 ‘John and Maria are two teachers’
- c. Johana na Maria ba-**tula** ba-alimu ba-biri  
 John and Maria 2SM-be-<sub>TULA</sub> 2AGR-teacher 2AGR-two  
 ‘John and Maria are two teachers’ Kifuliiru

Likewise, many of the other structural and/or distributional differences discussed in Roy (2013) do not hold of the Bantu data.<sup>16</sup> For example, though Roy assumes that mass nouns, e.g., ‘water’, require the addition of a Classifier projection (CIP) to function as predicates, all three Kifuliiru copulas may appear with a mass noun depending on context.

<sup>16</sup>In general, the Bantu copulas discussed here are not sensitive to the syntactic category of the predicate, with the exception of *-ri* (in certain cases; see chapter 5) in *-ri/ni* languages, e.g., Kinyarwanda (JD61), Kinyamulenge (JD61a), and Kirundi (JD62).

- (204) a. Gano ga-<sup>✓</sup>**li** ma-aaji  
 6.DEM 6SM-be-<sub>LI</sub> 6-water  
 ‘This is water’
- b. Gano ga-mu-<sup>✓</sup>**ba** ma-aaji  
 6.DEM 6SM-PRS-be-<sub>BA</sub> 6-water  
 ‘This is water’
- c. Gano ga-<sup>✓</sup>**tula** ma-aaji  
 6.DEM 6SM-be-<sub>TULA</sub> 6-water  
 ‘This is water’

Kifuliiru

Alongside the aforementioned issues regarding the meaning of “situation-descriptive”, the lack of distributional differences among the Bantu copulas suggests that an outright application of this analysis would be insufficient. Still, there remain clear similarities between the Bantu copula contrasts and the Roy’s three-way interpretive contrast, particularly as it pertains to the presence of a third, maximal-type category.

As an alternative to approaches that attribute interpretive differences to the nature or structure of the predicate, accounts like Clements (1988, 2006) and Maienborn (2005) instead attribute them to the semantics/pragmatics; interpretively distinct copulas are treated as presuppositional variants. In Spanish for example, the use of *estar* to describe limited situations/contexts is encoded via a presupposition that *ser* lacks. Specifically, *estar* presupposes a link to a particular situation; by using *estar*, the speaker restricts the scope of the utterance to a particular, limited context. Adapted to the Bantu copulas, this provides a straightforward account of the interpretive profile of *-li* (also its cognate *-ri*); *-li* presupposes that the property ascription holds of a particular discourse context.

Compared to a structural analysis like that in (Roy 2013), a presuppositional analysis of this type has one major positive; it avoids tying copular forms to the syntactic category of the predicate. By encoding interpretive contrasts via presupposition, copulas should be free to appear with predicates of any size or category — as is (mostly) observed. That being said, this type of analysis also faces a number of significant issues, the most notable being its inability to account for non-binary contrasts, e.g., Kifuliiru *-li/-ba/-tula*. Moreover, this analysis fails to address some of core issues related to the meaning of “situation-descriptive” interpretations discussed above. By presupposing a link

to a particular situation, this analysis provides a means of deriving the limited/restricted readings of *-li*, however it is unclear how this differs from the analysis in Roy (2013), which attributes such meanings to the presence of a dense eventuality. In both cases, there is no indication that the size of the situation/eventuality is restricted in any way, nor is there information about how the situation/eventuality is defined. Just as before, the presuppositional analysis may be too vague to capture the full interpretive range of *-li*. Like Deo et al. (2017) argues in the case of Spanish *estar*, an analysis that takes *-li* to presuppose a link to a particular discourse situation à la Maienborn (2005) would fail to capture the nuanced ways in which a “particular” situation/eventuality (circumstance) may be defined, e.g., temporally, spatially, contextually, etc..

As an alternative, Deo et al. (2017) present a modified presuppositional account that specifies the size of the relevant circumstance at which the truth of the proposition is evaluated, and the contextual parameter by which it is defined, e.g., time, world, location, etc. as part of a boundedness presupposition (see subsection 4.2.2). Adapted to the Bantu data, the boundedness approach in Deo et al. (2017) provides a succinct account for the fact that *-li* describes properties that are not just temporally limited, but also spatially or contextually limited.

Regardless of which presuppositional account one adopts, each relies on the notion of competition to pragmatically generate interpretive contrasts. By not using a stronger (more constrained/specific) copula, the speaker communicates that the described situation does not fit the criteria needed to use the stronger form. In a binary copular system, e.g., *-li* vs *-ba*, the use of the (weaker) copula that lacks the restricted/boundedness presupposition (e.g., *-ba*) would trigger an implicature that the property is not limited. By choosing *-ba*, the speaker communicates that the property ascription holds of a non-limited context, leading to a “general” or “permanent” property interpretation. In a three-way contrast like Kifuliiru however, this sort of presuppositional system would not work, as there would be no way to differentiate a third copula, e.g., the maximal copula *-tula*. To capture a three-way distinction in this system, there would need to be a third presuppositional variant, e.g., one that involves a maximal circumstance/eventuality.

The following table provides a concise comparison of the various approaches outlined above.

Though each has its own merits, no single approach can account for the range of Bantu copular contrasts discussed in this chapter.

Account	Binary contrasts	Ternary contrasts	No distributional restrictions	Specifies size of situation/circumstance	Specifies contextual parameter
Stage/Individual (Kratzer 1995)	✓	✗	✗	✗	✗
Structural (Roy 2013)	✓	✓	✗	✗	✗
Presuppositional (Maienborn 2005)	✓	✗	✓	✗	✗
Boundedness (Deo et al. 2017)	✓	✗	✓	✓	✓

Table 4.6: Comparing analyses of interpretive contrasts among copulas

From this, it is clear that a comprehensive analysis must involve some components of each approach to fully capture the data. With this in mind, I will provide a pragmatically-oriented analysis for the Bantu copulas that largely follows the presuppositional “boundedness” analysis presented in Deo et al. (2017), albeit with the addition of a maximality presupposition for *-tula*.

#### 4.4.2 Formal System

The current proposal serves as extension of the analysis presented in Deo et al. (2017) that includes the notion of maximality as operationalized in Roy (2013). In what follows, I provide a simplified semantic background for the analysis based on that described in Deo et al. (2017).

I follow Deo et al. (2017) in assuming a Kaplanian theory of sentence meaning in which the content and truth-value of a sentence *S* are determined by the context *c* in which it is used. The context of use determines the proposition expressed by *S* as well as the reference of its indexical expressions, e.g., an R-expression like “John”. Importantly, the context of use also determines the circumstances that are relevant for evaluating the truth of the sentence; the context fixes the settings of relevant discourse parameters, e.g., time, location, etc., that are crucial for determining propositional truth MacFarlane (2014), Deo et al. (2017).

For Deo et al. (2017), the truth-value of  $S$  is evaluated at a circumstance of evaluation  $i$ , which is defined as the set of parameters relevant for evaluating propositional truth. They assume that propositions are not only evaluated relative to possible worlds ( $w$ ), but also times ( $t$ ), locations ( $l$ ), agents ( $a$ ), and contextual standards ( $c$ ). With this in mind, the ontology includes the following elements: a non-empty set of worlds  $W$  (type  $s$ ), a non-null set of temporal intervals  $T$  (type  $\iota$ ), a set of ordered spatial regions  $L$  (type  $r$ ), and a set of objects  $D$  (type  $e$ ) (Deo et al. 2017: 21). A single circumstance of evaluation corresponds to a tuple of five contextual parameters  $\langle w, t, l, a, d_c \rangle$ , where  $w \in W$ ,  $t \in T$ ,  $l \in L$ ,  $a \in D$ , and  $d_c$  is determined according to a “delineation function” that maps gradable predicates to a contextual standard in a context  $c$  (see Deo et al. 2017: 21).

Circumstances of evaluation  $i$  are accessible from discourse contexts  $c$  via the *Circ* function. This function takes as its complement a context  $c$  and returns the set of circumstances of evaluation that are relevant to evaluating the truth of  $S$ . Specifically, *Circ*( $c$ ) returns a set of circumstances that differ minimally in the values of their contextual parameters, e.g., world, time, location, etc., and whether  $P(x) = 1$ . A particular circumstance  $i \in \text{Circ}(c)$  is simply one set of possible values for the five contextual parameters. In this sense, a context  $c$  — e.g., the utterance context  $c_0$  — is itself a potential circumstance of evaluation.

Importantly, Deo et al. (2017) assume a strength relation ( $\succ_p$ ) on *Circ*( $c$ ) that orders circumstances by strength along a single parameter. Given two circumstances  $i \in \text{Circ}(c)$  and  $i' \in \text{Circ}(c)$ ,  $i$  is at least as strong as ( $\succeq$ )  $i'$  iff:

- (205) a.  $t_i \geq t_{i'}$ , OR  
 b.  $l_i \geq l_{i'}$ , OR  
 c.  $a_i \geq a_{i'}$ , OR  
 d.  $d_{c_i} \gg d_{c_{i'}}$ , OR  
 e.  $w_i \leq_{g_r(p, w_0)} w_{i'}$ , AND  
 f. For all other parameters  $p$ ,  $p_i = p_{i'}$  Deo et al. (2017: 24)

The strength relation on *Circ*( $c$ ) serves to order circumstances that vary along a single contextual

parameter. Circumstances  $i$  and  $i'$  may exhibit different values for a single parameter, e.g., time, world, location, etc., but all other parameter values must be the same (205f). In this sense, The relevant parameter chosen for the strength relation is determined by the context  $c$ . If ordered according to the time parameter,  $i \succeq i'$  iff the time interval of  $i$  ( $t_i$ ) is identical to or includes the time interval of  $i'$  ( $t_{i'}$ ). The same general mereological relation is true for the location and agent parameter as well. With the contextual standard parameter  $d_c$ , the strength comparison relates to the degree of the standard;  $d_{c_i} \gg d_{c_{i'}}$  iff  $d_{c_i}$  assigns a higher degree for the contextual standard in  $c$  than  $d_{c_{i'}}$ . As for the world parameter, (Deo et al. 2017) adopt a Kratzerian view in assuming that strength ordering on worlds where stronger worlds are more like the actual world. This is formally expressed through a totally realistic ordering source  $g_r(p, w_0)$ , which contains the subset of propositions that are true of the actual world and causally independent of  $p$ . This ordering source effectively ensures that worlds are ordered according to how similar they are to the actual world once the proposition  $p$  and its causal ancestors are excluded. In this sense, a world  $w$  is as strong as  $w_{i'}$  iff all propositions  $q$  in  $g_r(p, w_0)$  that are true in  $w_{i'}$  are also true in  $w_i$ .

### 4.4.3 Lexical entries

Across the Great Lakes Bantu languages investigated in this project, there are four interpretively distinct forms of the copula. Different languages exhibit a different subset of these forms in present tense predicational sentences, however their general meanings are largely the same across languages. These broad meanings are reflected in the lexical entries in (206).

- (206) a.  $\llbracket \text{-li} \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s : \text{Bound}(P(x), c_0, i). i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$
- b.  $\llbracket \text{-tula} \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s : \text{Max}(P(x), i). i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$
- c.  $\llbracket \text{-ba} \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s. i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$
- d.  $\llbracket \text{-ni} \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle}. P(x)(c_0)$

Though they are not directly in competition in any given language, the copulas in (206) are listed



by strength/complexity; use the more complex one first. The two strongest copulas, *-li* and *-tula*, are more complex than the remaining copulas *-ba* and *ni* since they each involve a presupposition.

Just like Spanish *estar* in Deo et al. (2017), the copula *-li* (realized as *-ri* in some languages) describes properties that are boundedly true at the context of use, i.e., the utterance context/situation  $c_0$ . A proposition  $P(x)$  that is true at a circumstance  $i$  accessible from the utterance context  $c_0$  is said to be boundedly true at  $c_0$  iff there are other accessible circumstances  $i'$  where  $P(x)$  is false, and there is no larger accessible circumstances  $i''$  where  $P(x)$  is true. In prose, this states that the property ascription expressed by *-li* holds of a restricted circumstance/situation; since there exist other circumstances  $i'$  where  $P(x)$  does not hold, and there exist no larger circumstances  $i''$  where  $P(x)$  is also true, then  $P(x)$  must only be evaluated true at  $i$ . Moreover, since  $\text{Circ}(c_0)$  generates a set of alternative circumstances used to evaluate the truth of the proposition in which alternatives vary along a single contextual parameter determined by the context of use  $c_0$ , the circumstance  $i$  at which  $P(x)$  is true may be limited in various ways, e.g., temporally, spatially, etc.. By using *-li*, the speaker restricts the proposition to a particular circumstance  $i$  that is limited by a single contextual parameter.

For example, if the speaker deems the relevant contextual parameter in  $c_0$  to be the time parameter, then  $\text{Circ}(c_0)$  would generate a set of alternative circumstances that minimally differ in their respective time frames, e.g.,  $i$  spans 2:00 p.m. - 2:05 p.m.,  $i'$  spans 1:54 p.m. - 1:59 p.m.,  $i''$  spans 1:00 p.m. - 3:00 p.m., etc.. Since *-li* encodes a boundedness presupposition, its use would express that the truth of  $P(x)$  is restricted to the circumstance  $i$ , i.e., the time frame 2:00 p.m. - 2:05 p.m.. In this case, the boundedness presupposition would require that  $P(x)$  be false at another non-overlapping circumstance  $i'$ , i.e., 1:54 p.m. - 1:59 p.m., and also not be true of a larger circumstance  $i''$ , i.e., 1:00 p.m. - 3:00 p.m.. As such, if alternative circumstances vary temporally, *-li* yields a temporary interpretation; the truth of  $P(x)$  is restricted to the temporal interval of  $i$ . By allowing the relevant contextual parameter to be determined from the utterance context, the various interpretive flavors of *-li*, e.g., to express a temporary property, a localized property, a context-specific property, etc., fall out naturally from the fact that the circumstance  $i$  at which  $P(x)$

is evaluated as true can be limited in different ways.

In contrast to *-li*, the copula *-tula* describes properties that are maximally true of a circumstance of evaluation *i* (206b). I follow Stechow (1996) and Roy (2013) in assuming that a maximal eventuality — or in this case, a maximal circumstance — is the largest eventuality (circumstance) at which  $P(x)$  is true. This maximality condition is expressed via presupposition; *-tula* presupposes that there are no larger circumstances  $i'$  such that  $i$  is a proper part of  $i'$  and  $P(x)(i) = 1$  and  $P(x)(i') = 1$ . Note that this maximality presupposition is identical to the second condition of the boundedness presupposition in Deo et al. (2017). In a sense then, the sole difference between the boundedness presupposition encoded by *-li* and the maximality presupposition encoded by *-tula* is that the former also requires there to be accessible circumstances where  $P(x)$  is false. It is this difference that ensures a more permanent/individual-like flavor for *-tula*; since there are no accessible circumstances where  $P(x)$  is false, the circumstance  $i$  where  $P(x)$  is true could be quite large.

Once again, since alternative circumstances generated by  $\text{Circ}(c_0)$  vary minimally along a single contextual parameter, the bounds of the relevant circumstance may be defined in various ways. If these alternatives were to vary along the time parameter, the maximality presupposition would express that the time frame of a circumstance  $i$  is the largest time frame where  $P(x)$  is true. In principle, a circumstance  $i$  could constitute a time frame of any size, from a span of a few minutes to a span of many years or decades. In practice however, the use of *-tula* tends to trigger readings involve longer (or otherwise ‘larger’) circumstances due to pragmatic competition with *-li* (see subsection 4.4.6).

Unlike both *-li* and *-tula*, the copula *-ba* does not involve an additional presuppositional component. In all other ways however, it is identical to the previous copulas; *-ba* asserts that  $P(x)$  is true at a circumstance of evaluation  $i$  that is accessible from the utterance context  $c_0$  (206c). In this sense, *-ba* functions as a sort of default copula in that it is used in cases where *-li* or *-tula* would be too strong. Importantly, this aligns with the overall distribution of *-ba* across Bantu languages, as it appears in a much broader set of environments than the other copulas discussed here, including irrealis environments.

With respect to interpretation, the various meanings and uses of this copula are derived entirely pragmatically through competition. In a language featuring *-ba* and a more complex copula, e.g., *-li/-tula*, the use of *-ba* generates an implicature that the alternative copula is too strong. If chosen over *-li* for example, this would generate an implicature that  $P(x)$  is not boundedly true, and thus yield a more general, permanent, or individual-like reading. Given its function as a default copula, the exact meaning of *-ba* depends on the copular forms it is in competition with, and therefore the language. This is the reason why the interpretive profile of *-ba* differs when part of a binary system (e.g., *-li/-ba*) as opposed to a ternary system (e.g., *-li/-ba/-tula*).

Lastly, the copula *ni* is treated as a simple copula of predication (206d). As compared to the other copulas, *ni* is very simplistic in meaning, as it does not generate a set of alternative circumstances of evaluation accessible from the utterance context  $c_0$ . Instead, I suggest that *ni* directly asserts  $P(x)$  of the utterance context  $c_0$  itself. In this sense, *ni* expresses that the proposition  $P(x)$  holds of the utterance context, which necessarily includes the speaker and addressee's shared knowledge about the actual world.

In the *-ri/ni* languages (see subsection 4.3.2), the use of *ni* over *-ri* is associated with more general, well-known property descriptions, e.g., facts about the world. Under the proposed analysis, this interpretive effect arises because *ni* ties the truth of the proposition directly to the utterance context  $c_0$ . Assuming that  $c_0$  includes the discourse participants' shared knowledge about the actual world, *ni* could be used to describe propositions that are true at  $c_0$ , including those that extend beyond the immediate utterance context. Once again however, the full interpretive profile of *ni* is largely dependent on pragmatics; the use of *ni* instead of an alternative copular form triggers an implicature that the alternative forms are too strong, and therefore are not applicable.

It is important to note here that the lexical entries in (206) are meant to capture the general interpretive profile of the relevant copula forms irrespective of the language(s) they appear in. As discussed, the full interpretive behavior of each form depends heavily on the language it appears in and the other copular forms it competes with. In the following subsections, I will illustrate how the proposed meanings interact to yield the various interpretive contrasts that characterize the *-li/-ba*,

*-ri/ni*, and *-li/-ba/-tula* contrasts.

#### 4.4.4 *-li vs -ba*

For the two languages that exhibit a “true” *-li/-ba* contrast, e.g., Kihavu (JD52) and Mashi (JD53) (see subsection 4.3.1), there are only two copular forms in competition: *-li* and *-ba*. As in (206), these two copulas differ in strength, with *-li* being the stronger (i.e., more specific) of the two; *-li* includes an additional boundedness presupposition that is not present with *-ba*.

- (207) a.  $\llbracket -li \rrbracket = \lambda P_{\langle s,et \rangle} \lambda x_{\langle s,e \rangle} \lambda i_s : \text{Bound}(P(x),c_0,i). i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$   
b.  $\llbracket -ba \rrbracket = \lambda P_{\langle s,et \rangle} \lambda x_{\langle s,e \rangle} \lambda i_s. i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$

Interpretively, the *-li/-ba* contrast distinguishes properties that hold of a limited context from properties that hold of a more general context. The copula *-li* is used to describe limited properties, e.g., temporary properties, localized properties, context-specific properties, etc.. Following (Deo et al. 2017), these restrictive interpretations arise from the boundedness presupposition on *-li*, which ties the truth of the proposition to a circumstance of evaluation *i* that is limited by a single contextual parameter, e.g., time, world, location, etc..

In contrast, the copular form *-ba* is used to describe more general properties, and is often interpretively associated with permanent, individual-like readings. Unlike *-li*, the lexical entry for *-ba* in (207) does not include an additional presuppositional component, making it the less complex of the two forms. The interpretive effects of *-ba* fall out from this difference. Assuming that these two forms are in competition, and that speakers generally adhere to the Gricean maxim of quantity, i.e., be as informative as possible, the various meanings of *-ba* can be attributed to the pragmatics. By using *-ba* instead of *-li*, the speaker communicates that the proposition must not be boundedly true. This in turn gives rise to an implicature; since the proposition is not boundedly true, it must be true of a more general circumstance. How this “general” circumstance is portrayed depends on which contextual parameter is deemed relevant when generating the set of alternative circumstances. If it is the time parameter, *-ba* yields a permanent interpretation. If it is the location parameter, *-ba*

yields a non-localized interpretation. If it is the world parameter, *-ba* yields an expected/known property interpretation, and so on.

#### 4.4.5 *-ri vs ni*

In the *-ri/ni* languages, e.g., Kinyamulenge (JD61a), and Kirundi (JD62), there are three copular forms in competition: *-ri*, *-ba*, and *ni* (see subsection 4.3.2). These three forms are heavily constrained by morphosyntactic environment, and there can be at most two forms in competition in a single environment. Just as before, the three copulas are listed by strength, with *-li* being the strongest of the three due to its boundedness presupposition. Since the copulas *-ba* and *ni* are never in contrast in any of the three languages, they are assumed to be unordered.

- (208) a.  $\llbracket -ri \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s : \text{Bound}(P(x), c_0, i). i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$   
 b.  $\llbracket -ba \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s. i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$   
 c.  $\llbracket ni \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle}. P(x)(c_0)$

The primary interpretive contrast in these languages is that between *-ri* and *ni*. This contrast occurs in the presence of a locative predicate and certain location-denoting subjects, e.g., the city ‘Kigali’. In this environment, the two copulas distinguish context-specific locative descriptions and known properties. With *-ri*, the locative description is interpreted relative to a particular context, e.g., on a map, in a taxonomic set, what can be seen, etc.. However with *ni*, the locative description is interpreted as a widely known property of the subject.

- (175) *Context: You have never heard of Kigali or Rwanda before, but you see an official-looking map of East Africa and there is a city named ‘Kigali’ marked inside the borders of ‘Rwanda’.*

Kigali #ni/√i-ri	mu	Rwanda	
Kigali be <sub>NI</sub> /9SM-be- <sub>RI</sub>	18LOC	Rwanda	
‘Kigali is in Rwanda’			Kinyamulenge

- (176) *Context: You just got back from visiting Kigali, and you are listing all of the things you know about the city.*

Kigali ✓*ni/#i-ri* mu Rwanda  
 Kigali *be<sub>NI</sub>/9SM-be<sub>RI</sub>* 18LOC Rwanda  
 ‘Kigali is in Rwanda’

Kinyamulenge

Aside from the locative restriction, the interpretive profile of *-ri* closely aligns with the behavior of its cognate *-li* in the *-li/-ba* languages. For this reason, I treat *-ri* (208a) as a semantically identical allomorph of the copula *-li* seen previously. Just like *-li*, *-ri* encodes a boundedness presupposition that restricts the truth of the proposition to a particular circumstance of evaluation. Given its co-occurrence with locative predicates, this presupposition generally results in *-ri* yielding spatially-localized readings. In (175), the use of *-ri* restricts the truth of the proposition in Rwanda(Kigali) to the map; the speaker comments solely on the location of ‘Kigali’ relative to the map.

On the other hand, the interpretive effects of *ni* arise from a combination of its meaning and pragmatics. Unlike *-ri*, the copula *ni* in (208) does not involve any presuppositional component, nor the function  $\text{Circ}(c_0)$ ; instead, *ni* directly asserts the proposition of the utterance context  $c_0$ . Under the assumption that  $c_0$  contains information about the speaker and addressee’s shared knowledge, this alone could generate the “known-property” readings in examples like (176); one property that we know to be true of Kigali is that it is in Rwanda.

Since the utterance context  $c_0$  is a possible circumstance of evaluation, and would therefore be included as a member of the set generated by  $\text{Circ}(c_0)$ , one question that arises relative to the proposed analysis is why *ni* does not involve the  $\text{Circ}$  function. The motivation for this analysis is twofold. First, without  $\text{Circ}$  there are no alternative circumstances of evaluation; by using *ni*, the speaker simply asserts that  $P(x)$  is true at the utterance context. The absence of such alternatives is reflected in the “known property” uses of *ni*, which describe the proposition  $P(x)$  as one member of a set of propositions that are true at  $c_0$ , i.e., one thing that is known to be true of  $x$  is  $P$ . In this context, the truth of  $P(x)$  is not comparatively evaluated at multiple circumstances in order to restrict it to a particular circumstance (as is the case with the limited copula *-ri*), but rather directly asserted of the utterance context in isolation. In this sense, the relevant alternatives are not circumstances, but rather propositions;  $P(x)$  is true at  $c_0$ , but there may be other propositions that are also true at  $c_0$ . The second motivation for this analysis is that it accounts for the strictly present tense distribution

of *ni*. In the absence of the Circ function, there are no accessible alternative (temporally ordered) circumstances at which the truth of P(x) can be evaluated. As such, *ni* can only assert the truth of P(x) at the utterance context  $c_0$ , i.e., in the present. Importantly, this aligns with the observation in Güldemann (2003) that cognates of *ni* in other Bantu languages are often linked to the present with some even serving as overt markers of present progressive aspect.

Another important aspect of this analysis is its pragmatic component. In isolation, the proposed semantics for *ni* does not work, as there is no way to ensure that P(x) holds beyond the immediate utterance context in “known-property” readings. As it stands, the meaning of *ni* should allow it to describe properties that are strictly true of the utterance context, e.g., a limited property, contra (176). This issue is resolved however if some of the meaning of *ni* is derived pragmatically; the use of *ni* instead of *-ri* generates an implicature that P(x) is not boundedly true, and therefore may be true of a more general context.

In addition to the *-ri/ni* contrast, the three JD60 languages in question also exhibit a distinction between *-ri/-ba* in one particular context where *ni* is unavailable, namely present tense predicational sentences featuring a locative predicate and a class 1 (human) subject. Interpretively, the contrast between *-ri/-ba* in this environment is identical to the *-li/-ba* contrast (see subsection 4.4.4), though limited to locative descriptions; *-ri* expresses a limited locative description, e.g., a temporary location, while *-ba* expresses a more general locative description, e.g., a permanent location. Since *ni* is unavailable in this environment, this contrast is derived in the same way as the *-li/-ba* contrast. The limited interpretation of *-ri* arises from the boundedness presupposition, while the general interpretation of *-ba* arises via pragmatic competition.

Before moving on, there are two issues related to the proposed analysis that need to be addressed. The first issue is that the analysis in its current form does not fully account for the restricted morphosyntactic environment in which the *-ri/ni* (and *-ri/-ba*) contrast occurs, e.g., present tense sentences featuring a third person subject and a locative predicate. While the present tense restriction can be accounted for by tying *ni* to the utterance context  $c_0$ , the current analysis offers no further explanation of this restrictive distribution. For the purposes of this chapter, I assume that this is

not an issue to be directly addressed by the semantics. This is in part motivated by the similar interpretive profiles of related copular forms across the Bantu languages investigated in this project. In an effort to provide a overarching analysis of the copular forms that captures their use across languages, the proposed semantics for the JD60 copulas do not impose any restrictions on the category of the predicate. However, as I discuss in chapter 5, it is possible that the correlation between locative predicates and *-ri* arose from its ability to describe locationally-delimited property ascriptions.

The second issue to be addressed relates to the complementary distribution of *ni* and *-ba*. Since these two copular forms never appear in the same environment, and both participate in a contrast with the “limited” copula *-ri*, it is tempting to treat them as equivalents. There are however a few reasons to distinguish *ni* and *-ba*. The first is the two exhibit rather distinct distributions; *ni* is largely limited to third person predicational sentences in the present tense, while *-ba* appears in the present tense, future tense, infinitival mood, and subjunctive mood, among other contexts. The second reason to distinguish the two is that they yield distinct interpretations of locative predications when compared to *-ri*; while *ni* yields “known” property readings, e.g., facts about the world, *-ba* tends to yield more permanent interpretations, e.g., “John is in Kigali” to mean that John *lives* in Kigali. Though similar, I argue that these interpretations are distinct, and signal that *ni* and *-ba* assert  $P(x)$  of different circumstances. With *ni*, the speaker makes a direct statement about the utterance context; the proposition  $P(x)$  is true at  $c_0$ . If the utterance context includes the shared world knowledge of the discourse participants, this could involve known facts about the world, which would necessarily be true at  $c_0$ . With *-ba* however, the speaker does not necessarily make a statement about the utterance context, though they might. Instead, the speaker asserts that there is some circumstance  $i$  accessible from  $c_0$  at which  $P(x)$  is true. This circumstance could in theory be the utterance context  $c_0$ , but it need not be; it may also be a “larger” circumstance, e.g., a human lifetime, a world-level circumstance, etc.. For these reasons, I opt to distinguish *ni* as a simple copula of predication that is tied to the present tense, and asserts the truth of the proposition at the utterance context.



#### 4.4.6 *-li vs -ba vs -tula*

In Kifuliiru (JD63), there are three distinct copular forms that are mutually contrastive: *-li*, *-ba*, and *-tula*. As in the other copular systems, these forms are taken to differ in strength. The first form *-li* is the strongest of the three, as it includes a boundedness presupposition. The next strongest form *-tula* also involves a presuppositional component, e.g., a maximality presupposition, albeit one that is less complex. The weakest of the three forms is the copula *-ba*, which does not involve a presuppositional component.

- (209) a.  $\llbracket -li \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s : \text{Bound}(P(x), c_0, i). i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$
- b.  $\llbracket -tula \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s : \text{Max}(P(x), i). i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$
- c.  $\llbracket -ba \rrbracket = \lambda P_{\langle s, et \rangle} \lambda x_{\langle s, e \rangle} \lambda i_s. i \in \text{Circ}(c_0) \wedge P(x)(i) = 1$

The interpretation of Kifuliiru *-li* closely aligns with the interpretation of its cognates in the other languages discussed previously. In general, *-li* is used to describe properties that are true of limited circumstances, e.g., temporary properties, localized properties, etc.. As before, this is modeled via a boundedness presupposition that serves to restrict the truth of the proposition to a particular circumstance.

In contrast, the copula *-tula* is characterized as a maximal copula in that it is used to describe properties that are true of maximal circumstances, e.g., permanent properties. This is formally expressed via the presence of a maximality presupposition, which presupposes that the circumstance *i* at which *P(x)* is evaluated is the largest circumstance where *P(x)* is true. This presupposition minimally differs from the boundedness presupposition on *-li* in that it does not require there to be accessible alternatives where *P(x)* is not true.

To illustrate the difference between the two, let us again consider a set of alternatives that differ only along the time parameter. To satisfy the boundedness presupposition on *-li*, two conditions must be met: first, there must be an accessible circumstance *i'* such that  $P(x)(i') = 0$ , and second, there must be no larger circumstance *i''* such that  $P(x)(i) = 1$  and  $P(x)(i'') = 1$ . For example, if the relevant circumstance *i* spans from 2:00 - 2:05 p.m. and  $P(x)(i) = 1$ , the first condition would require

that there be an alternative circumstance  $i'$  that spans a different, non-overlapping time-frame where  $P(x)$  is not true, e.g.,  $i'$  spans from 1:54 - 1:59 p.m. and  $P(x)(i') = 0$ . Moreover, the second condition of this presupposition would require that there be no larger circumstance where  $P(x)$  is also true, e.g., if  $i$  spans from 2:00 - 2:05 p.m. and  $P(x)(i) = 1$ , and there is an alternative circumstance  $i''$  that spans from 1:00 - 3:00 p.m., then  $P(x)(i'') = 0$ . Together, these conditions ensure that the truth of  $P(x)$  is limited to the circumstance  $i$ .

In contrast, the maximality presupposition associated with *-tula* involves only a single condition. In fact, it specifically involves the second condition of the boundedness presupposition; there must be no larger circumstance  $i'$  such that  $P(x)(i) = 1$  and  $P(x)(i') = 1$ . By using *-tula*, the speaker asserts that the circumstance  $i$  is the largest circumstance where  $P(x)$  is true. As such, the crucial difference with *-li* is that *-li* requires there to be an alternative circumstance where  $P(x)$  does not hold. Assuming that the two copulas are in competition, this difference allows their respective interpretations to be generated pragmatically. By using *-li*, the speaker identifies alternative circumstances where  $P(x)$  is false, leading to a limited property reading. However, by using the weaker copula *-tula*, there is an implicature that there are no alternatives where  $P(x)$  is false, thus yielding a more permanent-oriented interpretation.

Unlike the other Bantu languages discussed in this chapter, there is a third copular form in Kifuliiru that exists in contrast with *-li* and *-tula*, namely *-ba*. Interpretively, this form tends to be correlated with general, permanent-like interpretations like *-tula*, however in a more limited way. In the presence of an individual-level predicate, e.g., 'tall', *-ba* may only appear with a generic subject, in which case it yields a kind-reading. While distinct from the interpretative behavior of *-ba* in the other Bantu languages discussed, I argue that the meaning of Kifuliiru *-ba* follows from it being in competition with both *-li* and *-tula*.

Formally, I treat Kifuliiru *-ba* (209) the same as *-ba* in other languages; it is identical to *-li/-tula*, but does not feature a presuppositional component. As such, the meaning of *-ba* is once again derived via pragmatic competition. Since there are two stronger copular forms, this pragmatic effect yields a distinct meaning relative to other languages. When a speaker uses *-ba*, it generates an

implicature that the proposition is not boundedly true, nor maximal. This allows *-ba* to describe properties that may hold of various overlapping and non-overlapping circumstances, as opposed to properties that hold of a single limited circumstance (*-li*), or a single maximal circumstance (*-tula*). One way that this may manifest is as a kind-reading; *-ba* expresses that  $P(x)$  is generally true in that it holds of many accessible circumstances, e.g.,  $i, i', i''$ , but crucially not all. The more general permanent readings of *-ba* arise in a similar way as well; if the circumstance  $i$  at which  $P(x)$  is evaluated is both non-minimal and non-maximal, it may presumably be of any size, whether large or small. In this sense, the “permanent” flavor of *-ba* may arise simply because the stronger, “limited” copula *-li* was not used<sup>17</sup>

#### 4.5 Discussion and conclusion

The semantic analysis I provide for the Great Lakes Bantu copulas discussed in this chapter synthesizes the ideas of previous approaches to interpretive contrasts in copular systems. Building on the work of Clements (1988, 2006), Maienborn (2005), Roy (2013) and Deo et al. (2017), I present an analysis that ties interpretive differences among copulas to the situation (or rather, circumstance) at which the truth of a proposition is evaluated. The core contrast between the relevant copular forms is expressed as a presupposition (or lack thereof) that determines how restrictive said circumstances are; some forms restrict the truth of the proposition to a single, limited circumstance (e.g., *-li*), while others assert the truth of the proposition of more general circumstances (e.g., *-ba*), or even maximal circumstances (e.g., *-tula*). The specific interpretive contrasts observed in each language, e.g., temporary vs permanent, localized vs non-localized, etc., arise pragmatically from competition among the subset of copular forms in that language.

With respect to the broad theoretical question regarding the semantic contribution of the copula, this analysis follows a long line of research in making the claim that copular elements can encode meaning (Ramchand 1997, Greenberg 1998, Green 2000, Maienborn 2005, Bochnak et al. 2011, Deo et al. 2017, a.o.). More specifically, this analysis makes a claim about the kinds of meanings

---

<sup>17</sup>This reflects the observation in Roy (2013: 86) that “non-dense predicates may be long-lasting or not”.

that copular systems may encode. Given the general availability of the distinct copular forms discussed in this chapter, I reject a traditional Kratzerian stage/individual distinction in copular clauses. Following Clements (1988, 2006), Maienborn (2005), and Deo et al. (2017), I instead assume that apparent stage/individual-level readings arise from a distinction between property ascriptions that hold of particular situations/circumstances, and those that hold of more general contexts. Adopting the base semantics of Deo et al. (2017), I moreover suggest that stage/individual or temporary/permanent-type contrasts represent just one subtype of the limited vs non-limited contrast. Under the limited/non-limited approach proposed here, the situation/circumstance at which the property ascription is asserted may be limited according to various contextual parameters, e.g., time, world, location, etc.. Like Deo et al. (2017), I assume that many of the seemingly distinct copular contrasts reported across languages are simply subtypes of a more general limited vs non-limited contrast. Classic stage/individual (or temporary/permanent) contrasts arise when the limited/non-limited contrast is temporally defined; stage-level readings correspond to property ascriptions that hold of temporally limited circumstances, while individual-level readings correspond to property ascriptions that hold of non-limited circumstances. Likewise, locational contrasts like that proposed in the Kinyarwanda copular system (Jerro 2015) arise when this contrast is spatially defined; localized readings occur when property ascriptions hold of spatially limited circumstances. Given its ability to account for these various types of interpretive contrasts, I make the strong claim that the limited/non-limited contrast is a core interpretive contrast that may be encoded in copular systems.

With that being said, though I argue that the relevant Bantu copulas are meaningfully distinct, this does not necessarily mean that the functional element that is the “copula” is a semantically meaningful element. For those who assume that the copula is a meaningless predicative element, interpretive contrasts like that between Spanish *estar/ser* are attributed to a separate dedicated predicational head in the syntax, e.g., Pred (Heycock and Kroch 1999, Adger 2003, Baker 2003, Balusu 2014, Myler 2016). Under this view, there are two different notions of the term “copula”. The first corresponds to the copular head projected in the syntax, e.g.,  $v_{BE}$  (Myler 2016, 2018)). In

this case, the copula is assumed to be meaningless; it serves only to integrate the predication relation mediated by Pred with higher functional heads in the clause. The second notion of “copula” is the more familiar of the two, as this describes the predicative element that is actually pronounced. In this case, the copula can be meaningful in the sense that copular forms can be interpretively distinct in a single language, e.g., Spanish *estar/ser*. Importantly for these accounts, this interpretive difference is assumed to be derived from the existence of two Pred heads: Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub>. In languages like Spanish, the particular Pred head that projects determines the form (and interpretation) of the copula, e.g., the copula is pronounced as *estar* with Pred<sub>STAGE</sub>, but as *ser* with Pred<sub>INDIV</sub>.

In theory, the semantic analysis I propose for the relevant Bantu copulas is amenable to this style of approach, albeit with some slight modifications. The first thing that needs to be addressed for this analysis to be workable is that there are only two variants of the Pred head (e.g., Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub>), but up to three interpretively distinct copular forms in some languages (e.g., Kifuliiru). Assuming that the existing Pred variants distinguish limited (Pred<sub>STAGE</sub>) and non-limited (Pred<sub>INDIV</sub>) interpretations in the proposed analysis, a third Pred variant would need to be added to maximal interpretations associated with Kifuliiru *-tula*. We would also need to modify the PredP approach to account for the appearance of the invariant copula *ni*, which differs substantially from the other copular forms. As I will discuss in chapter 5, it is unclear whether *ni* instantiates the same structure as other (verbal) copular forms, making it difficult to adapt the PredP approach to *ni*.

There are many other morphosyntactic issues that need to be addressed to determine whether the PredP approach is a viable option for the semantic analysis presented here. One major issue discussed earlier in this chapter is environmental restrictions on copular contrasts. In some cases, copular contrasts are restricted to highly specific morphosyntactic environments and show sensitivity to factors such as the person features of the subject, tense environment, and the category of the predicate. For example, in the *-ri/ni* languages, the contrast between *-ri* and *ni* is restricted to present tense predicational sentences featuring a third person subject and a locative predicate. Of these factors, only the tense specification is directly accounted for in the proposed semantics; *ni* is “tied” to the utterance context and therefore limited to the present tense. The other relevant factors would

need an alternative explanation, e.g., morphosyntactically-driven co-occurrence restrictions and/or diachronic changes that resulted in the correlation between certain forms and their environment. I will return to the discussion of structure and the form of the copula in chapter 5. What is important for this chapter is the observation that there are salient interpretive contrasts among the copulas in the Great Lakes Bantu languages discussed, and that each of these contrasts involves a general distinction between property ascriptions that hold of limited circumstances, and those that do not.

## Chapter 5

### Form

One of the central goals of any serious inquiry into the nature of non-verbal predication is to explain the form of the copula; if multiple copular forms are attested in a given language, what factors determine when one form appears instead of another?

In the previous chapter, I showed that different forms of the copula in some languages encode distinct interpretations of the predication relation they help realize; in these languages, the form of the copula is heavily dependent on meaning. That being said, there are many languages featuring complex copular systems (i.e., multiple copular forms) that are not sensitive to interpretation in the same way. As I have shown in previous chapters (see section 2.3), there are a range of other factors that condition the realization of the copula crosslinguistically, including copular clause type, tense-aspect-mood environment, polarity, subject person features, and predicate category. Importantly, languages differ in which factors — and combination(s) thereof — are relevant to the surface form of the copula.

For example, the form of the copula in Thai is sensitive to both copular clause type in the sense of Higgins (1979) and the categorial identity of the predicate (Wongwattana 2015). In standard “pure” predicational sentences, two forms of the copula are attested; *jù:* appears in sentences with a locative or temporal predicate, while *pen* appears in all other cases (57)-(58). In all other copular clause types, e.g., specificational, identificational and equational clauses, a third form *k<sup>h</sup>i* is used.

(57) *Predication*

K<sup>h</sup>ǎw **pen**/\***jù:** dèk  
s/he COP child  
'S/he is a child'

(58) *Predication* (PP<sub>LOC/TEMP</sub>)

Mɔ:tə:saj    **jù:/\*pen**    naj    bâ:n  
motorcycle    COP    in    house  
'The motorcycle is in the house'

(59) *Specification*

P<sup>h</sup>û:    tɕ<sup>h</sup>á<sup>?</sup>ná    **kh<sup>i</sup>/\*pen**    buak<sup>h</sup>ǎ:w    po:    pràmúk  
person    win    COP    Buakaw    Por    Pramuk  
'The winner is Buakaw Por Pramuk'

(60) *Identification*

Nî:    **kh<sup>i</sup>/\*pen**    ma:rí<sup>??</sup>ô  
this    COP    Mario  
'This is Mario'

(61) *Equation*

Sà<sup>?</sup>jǎ:m    **kh<sup>i</sup>/\*pen**    prà<sup>?</sup>t<sup>h</sup>ê:t    t<sup>h</sup>aj  
Siam    COP    country    Thai  
'Siam is Thailand'

Thai (Wongwattana 2015)

In contrast, the form of the copula in Russian exhibits a sensitivity to tense; while the zero (or pronominal) copula appears in present tense predicational sentences, only an inflected form of *byt'* (e.g., *byl*), may appear in the past tense.

(68) a. Segodnja reka            spokojna  
today    river.FEM    calm.FEM.SF  
'Today the river is calm'

b. Ivan **byl**                    goloden  
Ivan    be.PAST.MASC    hungry.MASC.SF  
'Ivan was hungry'

Russian (Adapted from Roy 2013: 119)

An even more striking alternation of this sort can be seen in the Bantu language Lamba (M54), where the person features of the subject determine which predicational strategy is employed; while an agreeing copular form *-li* is used with first/second person subjects, tonal marking of the predicate is used with third person subjects.



- (41) a. U-li                    muntu  
           2SG.SM-be-<sub>LI</sub> 1.person  
           ‘You are a person’
- b. Mù-ntu  
           1.person (lowered tone)  
           ‘S/he is a person’

Lamba (Doke 1922: 94)

The point made here is that there is a tremendous degree of cross-linguistic variation in the morphosyntactic expression of non-verbal predication and the various factors that constrain it. In some languages (e.g., those above), there are multiple strategies for encoding non-verbal predication that are in some way modulated by a particular combination of factors, e.g., copular clause type, tense-aspect-mood environment, etc.. In other languages however, one single strategy is seemingly used in all cases regardless of such factors, e.g., copular *be* in English. This variation naturally brings to light a number of typological and theoretical questions regarding the expression of non-verbal predication.

From a typological perspective, one major question is what predicational strategies and limiting factors are attested across languages; what kinds of complex copular systems are attested, and what are the factors (and combinations thereof) that play a role in determining the form of the copula/predicational strategy in these systems? From a more theoretical standpoint, copular variation raises a number of questions regarding the structure of non-verbal predication and the realization of the copula, including what head(s) the copula realizes, and whether different copular forms and/or predication strategies correspond to distinct structures. For example, one particularly prevalent question in the literature is whether different types of copular clauses involve semantically distinct copulas. The basis for this question lies in the observation that there exist multiple types of copular clauses that are structurally and interpretively distinct, with some positing four distinct types copular clauses — predicational, specificational, equational, and indentificational (Higgins 1979) — and others positing a proper subset of these. To account for the differences between clause types, some have proposed that there is more than one type of copula; in addition to a standard copula of predication that attributes properties to individuals, there is a second copula that equates the

reference of two individuals, and potentially even a third copula that equates two propositions (see discussion of Mikkelsen (2011) in subsection 3.2.2). For others however, different clause types do not arise from differences in the copula itself, but rather differences in structure of the predication relation. For these accounts, the copula is not directly involved in relating two elements in a predication relation, but is rather a meaningless element that helps integrate a predicational small clause with the extended verbal projection (Myler 2016, 2018), meaning it is uniformly present across different copular clause types. In this sense, the question of whether there are multiple semantically distinct copulas — and if so, how many — is tied to the fundamental question of what the copula is syntactically; does the copula serve a direct predicational function and realize a (dedicated) predicational head in the syntax, or does the copula serve a largely inflectional purpose and realize a higher functional head?

In this chapter, I explore these questions — among others — through an investigation of copular allomorphy in the select group of Great Lakes Bantu languages investigated in this dissertation. In the first half of the chapter, I provide a description of the copular systems of the relevant languages and show that copular allomorphy is sensitive to a number of different factors, including the (sub-)type of the copular clause, the tense-aspect-mood environment, the person features of the subject, the categorial identity of the predicate, and the interpretation of the predication relation. In so doing, I show that there are consistent patterns between the form of the copula and its morphosyntactic and interpretive environment, with cognate forms exhibiting similar distributions across languages. From these descriptive observations, I provide an analysis of copular allomorphy in the relevant Bantu languages that attempts to capture the distribution of the observed copular forms through a (broadly) generalizable set of Vocabulary Insertion rules. Ultimately I suggest that the general patterns of copular allomorphy reflect the existence of two distinct pressures on the form of the copula: one that ties the form of the copula to the interpretation of the predication relation, and another that ties the form of the copula to its morphosyntactic environment.

## 5.1 Background

### 5.1.1 Non-verbal predication in Bantu

As discussed in section 2.2, Bantu languages exhibit a wide variety of non-verbal predication strategies. A brief summary of the strategies mentioned previously can be seen below in Table 5.1.

Strategy	Example Language	Example	Translation
Verbal copula	Mongo (C61)	(16c) <b>Ba-le</b> ba-laki	‘They are teachers’
Invariant copula	Digo (E73)	(20) Mutu hiyu <b>ni</b> daktari	‘This person is a doctor’
Zero copula	Swahili (G41)	(25a) Hamisi mpishi	‘Hamisi is a cook’
Locative copula	Swahili (G41)	(27a) Yeye <b>yu-ko</b> Ukerewe	‘S/he are in Ukerewe’
Pronominal copula	Digo (E73)	(31) Kulungu <b>ndi-ye</b> wa-ngu	‘The antelope is mine’
Inflectional copula	Digo (E73)	(33) Chitabu <b>chi</b> tayari	‘The book is ready’
Possessive copula	Mashi (JD53)	(39) Johne a- <b>dwiire</b> iwazo linja	‘John has a good idea’
Tonal marking	Shona (S10)	(40) mú-nhu / <b>mù</b> -nhu	‘person’ / ‘It is a person’
Augment deletion	Lusoga (JE16)	(42) <b>o</b> -mú-géni / mú-géni	‘guest’ / ‘It is a guest’

Table 5.1: Summary of Bantu non-verbal predication strategies

Despite this tremendous degree of variation, there are nonetheless some consistent patterns observed across Bantu languages. One such pattern involves a sensitivity to tense. For example, it is regularly reported that invariant copulas, e.g., *ni* in Swahili (G41), Kinande (JD42), Kinyarwanda (JD61), etc., are restricted to the present tense — likely as a result of their inability to bear inflectional morphology (Jerro 2015, Schneider-Zioga 2018, Gibson et al. 2019). Another pattern involves an apparent sensitivity to predicate category/meaning; locative predication is often associated with variants of the verbal copula *-li* (\*-de) (Lanham 1953, Schneider-Zioga 2018), with some languages employing such forms as a dedicated locative copula, e.g., Kinyarwanda *-ri*, (Jerro 2015). Yet another pattern involves a sensitivity to matrix subject person features. Many languages employ different non-verbal predication strategies for third person subjects and first/second person subjects, e.g., Lamba (M54) uses a verbal copula with first/second persons but tonal marking on the predicate with third person subjects (Doke 1922).

One particularly relevant pattern for the purposes of this work involves a sensitivity to copular clause type à la Higgins (1979). As reported by Schneider-Zioga (2018), many Eastern Bantu languages distinguish pure predication from specification, identification, and equation using different

predicational strategies. In Kinande (JD42) for example, predicational clauses are expressed with the invariant copula (62), while specificational, identificational, and equational clauses are expressed via an agreeing relational element (glossed below as FOC) derived from the pronominal stem *-o* (63)-(65).

(62) *Predication*

Kambale **ni** mwimbi  
 Kambale be<sub>NI</sub> 1.thief  
 ‘Kambale is a thief’

(63) *Specification*

ómwira w-age **k’** ákákekulú k’ omo kisomó  
 1.friend 1AGR-1SG.POSS 12FOC 12.old woman 12LNK 18LOC 7.church  
 ky-etu  
 7AGR-1PL.POSS  
 ‘My (best) friend is a little old lady from our church’

(64) *Identification*

omúlumy’ óliá **yó** omukolo owe’ departement  
 1.man 1DEM 1FOC 1.head 1LNK 1.department  
 ‘That man is the head of the department’ (pointing)

(65) *Equation*

Eririma **ky’** ekihugo (Matthew 13:38)  
 5.field 7FOC 7.world  
 ‘The field is the world’

Kinande (Schneider-Zioga 2018)

Though sometimes described as a pronominal copula (see subsection 2.2.5) given its source as a demonstrative pronoun (Persohn 2017, Schneider-Zioga 2018), the relational element that appears in examples like (63)-(65) is generally analyzed as a focus marker rather than a dedicated copular/predicational element. In Nyakusa-Ngonde (M31), Persohn (2017) reports that the same element (what he describes as a referential demonstrative) strongly correlates with focus; though predication involving a class 3 subject most commonly involves a zero copula (30a), an agreeing *-o* pronominal, e.g., *gyo*, can appear when the subject is fronted (30b). Importantly, the same element is also reported to appear in cleft sentences in the language.

- (30) a. I-mi-piki mi-nywamu  
 AUG-4-tree 4AGR-big  
 ‘These trees are big’
- b. I-mi-piki i-gi **gyo** mi-nywamu  
 AUG-4-tree AUG-4DEM\_PROX 4DEM\_REF 4AGR-big  
 ‘These trees, they are big’ Nyakyusa-Ngonde (Adapted from Persohn 2017: 305)

Building on these observations, Schneider-Zioga (2018, 2021) suggests that the appearance of the *-o* pronominal in specificational, identificational, and equational sentences in Kinande also relates to focus: inversion structures (i.e., PREDICATE-SUBJECT configurations) arise from focus.

As I will show in section 5.2, many of the patterns discussed above, e.g., sensitivity to tense, predicate category, copular clause type, etc., are reflected to some degree in the realization of non-verbal predication in the cluster of Great Lakes Bantu languages focused on in this work.

## 5.2 Data

In this section, I will discuss the effect of various morphosyntactic and interpretive factors on the form of the copula in the Great Lakes Bantu languages at the center of this work, namely Kihavu (JD52), Mashi (JD53), Kinaymulenge (JD61a), Kirundi (JD61), and Kifuliiru (JD63). Though I discuss data from Kinyarwanda (JD61) — specifically that presented by Jerro (2015) — extensively in chapter 4, I omit any overt discussion of Kinyarwanda in this chapter since I do not have any novel data to contribute.

### 5.2.1 Interpretation

In chapter 4, I demonstrated that different copular forms encode distinct interpretations in present tense predicational clauses in the Great Lakes Bantu languages discussed in this work. In general, copular contrasts in these languages were shown to generally distinguish between property ascriptions that hold of limited circumstances, e.g., a short time frame or specific location, and those that hold of more general circumstances, e.g., a broader time frame or more general location.

Moreover, the interpretive profile of individual copular forms was shown to be largely consistent across languages, with related copular forms encoding similar meanings; reflexes of proto-Bantu *\*-de*, e.g., *-li/-ri* express limited readings, while reflexes of proto-Bantu *\*-bà* express more general, non-limited readings.

Despite the considerable overlap in their copular systems, individual languages sometimes differ in the particular forms they employ to interpretive effect. In total, there are three different combinations of contrastive forms attested in the languages under discussion. The first contrast, *-li* vs *-ba* appears in primarily in Kihavu (JD52) and Mashi (JD53), although it is also visible in certain contexts in other languages as well. In systems of this type, the two forms broadly distinguish between limited (*-li*) and non-limited readings (*-ba*). The second contrast, *-li* vs *-ba* vs *-tula*, is the only three-way contrast in the data, and is only attested in Kifuliiru (JD63). In this case, the forms again distinguish between property ascriptions that hold of limited circumstances (*-li*) and those that hold more generally, however there is an additional distinction between property ascriptions that simply hold of non-limited circumstances (*-ba*) and those that hold of maximal circumstances (*-tula*). The final contrast, *-ri* vs *ni*, occurs in Kinyarwanda (JD61), Kinyamulenge (JD61a), and Kirundi (JD62). In this case, the forms do not distinguish between limited and non-limited readings as before, but instead distinguish between limited readings (*-ri*) and “known-property” readings (*ni*). This difference is motivated by the availability of the *-ba* in certain cases where *ni* is unavailable. I argue that the meaning of *-ba* in this environment is distinct from the meaning of *ni* (see chapter 4 for further discussion).

### **5.2.2 Copular clause type**

Similar to what Schneider-Zioga and Mutaka (2015a) and Schneider-Zioga (2018, 2021) report in Kinande (JD42), many of the languages explored in this work utilize distinct predicational strategies across Higgins’ (1979) four sub-types of copular clauses. In particular, we find that Kihavu (JD52), Mashi (JD53), Kinyamulenge (JD61a), and Kirundi (JD62) exhibit the same contrast between pure predicational clauses and specificational, identificational, and equational clauses that is observed in

Kinande; while various copular forms may be used in isolation to express non-verbal predication in pure predicational clauses, the copula tends to co-occur with an agreeing focus marker in all other copular clause types.

### 5.2.2.1 JD50 languages

Kihavu (JD52) and Mashi (JD53) exhibit significant overlap in their copular systems. In pure predicational clauses, both languages employ two verbal copulas *-li* and *-ba* that differ interpretively (see subsection 4.4.4). In most other sub-types of copular clauses, both languages employ a pronominal copula like that found in Kinande.

Consider first the following examples from Kihavu. In predicational clauses, the verbal copulas *-li* and *-ba* broadly distinguish between properties that hold of a limited context and properties that hold of much broader circumstances (210).

(210) *Predication*

a. Mugisha a-**li** mu-lwala  
 Mugisha 1SM-be-<sub>LI</sub> 1AGR-sick  
 ‘Mugisha is sick’ (he has a cold)

b. Mugisha a-**ba** Muhavu  
 Mugisha 1SM-be-<sub>BA</sub> Muhavu  
 ‘Mugisha is Muhavu’ (his ethnicity)

Kihavu

In specificational and equational clauses however, a different predicational strategy is employed — focus marker (derived from the pronominal stem *-o*) obligatorily appears in such environments Schneider-Zioga (2021). Importantly, this copular element agrees with the noun class of the post-copular nominal (DP2).<sup>1</sup> This pattern coincides with what is generally reported of agreement in specificational copular clauses cross-linguistically, where agreement tends to be with the focused, post-copular nominal (Fischer 2003, Heycock 2012).

(211) *Specification*

<sup>1</sup>In some cases, the surface form of the focus marker can be transparently derived from the pronominal stem, *-o*, i.e., it surfaces as AGR-*o*. An exception to this is the class 1 pronominal copula, which surfaces as *ye* in Kihavu and Mashi. Following Schneider-Zioga (2021), I choose not to decompose this element, and instead gloss it as a monomorphemic focus marker.

- a. Gi-shambo **ye** omw-ana  
 7-thief 1FOC 1-child  
 ‘The thief is the child’
- b. O-mu-rhani **kyo** gi-shambo  
 AUG-1-murderer 7FOC 7-thief  
 ‘The murderer is the thief’

(212) *Equation*

- a. Nyina wa Mugisha **ye** nyina wa Murhulla  
 1.mother 1SM.LNK Mugisha 1FOC 1.mother 1SM.LNK Murhulla  
 ‘Mugisha’s mother<sub>i</sub> is also Murhulla’s mother<sub>i</sub>’
- b. A-ba-alimu **bo** ba-vuya  
 AUG-2-teacher 2FOC 2-cook  
 ‘The teachers<sub>i</sub> are the cooks<sub>i</sub>’

Kihavu

Identificational clauses in Kihavu largely follow the same pattern, as shown in (213a). Unlike specificational clauses however, certain identificational clauses (e.g., those involving a proximal demonstrative) may be realized with no overt copular form. As I will show, this is a property that is unique to Kihavu, as no other language in our sample permits a zero copula in this context.

(213) *Identification*

- a. Oyu u-mw-aana **kyo** gi-shambo  
 1DEM\_DIST AUG-1-child 7FOC 7-thief  
 ‘That child is the thief’ (pointing at child)
- b. Ono Mugisha  
 1DEM\_PROX Mugisha  
 ‘This is Mugisha’ (pointing at a person in a photo)
- c. Ono nyono  
 1DEM\_PROX 1.SG.PRO  
 ‘This is me’ (pointing at a person in a photo)

Kihavu

In almost all ways Mashi patterns identically to Kihavu, as it similarly employs the copular verbs *-li/-ba* in predicational sentences, and a pronominal copula in all other clausal sub-types. The sole difference between the two languages in this regard is that there is no evidence that Mashi permits a zero copula in identificational clauses.



(214) *Specification*

- a. Ci-shambo **ye** Johne  
7-thief 1FOC John  
'The thief is John'
- b. Johne **co** ci-shambo  
John 1FOC 7-thief  
'John is the thief' (answers the question "who is John?")
- c. o-mu-rhanishambo **kyo** gi-shambo  
AUG-1-murderer 7FOC 7-thief  
'The murderer is the thief'

(215) *Equation*

- a. Nyina wa Mugisha **ye** (na) nyina wa Murhulla  
1.mother 1SM.LNK Mugisha 1FOC (and) 1.mother 1SM.LNK Murhulla  
'Mugisha's mother<sub>i</sub> is also Murhulla's mother<sub>i</sub>'
- b. A-ba-alimu **bo** ba-vuya  
AUG-2-teacher 2FOC 2-cook  
'The teachers<sub>i</sub> are the cooks<sub>i</sub>'

(216) *Identification*

- a. Uyu **ye** Johne  
1DEM\_PROX 1FOC John  
'This is John' (pointing at a person in a photo) Mashi

### 5.2.2.2 JD60 languages

Unlike what is observed in Kinande and the JD50 languages Kihavu and Mashi, the distinction between the four sub-types of non-verbal predication is Kinyamulenge (JD61a) and Kirundi (JD62) is less pronounced. In fact, across the four sub-types of copular clauses, there is some degree of uniformity in both languages; all copular clause types (with the exception of certain predicational clauses that tend to employ *-ri*) require the invariant copula *ni* to be used when a third person subject is present. Nonetheless, both Kinyamulenge and Kirundi still pattern similarly to the languages mentioned previously in the sense that they distinguish pure predication from specification, equation, and identification, albeit in a slightly different way. In these two languages, non-predicational

clauses are optionally expressed via the combination of the invariant copula *ni* and the same focus marker seen in other languages, *-o*.

In Kinyamulenge, two copulas are observed in pure predicational clauses in the present tense: invariant *ni* and verbal *-ri*. Just as Jerro (2015) reports in Kinyarwanda (JD61), the contrast between these two forms is sensitive to two factors (in addition to tense), namely the person features of the subject, and the categorial identity of the predicate. The verbal copula *-ri* is used with all predicational sentences involving a first/second person subject (see subsection 5.2.3.2), and in sentences involving a third person subject and locative predicate (217b). The invariant form *ni* appears in all other third person predicational sentences.

(217) *Predication*

a. Johne **ni** mu-ganga  
 John be<sub>NI</sub> 1-doctor  
 ‘John is a doctor’

b. Kigali **i-ri** mu Rwanda  
 9.Kigali 9SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’

Kinyamulenge

Specificational, identificational, and equational clauses involve the same invariant copula *ni*, however they optionally allow a co-occurring focus marker (e.g., the pronominal copula) to appear after *ni*. As before, this focus marker typically agrees in noun class with DP2 (although this is sometimes surface opaque if DP1 and DP2 belong to the same noun class).<sup>2</sup>

(218) *Specification*

a. U-mu-aarimu **ni(=we)** Johne  
 AUG-1-teacher be<sub>NI</sub>=1FOC John  
 ‘The teacher is John’

b. Johne na Maria **ni(=bo)** a-ba-ibyi  
 John and Maria be<sub>NI</sub>=2FOC AUG-2-thief  
 ‘John and Maria are the thieves’ (answers the question “who are John and Mary?”)

(219) *Equation*

<sup>2</sup>In some cases, DP2 includes an augment, however there is no consistent pattern to be found in the data. More work is necessary to determine the role and/or necessity of the augment across copular clause types.

- a. Mama wa Johne **ni(=we)** mama wa Maria  
 1.mother 1LNK John be<sub>NI</sub>=1FOC 1.mother 1LNK Maria  
 ‘John’s mother<sub>i</sub> is Maria’s mother<sub>i</sub>’
- b. Inzu ya Johne **ni(=yo)** inzu ya Maria  
 9.house 9LNK John be<sub>NI</sub>=9FOC 9.house 9LNK Maria  
 ‘John’s mother<sub>i</sub> is Maria’s mother<sub>i</sub>’

(220) *Identification*

- a. Uriya mu-ntu **ni(=we)** Johne  
 1DEM\_DIST 1-person be<sub>NI</sub>=1FOC John  
 ‘That man is John’ (pointing at a person in the distance)
- b. aba **ni(=bo)** a-ba-ana ba-angi  
 2DEM\_PROX be<sub>NI</sub>=2FOC AUG-2-child 2AGR-POSS1SG  
 ‘These are my children’ (pointing at a group of people in a photo) Kinyamulenge

The same general observations hold of the copular system in neighboring Kirundi. In present tense predicational clauses, both *-ri* and *ni* exhibit the same distributional restrictions as in Kinyarwanda and Kinyamulenge.

(221) *Predication*

- a. Jean **n’** u-mu-ganga  
 Jean be<sub>NI</sub> AUG-1-doctor  
 ‘Jean is a doctor’
- b. Kigali **i-ri** mu Rwanda  
 9.Kigali 9SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’ Kirundi

Similarly, specificational, identificational, and equational clauses in Kirundi similarly involve an invariant copula and optional focus marker. Interestingly however, the presence of the focus marker appears slightly less acceptable in Kirundi than in Kinyamulenge, as evidenced by the fact that it is unavailable in (224).<sup>3</sup>

(222) *Specification*

<sup>3</sup>It is unclear why the focus marker is less acceptable in identificational clauses (224) than other clause types (224)-(223). Further work is needed to determine if this is a generalizable difference between identificational clauses and other clause types in Kirundi. For the purposes of this section, it suffices to say that this is simply a minor point of variation between Kirundi and Kinyamulenge.

- a. O-mu-kuru w' igihugu w' uRwanda **ni=we** Kagame  
 AUG-1-elder 1LNK 7-country 1LNK Rwanda be<sub>NI</sub>=1FOC Kagame  
 'The president of Rwanda is (Paul) Kagame'
- b. Jean na Marie **ni=bo** a-ba-ganga  
 Jean and Marie be<sub>NI</sub>=2FOC AUG-2-doctor  
 'John and Maria are the doctors' (answers the question "who are John and Mary?")

(223) *Equation*

- a. Mama wa Jean **ni=we** mama wa Marie  
 1.mother 1LNK Jean be<sub>NI</sub>=1FOC 1.mother 1LNK Marie  
 'Jean's mother<sub>i</sub> is Marie's mother<sub>i</sub>'
- b. A-ba-byeyi ba Marie **ni=bo** (a-ba-byeyi) ba Jean  
 AUG-2-parent 2LNK Marie be<sub>NI</sub>=2FOC AUG-2-parent 2LNK Jean  
 'Marie's parents<sub>i</sub> are Jean's parents<sub>i</sub>'

(224) *Identification*

- a. Uriya mu-gabo **ni(\*=we)** Jean  
 1DEM\_DIST 1-man be<sub>NI</sub>=1FOC Jean  
 'That man is John' (pointing at a person in the distance) Kirundi

### 5.2.3 Tense and Aspect

Across the languages surveyed in this project, there is a nearly identical effect of tense and aspect. In the present tense, languages exhibit the most substantial copular variation. This is the only context where multiple copular forms are in contrast, and therefore the only context where interpretive contrast are attested. In other tense/aspect environments, only a single form of the copula is attested. In general, the copula is obligatorily realized as *-li* (or its cognate *-ri*) in the past tense, and as *-ba* in the future tense.

#### 5.2.3.1 JD50 languages

In both Kihavu and Mashi, there are three primary strategies for expressing non-verbal predication in the present tense. In pure predicational clauses, one of two verbal copulas (e.g., *-li/-ba*) is used depending on the interpretation of the predication relation. In principle, either copula is available

with any predicate and subject (though see chapter 4 for more on their interpretive restrictions). This can be seen in the representative examples from Mashi below (225).<sup>4</sup>

(225) *Predication - present tense*

- a. N-**∅**-**di**                      mu-igiriza  
    1SG.SM-PRS-be-<sub>LI</sub> 1-teacher  
    ‘I am a teacher’ (in this context)
- b. M-**∅**-**pa**                      mu-igiriza  
    1SG.SM-PRS-be-<sub>BA</sub> 1-teacher  
    ‘I am a teacher’ (speaker’s career)
- c. Johne a-**∅**-**li**                      mu-igiriza  
    John 1SM-PRS-be-<sub>LI</sub> 1-teacher  
    ‘John is a teacher’ (in this context)
- d. Johne a-**∅**-**ba**                      mu-igiriza  
    John 1SM-PRS-be-<sub>BA</sub> 1-teacher  
    ‘John is a teacher’ (his career)

Mashi

On the other hand, present tense specificational, equational, and identificational clauses are expressed via a focus marker, as shown in Kihavu (211)

(211) *Specification - present tense*

- a. Gi-shambo **ye**      omw-ana  
    7-thief      1FOC 1-child  
    ‘The thief is the child’
- b. O-mu-rhani      **kyo**      gi-shambo  
    AUG-1-murderer 7FOC 7-thief  
    ‘The murderer is the thief’

Kihavu

In other tense-aspect environments however, the realization of the copula is slightly different. Due to data limitations, we will focus on the effect of tense and aspect on the form of the copulas specifically in pure predicational clauses, with a particular focus on Mashi. Note however that unless overtly stated otherwise, the Mashi and Kihavu systems are nearly identical.

---

<sup>4</sup>The surface form of *-li* and *-ba* is sensitive to a number of phonological factors as well. This is particularly salient in examples like (225a) and (225b), where a preceding nasal triggers place assimilation, fortition, and/or devoicing of the root initial consonant.

As a preface, it is unclear whether copular forms in Kihavu and Mashi bear overt tense morphology, and — if they do — what that tense morphology is. In their typology of tense-aspect morphology in Bantu languages, Nurse and Philippson (2006) report that Mashi, among other Eastern Bantu languages, typically marks the past tense through the joint use of two distinct morphemes: a prefixal past tense marker *-aá*, and a reanalyzed suffixal aspectual (perfective) marker *-ile*.

- (226) *Áá-bá-lumé bá-á-shákw-iire dúbá*  
 AUG-2-men 2SM-DPST-pound-PFV fast  
 ‘The men pounded fast’ (remote/distant past)

Mashi (Adapted from Nurse and Philippson 2006:160)

In past tense copular clauses however, it is less clear that this is the strategy Mashi employs. Consider for example the past tense sentences in (227), which involve a possessive *have*-verb *-dwiire* and a copular *be*-verb *-li* respectively.

- (227) a. *Johne a-li-dwiire o-mu-lali oku a-ja e Kinshasa*  
 John 1SM-PST-have-DWIIIRE AUG-3-plan COMP 1SM-go LOC Kinshasa  
 ‘John had a plan to go to Kinshasa’
- b. *Maria a-li-ri mu-lwala*  
 Maria 1SM-PST-be-LI 1AGR-sick  
 ‘Maria was sick’

Mashi

The issue in these examples is that the possessive in (227a) doesn’t straightforwardly adhere to the joint past marking strategy shown in (226); there is no dedicated tense-aspect prefix (e.g., *-a/-á*), and although it appears that the perfective marker *-ile* may be present (e.g., in *-dwiire*), this material is part of the lexical verb *-dwiire* which is invariant across tense/aspect environments (Finholt 2024). As such, the past tense interpretation in (227a) appears to correlate with the pre-root morpheme *-li*, which is identical in shape to the copular verb *-li*. Since this morpheme and the root *-dwiire* are directly adjacent with no intervening agreement or tense-aspect-mood morphology between them, I assume that this is not a periphrastic construction that includes an auxiliary *be*-verb. If correct, the most straightforward analysis of the morpheme *-li* is that it marks past tense.

Turning now to the copular example (227b), a similar issue arises; there is no overt past tense prefix *-a/-á*, and though there may be a correlate of the perfective suffix *-ile* present, it takes a

different form than seen previously (e.g., *-ri*). If however it is the case that *-li* can be a past tense marker, then it is possible that the *-ri* segment in (227b) is actually an allomorph of the copular verb *-li* that is preceded by the tense marker, *-li*.<sup>5</sup>

For the time being, I will not commit myself to one of these analyses, e.g., *-ri* as an allomorph of the copular verb *-li* or an allomorph of the perfective marker *-ire*. However, for the purposes of this section, I will gloss past tense copular constructions as shown in (227b). Ultimately the analytical difference between the two possible analyses are minimal, as they merely differ in whether tense or aspect is the relevant conditioning factor. In both cases, the pattern is the same: only an allomorph of *-li* may appear in the presence of past tense/perfective morphology.

Setting aside this ambiguity, there are two “past tense” environments in Mashi (and Kihavu) that have a salient effect on the form of the copula in pure predicational clauses. One of these involves the past tense/perfective morphology shown in (227b). Unlike the present tense, only one form of the copula is available in this environment: *-li* (or rather *-ri*).<sup>6</sup>

(228) Johne a-li-<sup>✓</sup>ri/#ba ci-shambo  
 John 1SM-PST-be-<sub>LI</sub>/be-<sub>BA</sub> 7-thief  
 ‘John was a thief’ Mashi

The second “past” environment involves the imperfective suffix *-ga*, which is transparently related to the cross-Bantu imperfective marker *-a(n)ga* (Nurse and Philippson 2006). Once again, we find that only one copular form may appear in this environment, however in this case it is the copular verb *-ba*. Note that when asked, our Kihavu and Mashi consultants stated that there is no interpretive difference between (228) and (229), and that these “past tense” strategies are completely interchangeable.

<sup>5</sup>The existence of such an allomorph could be attributed to a process of phonological dissimilation, as argued in Finholt (2024).

<sup>6</sup>Note that even under the alternative analysis, *-ba* would still be unavailable:

(1) Johne a-<sup>✓</sup>li/#ba-ri ci-shambo  
 John 1SM-be-<sub>LI</sub>/be-<sub>BA</sub>-PFV 7-thief  
 ‘John was a thief’

- (229) Johne a-#li/√ba-ga ci-shambo  
 John 1SM-be-<sub>LI</sub>/be-<sub>BA</sub>-IMPFV 7-thief  
 ‘John was a thief’ Mashi

The future tense in Kihavu and Mashi imposes a similar restriction on the form of the copula, as only *-ba* may appear. This is clearest in Mashi, where *-li* is incompatible with the future tense marker *-ishi* (230). In Kihavu, it is not clear that there is overt future tense marking in copular clauses, however *-ba* in isolation can be used to yield a future-oriented reading (231). Subjunctive morphology may also be used to express the future tense in Kihavu in some cases (see subsection 5.2.4).

- (230) Johne a-ishi-#li/√ba ci-shambo  
 John 1SM-FUT-be-<sub>LI</sub>/be-<sub>BA</sub> 7-thief  
 ‘John will be a thief’ Mashi

- (231) Johne a-ba mu-lwala  
 John 1SM-be-<sub>BA</sub> 1AGR-sick  
 ‘John is/will be a thief’ Kihavu

As noted previously, we do not have enough data to make any generalizations about the effect of tense/aspect in non-predicational clauses. However, initial data from Kihavu seems to suggest a similar neutralization pattern to what is seen in predicational clauses. Consider for example the following equational clauses. While present tense equational clauses are expressed via a focus marker in isolation, e.g., *ye*, past tense equational clauses involve the combination of the focus marker and a past inflected form of the copula *-li*, and future equational clauses involve the combination of the focus marker and an inflected (subjunctive marked) form of the copula *-ba*.

(232) *Equation - tense contrasts*

- a. Nyina wa Mugisha **ye** nyina wa Murhulla  
 1.mother 1SM.LNK Mugisha 1FOC 1.mother 1SM.LNK Murhulla  
 ‘Mugisha’s mother<sub>i</sub> is Murhulla’s mother<sub>i</sub>’
- b. Nyina wa Mugisha **ye a-li-ri** nyina wa  
 1.mother 1SM.LNK Mugisha 1FOC 1SM-PST-be-<sub>LI</sub> 1.mother 1SM.LNK  
 Murhulla  
 Murhulla  
 ‘Mugisha’s mother<sub>i</sub> was Murhulla’s mother<sub>i</sub>’



- c. Nyina wa Mugisha **ye** a-**be** nyina wa  
 1.mother 1SM.LNK Mugisha 1FOC 1SM-be-<sub>BA</sub>.SBJV 1.mother 1SM.LNK  
 Murhulla  
 Murhulla  
 ‘Mugisha’s mother<sub>i</sub> will be Murhulla’s mother<sub>i</sub>’

### 5.2.3.2 JD60 languages

As we have already seen, there are two primary copular forms that appear in present tense sentences in Kinyamulenge and Kirundi: verbal *-ri*, and invariant *ni*. In some copular clause types, the latter of these may co-occur with a focus marker derived from the pronominal stem *-o*. In other tense environments however the availability of these copular forms differs, and in some cases a third copular form (e.g., *-ba*) is found.

In the present tense, three copular forms are attested. The form of the copula in this environment depends on the person features of the subject, the category of the predicate, and the type of the copular clause (see subsection 5.2.2). Focusing specifically on pure predicational clauses, we see a binary distinction between the verbal copula *-ri* and the invariant copula *ni*. As illustrated below in Kinyamulenge (which exhibits an identical pattern to Kirundi), *-ri* is used in all predications involving a first and second person subject, and locative predications involving a third person subject. In contrast, *ni* is used in all non-locative predications involving a third person subject (however see subsection 5.2.5 for an example where *ni* is used with a locational predicate).

#### (233) Predication - present tense

- a. Njewe n-**Ø-di** u-mu-ganga  
 1SG.PRO 1SG.SM-PRS-be-<sub>RI</sub> AUG-1-doctor  
 ‘I am a doctor’
- b. Wewe u-**Ø-ri** u-mu-ganga  
 2SG.PRO 2SG.SM-PRS-be-<sub>RI</sub> AUG-1-doctor  
 ‘You are a doctor’
- c. Johne **ni** u-mu-ganga  
 John be-<sub>NI</sub> AUG-1-doctor  
 ‘John is a doctor’

- d. Johne a- $\emptyset$ -ri mu Rwanda  
 John 1SM-PRS-be-RI 18LOC Rwanda  
 ‘John is in Rwanda’

Kinyamulenge

In the past tense however, this contrast is fully neutralized, as only one copular form is attested. In the presence of overt past tense morphology (e.g., the recent past marker *-a*/distant past marker *-á*), the copula is obligatorily realized as *-ri* regardless of subject person, predicate category, or interpretive context. The distinct past tense pattern can be seen most clearly in contexts where invariant *ni* is used in the present, e.g., with non-third person subjects and non-locative predicates).<sup>7</sup> Once again, this can be seen in Kinyamulenge below, though the same pattern holds of Kirundi as well.

(234) *Predication - past tense*

- a. Njewe n-á-ri u-mu-ganga  
 1SG.PRO 1SG.SM-DPST-be-RI AUG-1-doctor  
 ‘I was a doctor’
- b. Wewe u-á-ri u-mu-ganga  
 2SG.PRO 2SG.SM-DPST-be-RI AUG-1-doctor  
 ‘You were a doctor’
- c. Johne y-á-ri u-mu-ganga  
 John 1SM-DPST-be-RI AUG-1-doctor  
 ‘John was a doctor’
- d. Johne y-á-ri mu Rwanda  
 John 1SM-DPST-be-RI 18LOC Rwanda  
 ‘John was in Rwanda’

Kinyamulenge

A similar neutralization pattern is observed in the future tense in both languages (marked by the future tense marker *-zaa/-zoo*), albeit with a different copular form. In the future tense, the copula is obligatorily realized as the verbal copula *-ba*, as illustrated for all persons in Kinyamulenge (235) and for third person in Kirundi (236).

<sup>7</sup>In the past tense, the class 1 subject marker *a-* is realized as a glide (represented orthographically as *y-*). This is likely the result of a general constraint on adjacent vowels in the language. A process of glide formation applies in (234c-d) due to the adjacency of the class 1 subject marker *a-* and the distant past tense marker *-á*. It is possible that a similar process occurs in (234b) as well; despite being presented as *u-*, the second person singular subject marker may be pronounced as [w].

(235) *Predication - future tense*

- a. Njewe n-zaa-**ba** u-mu-aarimu  
1SG.PRO 1SG.SM-FUT-be-<sub>BA</sub> AUG-1-teacher  
'I will be a teacher'
- b. Njewe n-zaa-**ba** u-mu-aarimu  
2SG.PRO 2SG.SM-FUT-be-<sub>BA</sub> AUG-1-teacher  
'You will be a teacher'
- c. Njewe n-zaa-**ba** u-mu-aarimu  
John 1SM-FUT-be-<sub>BA</sub> AUG-1-teacher  
'John will be a teacher'
- d. Johne a-zaa-**ba** mu Rwanda  
John 1SM-FUT-be-<sub>BA</sub> 18LOC Rwanda  
'John will be in Rwanda'

Kinyamulenge

(236) *Predication - future tense*

- a. Njewe n-zaa-**ba** u-mu-aarimu  
John 1SM-FUT-be-<sub>BA</sub> AUG-1-teacher  
'John will be a teacher'
- b. Johne a-zaa-**ba** mu Rwanda  
John 1SM-FUT-be-<sub>BA</sub> 18LOC Rwanda  
'John will be in Rwanda'

Kirundi

This in many ways correlates with the tense-aspect conditioned neutralization pattern observed in the JD50 languages Kihavu and Mashi above. Although the JD50 languages exhibit a slightly more complicated pattern due to the existence of two “past tense” marking strategies, the general pattern is the same across the JD50 and JD60 languages discussed here. In all cases, there is an overt contrast between copular forms in the present tense, and a full-neutralization of this contrast in other tense-aspect environments. What’s more, there is a significant overlap among the languages in terms of which forms appear in different environments. With the exception of the imperfective strategy, past tense correlates with reflexes of the proto-Bantu copula proto-Bantu *\*-de*, e.g., *-li/-ri* in both groups. Similarly, the future tense uniformly correlates with reflexes of proto-Bantu *\*-bà*, e.g., *-ba*. The tense-aspect patterns are summarized in Table 5.2 below.

	Past		Present	Future
	Perfective	Imperfective		
JD50 Kihavu and Mashi	<i>-li</i>	<i>-ba</i>	<i>-li, -ba</i>	<i>-ba</i>
JD60 Kinyamulenge and Kirundi		<i>-ri</i>	<i>-ri, ni</i>	<i>-ba</i>

Table 5.2: Copular forms in different tense/aspect environments

As with the JD50 languages in subsection 5.2.3.1, there is insufficient data from non-predicational copular clauses to make any concrete generalizations about the effect of tense-aspect on the form of the copula. However, initial data in Kinyamulenge suggests that the neutralization pattern seen in predicational copular clauses also holds of other clause types.

In the present tense, specificational, equational, and identificational clauses are expressed via the invariant copula *ni* and an optional post-copular focus marker in Kinyamulenge (219).

(219) *Equation - present tense*

- a. Mama wa Johne **ni(=we)** mama wa Maria  
 1.mother 1LNK John be<sub>NI</sub>=1FOC 1.mother 1LNK Maria  
 ‘John’s mother<sub>i</sub> is Maria’s mother<sub>i</sub>’.
- b. Inzu ya Johne **ni(=yo)** inzu ya Maria  
 9.house 9LNK John be<sub>NI</sub>=9FOC 9.house 9LNK Maria  
 ‘John’s mother<sub>i</sub> is Maria’s mother<sub>i</sub>’

In the presence of overt tense morphology however, the invariant copula is no longer available. Instead a verbal copula, e.g., *-ri/-ba*, surfaces together with the optional focus marker. As before, the particular verbal copula that appears directly correlates to tense; *-ri* obligatorily appears in the past tense, and *-ba* in the future.

(237) *Equation - past and future tense*

- a. U-mu-byeyi wa Peter y-á-**ri(=we)** mu-byeyi wa Maria  
 AUG-1-father 1LNK Peter 1SM-DPST-be<sub>RI</sub>=1FOC 1-father 1LNK Maria  
 ‘Peter’s father<sub>i</sub> was Maria’s father<sub>i</sub>’ (when he was alive)
- b. U-mu-byeyi wa Peter a-zaa-**ba(=we)** mu-byeyi wa Maria  
 AUG-1-father 1LNK Peter 1SM-FUT-be<sub>BA</sub>=1FOC 1-father 1LNK Maria  
 ‘Peter’s father<sub>i</sub> will be Maria’s father<sub>i</sub>’ (when she is born/adopted)

Though there is more work to be done regarding the effect of tense-aspect in non-predicational copular clauses, the same general copular neutralization pattern appears to apply, at least in Kinyamulenge.

## 5.2.4 Mood

In addition to tense and aspect, mood is another factor that plays a salient role in determining the form of the copula in Eastern Bantu languages. Across the languages surveyed in this work, there is a uniform distinction between realis and irrealis mood in the copular system: while a variety of copular contrasts and forms are attested in realis contexts (e.g., present/past indicative), only the copular verb *-ba* may appear in certain irrealis contexts (e.g., subjunctive/infinitival mood).

As illustrated in previous sections, languages of the JD50 and JD60 subgroups exhibit substantial variation in the form of the copula in the indicative mood. This is particularly apparent in the present indicative, where languages employ two or even three distinct forms (that not including the focus marker/pronominal copula). In contrast, there is no variation in the form of the copula in irrealis environments like the subjunctive and infinitival moods, where only *-ba* is available.

In many Bantu languages, the subjunctive mood is expressed via a suffixal final vowel *-e* (Nurse and Philippson 2006). Subjunctive morphology (e.g., final vowel *-e*) can be seen in the JD50 and JD60 languages under consideration in cases of uncertainty, e.g., (238a) and (238d), as well as cases of obligation/requirement (239b-c).

### (238) *Subjunctive mood*

- a. Johne ye-**\*l(e)ʷbe** mu-aalimu  
 John 1SG.SM-be<sub>-LI</sub>.SBJV/be<sub>-BA</sub>.SBJV 1-teacher  
 ‘John will/could/may be tall’ Kihavu
- b. O-shinganine o-**\*l(e)ʷbe** mu-liri o-twa  
 2SG.SM-must.SBJV 2SG.SM-be<sub>-LI</sub>.SBJV/be<sub>-BA</sub>.SBJV 1AGR-tall 2SG.SM-pick  
 a-ma-lehe oku-gu-murhi  
 AUG-6-fruit 15LOC-3DEM-3-tree  
 ‘You must be tall to pick fruit from this tree’ Mashi

- c. U-gomba      u-**\*r(e)ʼbe**                      muremure  
 2SG.SM-must 2SG.SM-be-<sub>RI</sub>.SBJV/be-<sub>BA</sub>.SBJV 1AGR.tall  
 ‘You must be tall’ (to do this task)                      Kinyamulenge
- d. U-zoo-**\*r(e)ʼbe**                      mu-kuru  
 2SG.SM-FUT-be-<sub>RI</sub>.SBJV/be-<sub>BA</sub>.SBJV 1AGR-tall  
 ‘You will/could/may be tall’                      Kirundi

In the infinitive (marked by the prefix *-ku*, we see an identical pattern. As illustrated below in Mashi and Kinyamulenge, only the copular form *-ba* may appear together with infinitival morphology. This pattern can be seen more broadly in the form of the infinitive ‘to be’ across Bantu languages, which generally takes the form *kuba* (or something close to it).

(239) *Infinitival mood*

- a. Bi-li      bi-nja      o-ku-**ba**      mu-liri  
 8SM-be-<sub>LI</sub> 8AGR-good AUG-INF-be-<sub>BA</sub> 1AGR-tall  
 ‘It is good to be tall’                      Mashi
- b. U-saba      ku-**ba**      muremure  
 2SG.SM-request INF-be-<sub>BA</sub> 1AGR.tall  
 ‘You have to be tall’ (to do this task)                      Kinyamulenge

The correlation between irrealis environments and *-ba* becomes even clearer if one also considers the future tense to be irrealis. Assuming that irrealis mood(s) is used to describe events or situations that have not and/or may not come to be in the actual world, then one could argue that the future tense has an irrealis flavor in the sense that it is used to describe events or situations that have not yet occurred (see for example Mithun 2001). As discussed in subsection 5.2.3, future tense morphology strictly co-occurs with the copular form *-ba* in Kihavu, Mashi, Kinyamulenge, and Kirundi, just as we see in the subjunctive and infinitive. Further research is necessary to determine whether this pattern extends to all irrealis contexts, however the fact that *-ba* is the only available form in the subjunctive, infinitive, and future tense provides strong evidence of a correlation between irrealis mood and cognates of *-ba* in Eastern Bantu languages.

### 5.2.5 Other

In addition to those discussed above, there are logically many more factors that could affect the expression of non-verbal predication in the languages under discussion. Indeed, it has been shown that the form of the copula in some Bantu languages is sensitive to negation (Gibson et al. 2019), clausal embedding and/or choice of complementizer (Gluckman 2024), and whether or not the predication relation involves a possessive configuration (Finholt 2024). Though I will leave any discussion of these additional factors to future work, there are two factors mentioned briefly in previous sections — namely the person features of the subject and predicate category — that are worth discussing here, as they play a critical role in determining the form of the copula in many JD60 languages, including Kinyamulenge and Kirundi.

Previous work on Bantu copular systems has shown a regular sensitivity to both subject person and predicate category (or something similar) in languages of the JD group, most notably in Kinande (JD42; Schneider-Zioga and Mutaka 2015a) Kinyarwanda (JD61; Jerro 2015), and Kirundi (JD62; Lafkioui et al. 2016, Gatchalian 2023). In each of these languages, there are two distinct copular forms that appear in present tense (pure) predicational clauses, namely invariant *ni* and verbal *-ri*. In this context, the distinction between the two forms is primarily driven by the person features of the subject, with *-ri* being used in all sentences involving a first or second person subject, and *ni* appearing strictly in predications involving a third person subject. That said, *-ri* may also appear in third person predications involving a locative predicate, meaning there is a categorial distinction between nominal/adjectival predicates (which require invariant *ni*) and locative predicates (which require verbal *-ri*) that is unique to the third person.<sup>8</sup>

In previous sections I provided further evidence of a sensitivity to subject person and predicate category in the copular system of Kirundi, and demonstrated that an identical pattern generally

---

<sup>8</sup>As an argument against a selectional approach, Jerro (2015) demonstrates that Kinyarwanda *-ri* also appears with locative adverbials in the third person, suggesting it does not subcategorize for locative (PP) predicates. Instead Jerro provides a semantic analysis, suggesting that *-ri* has an inherently locative semantics. Ultimately I will not directly address whether the co-occurrence between *-ri* and locative predicates in Kinyarwanda (or other JD60 languages) is a selectional issue or not, however I do suggest that the reason for the correlation between *-ri* and locatives may have something to do with its semantics.

holds of related Kinyamulenge (JD61a) as well, as illustrated below (see also Table 4.3).

(233) *Predication - present tense*

- a. Njewe n- $\emptyset$ -**di** u-mu-ganga  
 1SG.PRO 1SG.SM-PRS-be-<sub>RI</sub> AUG-1-doctor  
 ‘I am a doctor’
- b. Wewe u- $\emptyset$ -**ri** u-mu-ganga  
 2SG.PRO 2SG.SM-PRS-be-<sub>RI</sub> AUG-1-doctor  
 ‘You are a doctor’
- c. Johne **ni** u-mu-ganga  
 John be-<sub>NI</sub> AUG-1-doctor  
 ‘John is a doctor’
- d. Johne a- $\emptyset$ -**ri** mu Rwanda  
 John 1SM-PRS-be-<sub>RI</sub> 18LOC Rwanda  
 ‘John is in Rwanda’

Kinyamulenge

That being said, the effect of predicate category in Kinyamulenge is slightly more complicated than what is reported in other closely-related languages. As in Kinyarwanda and Kirundi, locational predicates generally co-occur with the verbal copula *-ri* even in sentences including a third person subject where *ni* is typically used (233c-d). However, as discussed in subsection 4.4.5, locational predicates may also occur with the invariant copula *ni* in some cases, resulting in an interpretive contrast between *-ri* and *ni*.

(175) *Context: You have never heard of Kigali or Rwanda before, but you see an official-looking map of East Africa and there is a city named ‘Kigali’ marked inside the borders of ‘Rwanda’.*

- Kigali #**ni**/<sup>✓</sup>**i-ri** mu Rwanda  
 Kigali be-<sub>NI</sub>/<sub>9</sub>SM-be-<sub>RI</sub> 18LOC Rwanda  
 ‘Kigali is in Rwanda’

Kinyamulenge

(176) *Context: You just got back from visiting Kigali, and you are listing all of the things you know about the city.*

- Kigali <sup>✓</sup>**ni**/**#i-ri** mu Rwanda  
 Kigali be-<sub>NI</sub>/<sub>9</sub>SM-be-<sub>RI</sub> 18LOC Rwanda



This variation contrasts with what is generally reported in Kirundi (Lafkioui et al. 2016, Gatchalian 2023) and closely-related Kinyarwanda (Jerro 2015), where locational predicates obligatorily occur with *-ri*. Given the close genealogical relationship between these languages and the tremendous degree of overlap in their copular systems, the Kinyamulenge data suggests that the effect of predicate category on the form of the copula in these languages may be slightly more nuanced than originally thought, and may even relate to the semantics of the copula.

### 5.3 Analysis

In the previous sections of this chapter, I provided a descriptive overview of copular variation in the Great Lakes Bantu languages Kihavu (JD52), Mashi (JD53), Kinyamulenge (JD61a), Kirundi (JD62), and Kifuliiru (JD63). I showed that the form of the copula in these languages is sensitive to a number of different factors, including interpretive context, copular clause type, tense-aspect-mood environment, and to a lesser extent, predicate category and subject person features. Although language-specific differences in the exact combination of copular forms and relevant conditioning factors are attested, there are significant similarities across languages in terms of how particular factors affect the form of the copula. For example, languages exhibit the same tense/mood-conditioned neutralization pattern regardless of which copular forms they exhibit; though languages often permit multiple interpretively distinct copulas in the present tense, the copula is almost always realized as *-li/-ri* in the past tense, and as *-ba* in the future tense/irrealis mood. Similarly, although languages exhibit substantial variation in the present tense, cognate forms across languages tend to encode similar meanings; *-li/-ri* generally encodes a limited reading of property ascriptions, while *-ba* encodes a more general, non-limited reading. The effect of copular clause type is similarly consistent across languages, albeit not in the same way. In all of the languages surveyed, there is a clear distinction between pure predicational clauses on the one hand and specificational, identificational, and equational clauses on the other; while predicational clauses are expressed by a

copular form in isolation, all other clauses involve the combination of a copular form and a focus marker.

From a descriptive point of view, these similarities are quite salient; conditioning factors tend to have the same effect across languages, and cognate forms tend to exhibit similar distributions. At the same time however, individual languages sometimes differ quite radically from each other in how the form of the copula is affected by its environment. For example, Kifuliiru is the only language surveyed that exhibits a three-way interpretive contrast in the present tense; there is no direct equivalent to the form *-tula* in other languages. Likewise, Kinyamulenge and Kirundi exhibit an invariant copula (e.g., *ni*) in present tense copular clauses featuring a third person subject — something that is not observed in the other languages surveyed. An adequate analysis of copular variation in these languages must therefore capture the broadly generalizable patterns that hold across languages while also capturing the unique differences between them.

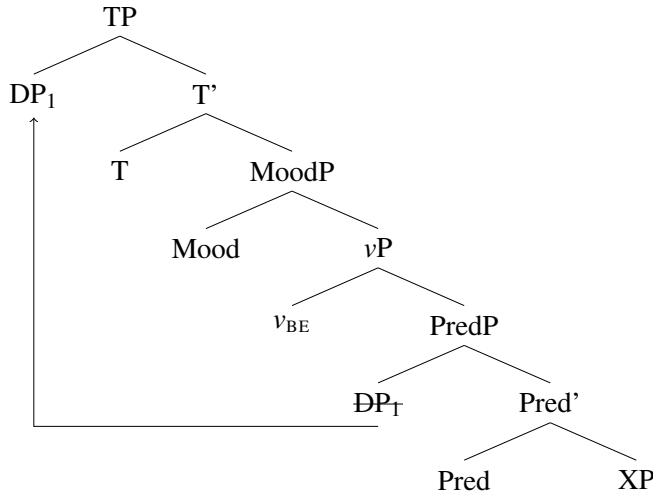
In this section, I will provide an analysis of copular allomorphy in the Bantu languages discussed in this work that accounts for the overlapping distributions of related copular forms while also capturing language-specific forms and effects. Specifically, I will propose an analysis that treats the observed copular variants as context-specific realizations of the light verbal head  $v_{BE}$  (Heycock and Kroch 1999, Adger 2003, Baker 2003, Balusu 2014, Myler 2016, 2018). To account for the similar patterns of allomorphy, I provide a generalizable set of VI rules for  $v_{BE}$  that ties the distribution of various copular forms to the presence and/or features of particular heads in the syntax, e.g., Mood, Tense, Pred, and the focus head *f*. Differences across languages are accounted for by the addition of language-specific VI rules to the base set of rules.

### **5.3.1 Background assumptions and schematics**

Following a long tradition of work on the syntax of copular clauses, I assume a fairly standard view of non-verbal predication (at least in pure predicational clauses) in which the copula realizes a meaningless verbal head  $v_{BE}$  (or just  $v$ ) that takes as its complement a small clause headed by the predicational head Pred (Heycock and Kroch 1999, Adger 2003, Baker 2003, Balusu 2014, and

Myler 2016, 2018, a.o.). I take all predicational clauses in the data involving a verbal copula to correspond to the structure in (240).

(240) *Predicational clause structure*

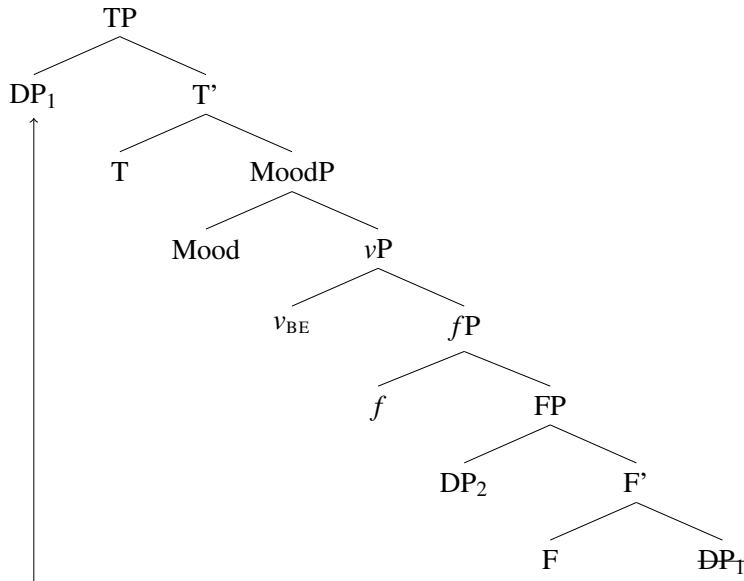


Under this approach, interpretive contrasts among copulas like that between Spanish *estar/ser* are attributed to distinct to the existence of distinct Pred heads, e.g., Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub> (Adger and Ramchand 2003, Markman 2008a, Balusu 2014). In this sense, the meaning of the copula is situated on Pred; the particular Pred variant that is merged determines the form (and meaning) of the copula. I will assume that the interpretive contrasts observed between verbal copular forms in Eastern Bantu languages are likewise derived from the existence of distinct Pred heads. For the sake of convention, I will largely describe this as a distinction between the heads Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub>. Importantly however I do not assume that the semantic distinction between these heads is equivalent to a traditional stage-individual distinction. Instead, I will use these heads as placeholders for the semantic analysis I proposed in chapter 4, with Pred<sub>STAGE</sub> encoding a bounded or limited property meaning, and Pred<sub>INDIV</sub> encoding a more general property meaning.

I assume a slightly different structure for specificational, equational, and identificational clauses. Following Schneider-Zioga’s (2021) analysis of specificational copular clauses in another Eastern Bantu language, e.g., Kinande (JD42), I assume that specificational, equational, and identificational clauses (which Schneider-Zioga collectively refers to as specificational copular clauses, or “SCCs”) involve an additional layer of structure above PredP, namely a focus phrase *fP*. In these clause types,

$v_{BE}$  takes an  $fP$  complement that contains a small clause FP headed by a relational head F that encodes a specificational semantics (Romero 2005, Heycock 2012, Schneider-Zioga 2021). The assumed structure for SCCs can be seen below (241).<sup>9</sup>

(241) *SCC structure*



(adapted from Schneider-Zioga 2021)

The primary motivation for Schneider-Zioga’s analysis in Kinande is that SCCs involve an agreeing focus marker that is distinct from the copula, just like the JD50 and JD60 languages discussed in this chapter. In the present tense, this focus marker appears in isolation (63)-(65), however it follows an inflected verbal copula in other tense environments (242).

- (242) ebyalya ebyo nabya nanzire kutsibu **byabya** i-lw’ olukondi  
 8.food 8.that 1SM.PST.be 1SG.SM.like best 8SM.PST.be 11FOC 11.beans  
 ‘The food that I liked the best was beans’ Kinande (adapted from Schneider-Zioga 2021)

With these structural assumptions in mind, I take the various verbal copular forms, e.g., *-li*, *-ba*,

<sup>9</sup>The structure I adopt for SCCs in (241) differs from the original proposal in Schneider-Zioga (2021) — see (121) — in that it does not assume that DP<sub>1</sub> moves through Spec,  $fP$  nor include any information about the agreement relation between  $f$  and DP<sub>2</sub>. These aspects of the original analysis were omitted due to data limitations. More work is necessary to determine whether the agreement patterns in the languages in this work are identical to that of Kinande and support a more faithful analysis. Note that this analysis also raises the question of why DP<sub>1</sub> can invert across DP<sub>2</sub>, as this would seemingly violate the Minimal Link Condition (Chomsky 1995). In the absence of the agreement data that would motivate an intermediate movement step for DP<sub>1</sub> in Spec,  $fP$  as in Schneider-Zioga (2021), I will ignore this issue here. The simplified structure I propose for SCCs serves simply to model the basic word order and allomorphy facts, and does not provide a concrete account of movement.

*-tula* (but not necessarily invariant *ni*, see subsection 5.3.4), to be allomorphs of  $v_{BE}$  that surface in particular environments. Specifically, I assume that  $v_{BE}$  allomorphy is sensitive to the presence or absence of the following heads: Mood<sub>[+indic]</sub>, T<sub>[+prs]</sub>, Pred<sub>STAGE</sub>/Pred<sub>INDIV</sub>, and *f*.

To account for the conditioning effect of these different heads, I partially adopt a spanning insertion approach to  $v_{BE}$  allomorphy à la Merchant (2015). In particular, I follow the analysis of *be* and *have* allomorphy in Mashi presented in Finholt (2024) in which copular verbs jointly realize the heads  $v_{BE}$  and Pred. The motivation for this analysis in Finholt (2024) stems from a complex pattern of  $v_{BE}$  allomorphy that is simultaneously sensitive to multiple higher and lower heads, e.g., Mood and Tense above, and Pred and P<sub>HAVE</sub> below. The idea in this analysis is that multiple heads can be jointly targeted for vocabulary insertion and/or simultaneously serve to condition allomorphy of an adjacent series (span) of heads. In this case, the argument is that there are two relevant spans at play in the realization of *be*/*have*-verbs in Mashi: a lower span [ $v_{BE}$ +Pred(+P<sub>HAVE</sub>)], and a higher span [Mood-Tense]. The realization of the lower span in particular is argued to be sensitive to both the heads involved in that span, e.g.,  $v_{BE}$ , Pred, and (optionally) P<sub>HAVE</sub>, and also the heads in the higher span, e.g., Mood and Tense (which are together discretely exponed relative to the lower span). Though a spans analysis is not as integral to the copular allomorphy data presented in this work, it is particularly adept at accounting for the various interpretive contrasts that appear in the present tense in our data. Assuming that the semantics of the copula is attributed to the existence of distinct Pred heads, copular contrasts can be explained as cases where  $v_{BE}$  is jointly realized with a particular Pred head, e.g, Pred<sub>STAGE</sub> or Pred<sub>INDIV</sub>.

Though  $v_{BE}$  allomorphy differs slightly across the various languages investigated in this work, the general analysis follows a similar formula. For each language (or group of languages), I propose a set of ordered Vocabulary Insertion rules for  $v_{BE}$  that determine the form of the copula according to the surface environment of  $v_{BE}$ . These rules are ordered according to the specificity of the contexts in which they apply, with the most specific rules applying first. As such, I assume that the different allomorphs of  $v_{BE}$  are in competition, with some applying in highly specific contexts, and others applying in more general contexts. Importantly, each set of ordered VI rules I propose in the

following sections consists of the same core contrasts. In most cases, there are two highly specific VI rules that apply uniquely in the present tense: one that realizes a non-limited or maximal copula (e.g., *-ba* or *-tula*) in the presence of  $\text{Pred}_{\text{INDIV}}$  (243a), and one that realizes the zero copula  $\emptyset$  in the presence of the focus head *f*, e.g., in specificational, equational, and identificational clauses (243b). These are the copular forms that are the most restricted in their distribution. The next most specific VI rule is that which realizes reflexes of the proto-Bantu copula *\*-de*, e.g., *-li* or *-ri*. This rule uniformly applies in the indicative mood across languages, which serves to capture the present/past tense distribution of *-li/-ri* (243c). Lastly, there is a general VI rule that applies in all environments that do not trigger any of the aforementioned VI rules (243d). In all languages, this elsewhere rule realizes the copular form *-ba*, which appears in irrealis environments, e.g., in the subjunctive/infinitival mood, or future tense.

(243) *General VI rules - all copular clauses*

- a.  $[v_{\text{BE}}] \rightarrow ba/tula / \text{Mood}_{[+\text{indic}]} + \text{T}_{[+\text{prs}]} \boxed{\text{Pred}_{\text{INDIV}}}$
- b.  $[v_{\text{BE}}] \rightarrow \emptyset / \text{Mood}_{[+\text{indic}]} + \text{T}_{[+\text{prs}]} \text{ } \_\_\_\_\_\_ f$
- c.  $[v_{\text{BE}}] \rightarrow li / \text{Mood}_{[+\text{indic}]} \text{ } \_\_\_\_\_\_$
- d.  $[v_{\text{BE}}] \rightarrow ba$

In the sections that follow, I will illustrate how this analysis plays out in the Great Lakes Bantu languages surveyed in this work, starting with Kifuliiru (JD63).

### 5.3.2 Kifuliiru copular system

Let us begin first with a discussion of the copular system in Kifuliiru (JD63), as our current data on this language is limited to present tense predicational clauses. In this environment, Kifuliiru exhibits three interpretively distinct verbal copulas: *-li*, *-ba*, and *-tula*. As discussed in chapter 4, these forms broadly encode a distinction between properties that hold of limited circumstances (*-li*), general circumstances (*-ba*), and maximal circumstances (*-tula*).

Given the limited nature of the data, there is only one factor that is relevant to the surface form

of the copula: interpretation. Following accounts like Adger and Ramchand (2003), Markman (2008a), Balusu (2014), I assume that interpretive differences among copular forms arises from the existence of distinct Pred heads, e.g., Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub>.<sup>10</sup> When Pred<sub>STAGE</sub> is merged, the predication relation receives a limited or bounded interpretation. When Pred<sub>INDIV</sub> is merged, the predication relation instead receives a maximal interpretation.

With this in mind, I propose the following set of partially ordered Vocabulary Insertion rules for Kifuliiru  $v_{BE}$  (244). Since the data is limited to one tense-aspect-mood environment, these rules are underspecified with respect to these features and their corresponding heads. To capture the interpretive differences between forms, I assume that  $v_{BE}$  and Pred are jointly realized by a single insertion span (this is modeled by the box in (244a)-(244b), which identifies the span of heads targeted for insertion). As such, the form of  $v_{BE}$  is sensitive to which Pred head is projected in the syntax;  $v_{BE}$  is pronounced as *-tula* if Pred<sub>INDIV</sub> is merged (244a), but as *-li* if Pred<sub>STAGE</sub> is merged (244b). A third form *-ba* appears in all other contexts, functioning as the default or elsewhere form (see discussion below).

(244) *Kifuliiru VI rules - predicational be (present tense only)*

- a. [  $v_{BE}$  ]  $\rightarrow$  *tula* / \_\_\_\_\_ Pred<sub>INDIV</sub>
- b. [  $v_{BE}$  ]  $\rightarrow$  *li* / \_\_\_\_\_ Pred<sub>STAGE</sub>
- c. [  $v_{BE}$  ]  $\rightarrow$  *ba*

The first two VI rules in (244) are unordered, and only differ in which Pred head is projected in the syntax. These rules account for the core interpretive contrast between *-li* and *-tula*. In contrast, the third rule (244c) only applies in cases where the structural conditions imposed by the first two rules are not met. Since (244a) and (244b) specify Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub> as the relevant structural conditions for their application, (244c) should only apply in cases where neither Pred variant is merged. Assuming that non-verbal predication always involves a small clause headed by Pred, there are two possible ways to interpret this.

<sup>10</sup>Following the semantic analysis I propose in chapter 4, I assume that the difference between these Pred variants does not correspond to a traditional stage-individual contrast, but rather a contrast between limited properties and maximal properties.

One possibility is that there is a third Pred variant that encodes the non-limited/non-maximal reading associated with Kifuliiru *-ba*, and  $v_{BE}$  is realized as *-ba* when this Pred head is merged. In this case, all three VI rules in (244) would be equally specific (and as such, unordered), as the only difference between them would be the Pred variant they specify. In this case, *-ba* would not be considered the default or elsewhere form. A second possibility is that *-ba* instead appears when  $v_{BE}$  and Pred are not jointly targeted for insertion; in certain cases,  $v_{BE}$  is realized in isolation, and Pred is left unexponed. Though it is unclear why  $v_{BE}$  and Pred would sometimes be jointly realized by a single insertion span and sometimes not, there is some support for treating *-ba* as a default form, at least in other Great Lakes Bantu languages. For example, an analysis of this type could straightforwardly explain cases of neutralization where  $v_{BE}$  is realized as *-ba* regardless of which Pred head is merged, e.g., in irrealis environments (see Finholt 2024). That being said, it is also possible to account for such data as a case of mood-conditioned allomorphy; in irrealis environments, the span [ $v_{BE}$ +Pred] is always realized as *-ba*. Ultimately I will not attempt to compare these possible analyses here. What is important for our current purposes is that the three Kifuliiru forms can more or less be captured simply by positing the existence of distinct Pred heads.

### 5.3.3 Kihavu and Mashi copular system(s)

In Kihavu and Mashi, we observe three factors that condition the realization of the copula, namely copular clause type, tense-aspect-mood environment, and interpretation. In pure predicational clauses, two copular forms are attested in each language: *-li* and *-ba*. In the present indicative, these two forms encode distinct interpretations of the predication relation, with *-li* yielding a bounded property reading, and *-ba* a more general reading. As before, I attribute this difference to the existence of distinct Pred heads, e.g., Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub>. In other indicative contexts, the form of the copula is directly tied to tense-aspect morphology; past tense/perfective morphology only appears with *-li*, while imperfective morphology only appears with *-ba*. In all non-indicative contexts, only a single copular form *-ba* is attested.

Building on the analysis proposed for Kifuliiru in (244), I propose the following set of ordered



Vocabulary Insertion rules for predicational  $v_{BE}$  in Kihavu and Mashi. In addition to specifying which Pred head is merged, these VI rules also include information related to tense and mood in order to capture the full distribution of the  $v_{BE}$  allomorphs.

(245) *Kihavu/Mashi VI rules - predicational clauses*

- a. [  $v_{BE}$  ]  $\rightarrow$  *ba* / Mood<sub>[+indic]</sub> + T<sub>[+prs]</sub> \_\_\_\_\_ Pred<sub>INDIV</sub>
- b. [  $v_{BE}$  ]  $\rightarrow$  *li* / Mood<sub>[+indic]</sub> \_\_\_\_\_ Pred<sub>STAGE</sub>
- c. [  $v_{BE}$  ]  $\rightarrow$  *ba*

In short, there are three distinct allomorphs of  $v_{BE}$  in predicational clauses. The first form *-ba* (245a) appears in a highly specific environment, i.e., present indicative sentences where Pred<sub>INDIV</sub> is merged. This form corresponds to the non-limited or “permanent” copula that appears exclusively in the present tense and yields general (or even maximal) property readings.<sup>11</sup> The second form *-li* (245b) appears in all indicative sentences (past and present) where Pred<sub>STAGE</sub> is merged. This corresponds to the limited copula that appears in the present tense and yields contextually bounded interpretations, as well as the past tense (or perfective) copula. The final form *-ba* (245c) — which is homophonous with the allomorph *-ba* in (245a) — serves as the default form, and applies in all other contexts (e.g., irrealis environments). For the time being, I will assume that aspect does not play a role in determining the form of the copula, but rather the semantics of the copula (e.g., which Pred head is merged) determines which aspectual markers are available.<sup>12</sup>

The analysis shown in (245) directly accounts for two of the three factors at play in the copular systems of Kihavu and Mashi, namely tense-mood environment (but not aspect), and interpretation. The most straightforward aspect of this analysis is that *-li* (245b) is underspecified for tense, which

<sup>11</sup>Since Kihavu and Mashi do not exhibit three distinct copular forms in the present tense like Kifuliiru, there is no distinction between general and maximal interpretations. Instead, there is a single binary contrast between limited and non-limited/general interpretations.

<sup>12</sup>Specifically, I assume that the semantics of Pred<sub>STAGE</sub> is incompatible with imperfective aspect, meaning that imperfective morphology can only appear with the default form *-ba*. The semantics I posit for the limited/bounded copula (which appears when Pred<sub>STAGE</sub> is merged) restricts the truth of a property-ascription P(x) to a limited circumstance of evaluation. Assuming that this circumstance is interpreted as the smallest circumstance where P(x) is true due to pragmatic factors, it is possible that the limited copula — and therefore Pred<sub>STAGE</sub> — is generally incompatible with imperfective aspect since presumably lacks internal structure. Ultimately however, I will omit any discussion of aspect from this analysis, as it is difficult to capture its effects while also accounting for other factors affecting the form of the copula. More work is necessary to fully address the role of aspect on the form of the copula.

allows it to surface in both the present tense (where it yields a limited/bounded interpretation) and the past tense. This rule also explicitly includes the presence of  $\text{Pred}_{\text{STAGE}}$  in order to capture the use of *-ba* in ostensibly past tense environments, e.g., alongside the imperfective marker *-ga*.

The motivation for the two homophonous *-ba* allomorphs is slightly more complex. In fact, assuming the VI rule for *-li*, one could capture copular allomorphy in predicational clauses with a single VI rule for *-ba* where it is the default (245c). That is, even if we exclude the VI rule in (245a), we can still capture the fact that the copula surfaces as *-ba* in both the present tense (when  $\text{Pred}_{\text{INDIV}}$  is projected) and the future tense ([*-indic*]), since the VI rule for *-li* includes  $\text{Mood}_{[+\text{indic}]}$  and  $\text{Pred}_{\text{STAGE}}$  in its environment. There are however a few reasons to believe that there are two distinct *-ba* allomorphs as opposed to just a single default form. The first reason is typological in nature. In at least one other Bantu language, e.g., Kifuliiru, there are three interpretively distinct allomorphs that appear in the present tense. Like Kihavu and Mashi, Kifuliiru encodes a general contrast between limited and non-limited (or in this case, maximal) property ascriptions using two distinct copular forms *-li* and *-tula*. The third copular form in Kifuliiru, *-ba*, is more default-like; it is associated with more general readings that are both non-limited and non-maximal. In a sense, the Kifuliiru copular system is therefore quite similar to the Kihavu/Mashi copular system. Both systems include one allomorph that encodes a limited interpretation (appears with  $\text{Pred}_{\text{STAGE}}$ ), a second allomorph that directly opposes it and encodes a non-limited/maximal interpretation (appears with  $\text{Pred}_{\text{INDIV}}$ ), and a third allomorph that occurs in all other contexts. The only difference between the two systems is that the non-limited/maximal allomorph in Kifuliiru *-tula* is not homophonous with the elsewhere allomorph *-ba* as in Kihavu and Mashi. It is possible that such a three way interpretive contrast is (or was) present in Kihavu/Mashi, but two of the forms have converged, leaving no overt contrast between maximal interpretations and more general non-limited interpretations.

An additional, albeit weaker piece of evidence for the existence of two homophonous allomorphs — or more specifically, for the existence of a present tense specific allomorph — is that the copula in specificational/equational/identificational clauses has a unique form in the present tense. In non-predicational clauses (which I will group together and refer to henceforth as “specificational

copular clauses” or SCCs), there are three forms of the copula that appear (see (232) for relevant examples involving equational clauses). In the present tense, SCCs are expressed via an agreeing focus marker in isolation. In contrast, SCCs in the past and future tenses (at least in the limited set of Kihavu data discussed previously) involve an overt verbal copula in addition to the focus marker, e.g., *-li* in the past and *-ba* in the future. Assuming then that the focus marker is not itself the copula in present tense SCCs, but is rather accompanied by a zero copula, there are three allomorphs of the copula that appear in SCCs:  $\emptyset$ , *-li*, *-ba*.

For the time being, let us assume that the copula in specificational clauses realizes a distinct variant of the light verbal head  $v_{BE}$ , e.g.,  $v_{BE-SPEC}$  (more on this below).

(246) *Kihavu/Mashi VI rules - specificational be*

- a. [  $v_{BE-SPEC}$  ]  $\rightarrow \emptyset / \text{Mood}_{[+indic]} + \text{T}_{[+prs]}$  \_\_\_\_\_
- b. [  $v_{BE-SPEC}$  ]  $\rightarrow li / \text{Mood}_{[+indic]}$  \_\_\_\_\_
- c. [  $v_{BE-SPEC}$  ]  $\rightarrow ba$

The proposed VI rules for  $v_{BE-SPEC}$  closely reflect those proposed for predicational  $v_{BE}$  in (245), with the only differences being that  $\text{Pred}_{\text{STAGE}}/\text{Pred}_{\text{INDIV}}$  are not explicitly included in any rules (which follows from the fact that there are no interpretive contrasts in SCCs), and that the present tense allomorph in (246a) is null. Under this analysis, both predicational  $v_{BE}$  and specificational  $v_{BE-SPEC}$  have three allomorphs: one that surfaces exclusively in the present indicative (e.g., *-ba/0*), one that surfaces in other indicative environments (e.g., *-li*), and one that surfaces in all non-indicative environments (e.g., *-ba*). In both cases, the present tense allomorph is to some degree exceptional; in predicational clauses it is homophonous with the default copula, and in specificational clauses it is null. I argue that this provides some support for the existence of two *-ba* allomorphs in (245), as it suggests that there are specific forms of the copula that are unique to the present tense.

It is important to note here that the proposal outlined above diverges from the literature quite strongly in one regard. Contra accounts like Heycock and Kroch (1999), Adger (2003), Baker (2003), Balusu (2014), and Myler (2016), a.o., the above VI rules assume that the copula in predicational

clauses realizes a different verbal head, e.g.,  $v_{BE}$ , than the copula in SCCs, e.g.,  $v_{BE-SPEC}$ . Assuming that the core function of the light verbal head  $v_{BE}$  is to link the small clause PredP into the extended verbal projection (Myler 2016), it is unclear why it would be necessary to posit a variant  $v_{BE-SPEC}$  that is unique to SCCs. There is however at least one argument in favor of such a variant. Citing the presence of an agreeing focus marker that co-occurs with (and crucially follows) the copula in SCCs in Kinande, Schneider-Zioga (2021) suggests that the copula in SCCs does not directly embed a small clause, e.g., PredP, but rather a larger focus projection  $fP$  that contains a small clause headed by a predicational head endowed with a specificational semantics, F (Heycock 2012) (121). Given the similarity between the expression of SCCs in Kinande and the languages discussed in this work, I adopt the same general analysis (241). Now, if the structure proposed by Schneider-Zioga (2021) is correct, then there is one clear difference between the light verbal head that realizes the copula in predicational clauses and SCCs:  $v_{BE}$  takes a PredP complement in predicational clauses, but a  $fP$  complement in SCCs. Given that  $v_{BE}$  is however assumed to be a type neutral identity function, it is unclear whether this difference warrants positing two selectionally distinct heads  $v_{BE}$  and  $v_{BE-SPEC}$ . Nonetheless, it is still true that under Schneider-Zioga’s proposal there is a difference in what the light verbal head takes as its complement in predicational clauses and SCCs.

With that being said, it is possible to characterize the allomorphy patterns in predicational clauses and SCCs as allomorphy of a single  $v_{BE}$  head, with a few caveats. As I argued above, the VI rules proposed for predicational  $v_{BE}$  (245) and specificational  $v_{BE-SPEC}$  (246) are largely similar. In both cases, the most specific VI rule applies in the present tense, followed by a less-specific rule that applies in other indicative contexts, and a third, elsewhere rule. More importantly, the latter two of these rules generate the same copular forms in both predicational clauses and SCCs, e.g., *-li* and *-ba*. Given this overlap, it makes sense to treat the zero copula that appears in SCCs as another possible realization of  $v_{BE}$  (specifically one that surfaces uniquely in the present tense) as opposed to a realization of a distinct verbal head  $v_{BE-SPEC}$ . Given the structure of SCCs I adopt from Schneider-Zioga (2021), the relevant conditioning environment for the zero copula involves  $Mood_{[+indic]}$ ,  $T_{[+prs]}$ , and an adjacent focus head  $f$ . The full set of VI rules for  $v_{BE}$  can be seen below

in (247).

(247) *Kihavu/Mashi VI rules - all copular clause types*

- a. [  $v_{BE}$  ]  $\rightarrow$  *ba* / Mood<sub>[+indic]</sub> + T<sub>[+prs]</sub> \_\_\_\_\_ Pred<sub>INDIV</sub>
- b. [  $v_{BE}$  ]  $\rightarrow$   $\emptyset$  / Mood<sub>[+indic]</sub> + T<sub>[+prs]</sub> \_\_\_\_\_ *f*
- c. [  $v_{BE}$  ]  $\rightarrow$  *li* / Mood<sub>[+indic]</sub> \_\_\_\_\_
- d. [  $v_{BE}$  ]  $\rightarrow$  *ba*

Since the copula in SCCs is directly adjacent to the focus head *f*, I assume that  $v_{BE}$  and Pred (or rather, the specificational head F) are not jointly targeted for insertion when *fP* is present. This coincides with the absence of any interpretive contrasts in SCCs.<sup>13</sup>

Aside from the addition of the zero copula (247b), the above VI rules differ from previous iterations in that the rule for *-li* in (247c) does not include Pred<sub>STAGE</sub> in the conditioning environment. By not specifying Pred<sub>STAGE</sub> in (247c), we capture the fact that *-li* appears in the present/past tense in predicational sentences, and also the past tense in SCCs; (247c) applies in non-present indicative environments even if *f* is present. Unfortunately, this comes at a cost. By omitting Pred<sub>STAGE</sub>, we cannot straightforwardly capture the fact that *-ba* may also appear in “past tense” indicative environments, e.g., in the imperfective. Despite this issue, I will adopt the single copula analysis presented in (247) in lieu of positing two distinct heads  $v_{BE}$  and  $v_{BE-SPEC}$ . For the purposes of this work, I will ignore the effect of (imperfective) aspect on the form of the copula. In truth, it is not clear how aspect could be integrated into the analysis without positing multiple distinct forms of both *-li* and *-ba*. With this in mind, I will leave any further discussion of this issue to further work.

To further illustrate how this analysis works, consider the following sentences and their corresponding structures. In the partial trees that follow, I provide only the structure that is immediately relevant to the realization of the copula. Importantly, this means that the subject DP and predicate XP are excluded. Nonetheless, I assume in all cases that the pre-copular DP<sub>1</sub> has moved from a

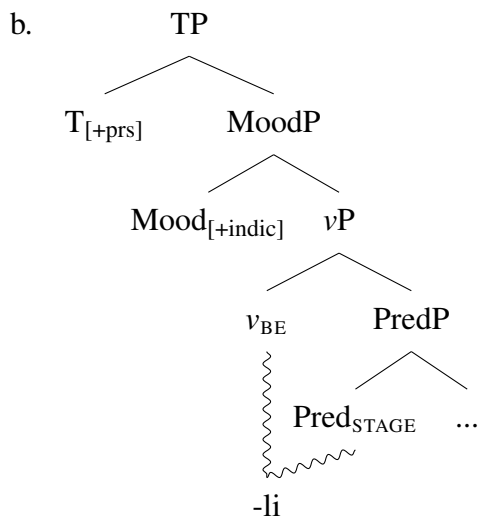
---

<sup>13</sup>I assume throughout that interpretive contrasts arise from the existence of multiple Pred heads that may be realized together with  $v_{BE}$ . Since SCCs do not involve a traditional small clause headed by Pred, interpretive contrasts are not possible.

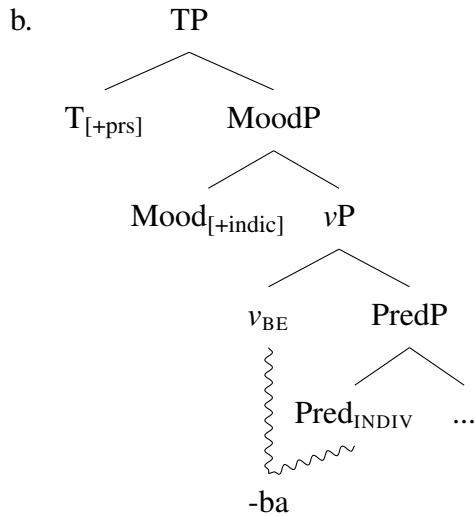
position in the small clause, e.g., PredP in predicational clauses (240), and FP in SCCs (241), to Spec, TP for EPP reasons (Carstens 2005).

In present tense predicational sentences, only two VI rules are relevant, e.g., (247b) and (247c). In this environment, the form of  $v_{BE}$  corresponds to which Pred head is merged:  $v_{BE}$  is realized as *-ba* when Pred<sub>INDIV</sub> is present, but as *-li* when Pred<sub>STAGE</sub> is present. To capture the interpretive contrast between the two forms, I assume that  $v_{BE}$  and Pred are realized by a single insertion span.

- (248) a. Johne a- $\emptyset$ -**li** mu-igiriza  
 John 1SM-PRS-be-<sub>LI</sub> 1-teacher  
 ‘John is a teacher’ (in this context) Mashi



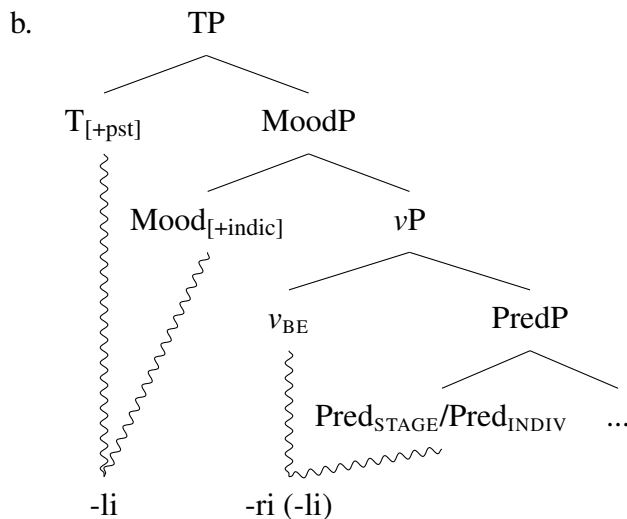
- (249) a. Johne a- $\emptyset$ -**ba** mu-igiriza  
 John 1SM-PRS-be-<sub>BA</sub> 1-teacher  
 ‘John is a teacher’ (his career) Mashi



In the past tense (excluding the imperfective), there is only one relevant rule: (247c). As such,  $v_{BE}$  surfaces as *-ri* (a phonological variant of *-li*) regardless of which Pred head is merged. This results in the absence of interpretive contrasts in the past tense.

- (250) a. Johne a-li-ri ci-shambo  
 John 1SM-PST-be-<sub>LI</sub> 7-thief  
 ‘John was a thief’

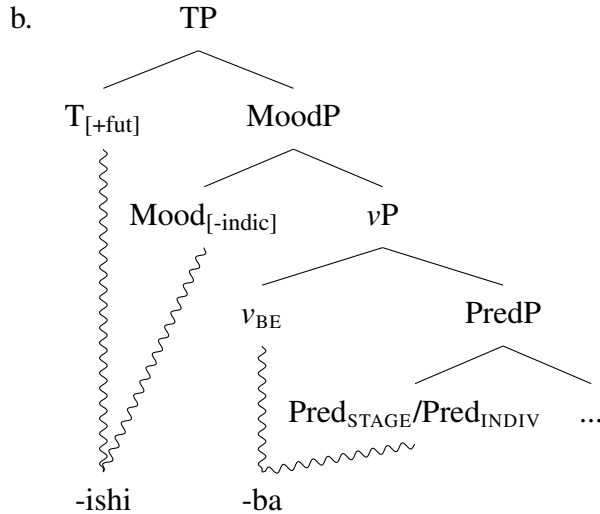
Mashi



In non-indicative (irrealis) environments, e.g., subjunctive mood, infinitival mood, and future tense, only the elsewhere VI rule is applicable (247d). Accordingly,  $v_{BE}$  is obligatorily realized as *-ba* in these environments.

- (251) a. Johne a-ishi-**ba** ci-shambo  
 John 1SM-FUT-be-<sub>BA</sub> 7-thief  
 ‘John will be a thief’

Mashi



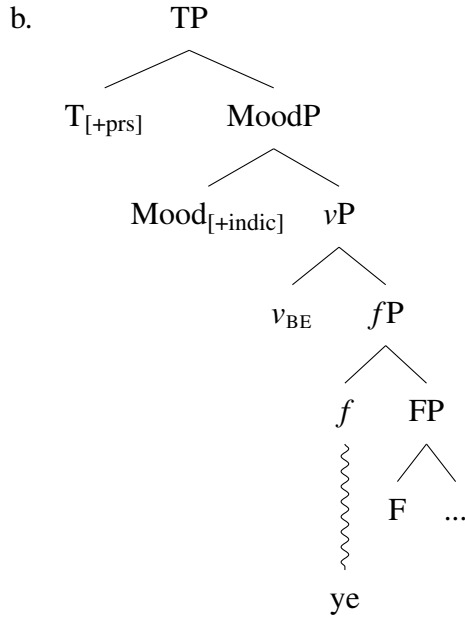
Things are slightly different in SCCs, where there is an additional focus phrase *fP* between the copula and the small clause (in this case FP). Given the intervening head *f*,  $v_{BE}$  is individually targeted for insertion in SCCs. In the present tense, the rule in (247b) always applies, and  $v_{BE}$  is realized as  $\emptyset$ . In other environments however, the form of the copula is unaffected by the presence of the focus phrase, and patterns the same as in predicational clauses:  $v_{BE}$  is realized as *-li* in the past tense, and as *-ba* in irrealis environments.<sup>14</sup>

- (252) a. Nyina wa Mugisha **ye** nyina wa Murhulla  
 1.mother 1SM.LNK Mugisha 1FOC 1.mother 1SM.LNK Murhulla  
 ‘Mugisha’s mother<sub>i</sub> is Murhulla’s mother<sub>i</sub>’

Kihavu

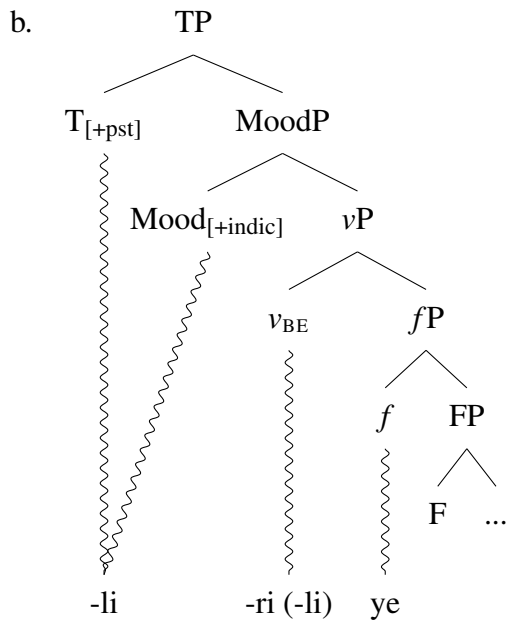
<sup>14</sup>Note that there is a mismatch between the surface order of the focus marker and copula in (253a)/(254a) and the structure(s) in (253b)/(254b). The surface position of the focus marker relative to the copula in Kihavu/Mashi is different to what is reported in Kinande (Schneider-Zioga 2021), suggesting that the order of these elements may vary across Eastern Bantu languages. Moreover, the status of the focus marker as an independent word is unclear; it is possible that the focus marker cliticizes onto the copula in some languages. Given these issues, I will leave any further discussion of the ordering facts to future work.





- (253) a. Nyina wa Mugisha **ye** a-li-ri nyina wa  
 1.mother 1SM.LNK Mugisha 1FOC 1SM-PST-be-LI 1.mother 1SM.LNK  
 Murhulla  
 Murhulla  
 ‘Mugisha’s mother<sub>i</sub> was Murhulla’s mother<sub>i</sub>’

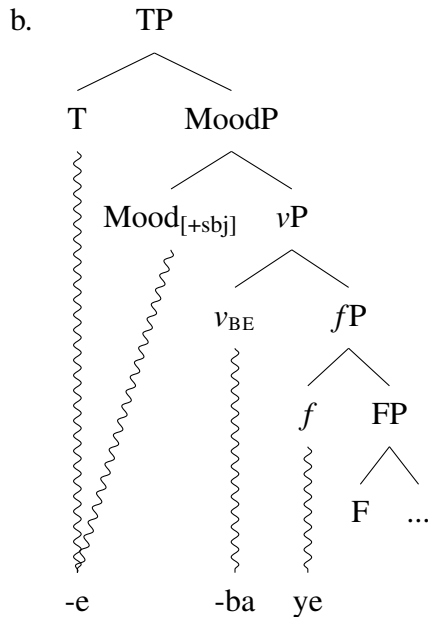
Kihavu



- (254) a. Nyina wa Mugisha **ye** a-be nyina wa  
 1.mother 1SM.LNK Mugisha 1FOC 1SM-be-LI.SBJV 1.mother 1SM.LNK

Murhulla  
Murhulla  
‘Mugisha’s mother<sub>i</sub> will be Murhulla’s mother<sub>i</sub>’

Kihavu



Across predicational clauses and SCCs in Kihavu and Mashi, there is a strong and consistent effect of tense and mood. In the present indicative, there is significant variation in the form of the copula. It is here where we see overt interpretive contrasts, e.g., *-li* vs *-ba* in predicational clauses, and forms that do not appear elsewhere, e.g.,  $\emptyset$  in SCCs. In contrast, the form of the copula is much more regular in other environments. In the past indicative (excluding the imperfective), the form of the copula is always *-li* (*[-ri]*) regardless of copular clause type or which Pred head is merged. Likewise in irrealis environments like the subjunctive, infinitive, or future tense, the form of the copula is obligatorily *-ba*, again regardless of other factors. Importantly, similar patterns are observed in other Bantu languages, including some of those investigated in this work, e.g., the JD60 languages Kinyamulenge and Kirundi. I will return to this topic in section 5.4.

### 5.3.4 Kinyamulenge/Kirundi copular system(s)

Turning now to the remaining two languages discussed in this work, we observe a slightly more complex pattern of copular allomorphy. In Kinyamulenge and Kirundi, there are at minimum three

factors that condition the realization of the copula, namely tense-aspect-mood environment, the person features of the subject, and the category of the predicate.

Unlike what we have seen in other languages, the form of the copula in Kinyamulenge and Kirundi is unaffected by the sub-type of the copular clause. Though SCCs again differ from pure predicational clauses in that they involve the presence of a focus marker *-o*, the form of the copula itself does not differ. Across all copular clause types, three copular forms are attested, including two verbal copulas *-ri* and *-ba*, as well as the invariant copula *ni*. As in Kihavu and Mashi, the distribution of these forms is largely contingent on tense and mood. The most significant variation is observed in the present indicative, where invariant *ni* and verbal *-ri* (and even *-ba* in some cases; see (184)-(185)) are attested. In this environment, there is significant sensitivity to both subject person and predicate category. In examples involving first and second person subjects, only *-ri* is available. With a third person subject however, *-ri* is only available with locative predicates, while *ni* is more widely available. Outside the present indicative the pattern is almost identical to what is observed in Mashi: *-ri* (cognate to *-li*) is obligatory in the past tense, while *-ba* is obligatory in irrealis environments like the future tense.

Among the languages surveyed in this work, the copular systems in Kinyamulenge and Kirundi are somewhat unique in that they involve an invariant copula and exhibit sensitivity to subject person and predicate category. Focusing first on the invariant copula, it is unclear whether this element is truly a “copular form” as operationalized in this work. To this point I have assumed that the various copular forms that appear in predicational sentences and SCCs are verbal elements that realize the light verbal head  $v_{BE}$  (or in some cases, the span  $[v_{BE}+Pred]$ ). By all accounts however, the invariant copula *ni* does not appear to be a true verbal element (Gatchalian 2023). This is perhaps most apparent in its limited distribution and inability to inflect for tense or subject agreement; copular *ni* appears strictly in the present tense, and never co-occurs with tense morphology or subject markers.

Based on these same observations, Gatchalian (2023) proposes that invariant *ni* in Kirundi is not a realization of the light verbal head  $v_{BE}$  like the verbal copula *-ri*. The difference between *-ri* and *ni* boils down to whether  $v_{BE}$  is projected or not; *ni* appears in cases where  $v_{BE}$  is not present.

Citing clause embedding data showing that the complementizer *ko* is incompatible with copular *ni* in embedded clauses in Kirundi, Gatchalian suggests that *ni* is instead located in the left-periphery, realizing the head C.

One of the major advantages of the analysis presented in Gatchalian (2023) is that it offers an account of another crucial aspect of the Kirundi (and Kinyamulenge) data, namely the sensitivity to subject person and predicate category. The effect of subject person is attributed to a broader distinction between speech act participants (first/second person) and non-speech act participants. The fact that *ni* appears exclusively in predications involving a third person subject is taken to be a consequence of the Person-Licensing Condition (PLC) as presented by Béjar and Rezac (2003).

(255) *Person-Licensing Condition (Béjar and Rezac 2003: 53)*

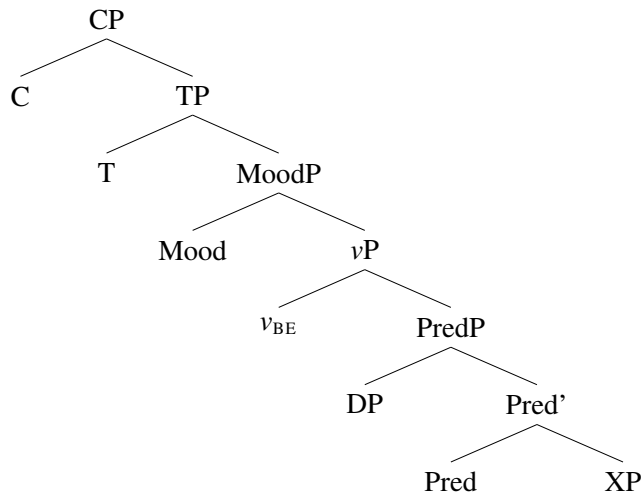
An interpretable 1st/2nd person feature must be licensed by entering into an Agree relation with a functional category.

The idea here is that under a general understanding of Agree, the PLC requires agreement only for first/second person subjects; third person subjects do not have to enter into an Agree relation in the same way. Importantly, subject agreement in non-verbal predication is dependent on the presence of additional verbal structure, e.g.,  $v_{BE}$ , that serves to integrate the small clause into the extended verbal projection. Without  $v_{BE}$ , there is no way to link the small clause to higher functional heads like T or Agr, making subject agreement impossible. In this sense, Gatchalian treats *-ri* — or rather  $v_{BE}$  — as inflectional support; its primary purpose is to integrate the small clause with higher functional heads that facilitate agreement. The effect of this is that  $v_{BE}$  must be projected in predication relations involving a first or second person subject in order to satisfy the PLC. With a third person subject however,  $v_{BE}$  is no longer required as the PLC is not relevant and agreement is not necessary. Assuming that *ni* appears in cases where  $v_{BE}$  is not projected, this means that *ni*-clauses are necessarily tenseless (which explains why they are limited to the present tense) and may only involve a third person subject.

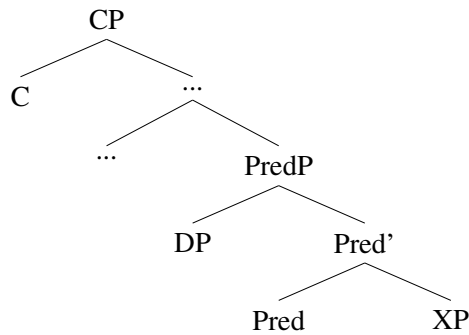
In a similar vein, this analysis also attempts to provide an account for the sensitivity to predicate category. Specifically, Gatchalian assumes that the correlation between *-ri* and locative predicates

arises because locational PP predicates introduce an eventuality variable that needs to be bound (Adger and Ramchand 2003). The idea here is that unlike nominal and adjectival predicates, locational PP predicates introduce an eventuality variable that must be bound by projecting  $v_{BE}$ . As such, though third person predicational clauses do not need to include  $v_{BE}$  to satisfy the PLC,  $v_{BE}$  is necessarily projected if they involve a locative PP predicate. The basic structural difference between clauses involving a verbal copula, e.g., *-ri/-ba*, and invariant *ni* can be seen below.<sup>15</sup>

(256) a. *Copular clause with verbal copula*



b. *Copular clause with invariant copula*



(adapted from Gatchalian 2023)

Though I agree with the general notion that invariant *ni* is not verbal in the same way as other copular forms and therefore may not realize  $v_{BE}$ , I do not fully adopt the analysis proposed by Gatchalian (2023) here. There are two reasons for this. The first is that it is not clear that *ni* realizes C. The motivation in Gatchalian (2023) for treating Kirundi *ni* as a realization of C is based on the

<sup>15</sup>These trees are presented “pre-movement” for ease of comparison and to avoid the assumption made in Gatchalian (2023) that *ni* realizes C.

observation that the complementizer *ko* never co-occurs with an embedded copular clause featuring invariant *ni*. This same co-occurrence restriction is also reported in Kinyamulenge, however *ni* is shown to occur with a complementizer other than *ko*, namely *ngo*, suggesting that *ni* does not realize C (Gluckman 2024). The second reason I do not adopt this analysis is that it assumes à la Kratzer (1995) that not all predicates are predicates of eventualities. To capture the sensitivity to predicate category, Gatchalian assumes that locational PP predicates — but not nominal or adjectival predicates — introduce an eventuality argument that needs to be bound. As discussed in chapter 4, a Kratzerian analysis of this kind cannot account for the fact that almost all non-verbal predicates may co-occur with different copular forms in Eastern Bantu languages, suggesting that differences in copular form cannot be reduced to differences among predicates. In fact, though it has been reported that locational predicates almost always co-occur with *-ri* in Kinyarwanda (Jerro 2015) and Kirundi (Gatchalian 2023), they may also appear with invariant *ni* in closely-related Kinyamulenge, in which case a subtle interpretive contrast arises (175)-(176).<sup>16</sup> Given the significant overlap between the copular systems in these languages, it is therefore unlikely that the correlation between locative/locational predicates and *-ri* can be reduced to a division among predicates that introduce an eventuality argument and predicates that do not. Instead, it is possible that this correlation relates to the semantics of *-ri* (see subsection 4.4.5).

Given the unclear status of invariant *ni*, let us briefly set it aside and instead focus on the other, transparently verbal copular forms in Kinyamulenge and Kirundi, e.g., *-ri* and *-ba*. When considered separately from *ni*, the distribution of *-ri* and *-ba* generally mirrors the distribution of their respective cognate forms, e.g., *-li* and *-ba*, in Kihavu and Mashi, with the only real difference being that the two are largely not in contrastive distribution in the present tense.<sup>17</sup> The effect of tense and mood is otherwise identical: *-ri* appears in the both the present and past tenses, while *-ba* is limited to irrealis environments. Ignoring the limited distribution of *-ri* in the present tense (i.e., the effect of

---

<sup>16</sup>Despite the general effect of predicate category on the form of the copula in Kinyarwanda, Jerro (2015) reports that *ni* is sometimes available with locational predicates in Kinyarwanda, in which case a similar interpretive contrast to that seen in Kinyamulenge arises (see *ri/ni*).

<sup>17</sup>I will again note that *-ba* can appear in one highly specific present tense context in both Kinyamulenge and Kirundi, namely when the subject of a predication relation involving a locative predicate is a class 1 subject (184)-(185). I will omit any discussion of this use of *-ba* here.

subject person and predicate category), this distribution can be captured by the simple VI rules in (257).

(257) *Kinyamulenge/Kirundi VI rules - simplified*

a. [  $v_{BE}$  ]  $\rightarrow$  *-ri* / Mood<sub>[+indic]</sub> \_\_\_\_\_

b. [  $v_{BE}$  ]  $\rightarrow$  *ba*

Outside constructions involving a locative predicate, there are no cases where multiple forms of  $v_{BE}$  are in contrast in the present tense unlike in Kihavu, Mashi, and Kifuliiru. As such, there is no need to invoke distinct Pred heads in the above VI rules, nor include a unique, interpretively distinct present tense form, e.g., Kihavu/Mashi *-ba*. Instead, the most specific VI rule in this case is (257a), which ensures that *-ri* (like its cognate *-li* in other languages) generally appears in both the present and past tense, but never in irrealis environments (e.g., the future tense). Once again, (257b) serves as the elsewhere rule, ensuring that  $v_{BE}$  is realized as *-ba* in irrealis environments across all copular clause types.

The analysis in in (257) presents a simple account of the verbal allomorphs *-ri* and *-ba* in Kinyamulenge and Kirundi that aligns quite nicely with the analysis proposed for Kihavu and Mashi in (247). As appealing as this analysis may be, its inability to account for the restrictive distribution of *-ri* in the present tense, and perhaps more glaringly, the complete omission of invariant *ni*, leaves much to be desired. A sufficient account of the Kinyamulenge/Kirundi copular system(s) should capture the full distribution of all copular forms, yet the analysis proposed above fails to capture the full distribution of even a subset of forms. In what follows, I will outline one possible alternative analysis that could account for the full distribution of all three copular forms, e.g., *-ri*, *-ba*, and *ni*, including subject person and predicate category effects.

As previously discussed, the status of invariant *ni* as a realization of the light verbal head  $v_{BE}$  is cast into doubt by its inability to bear inflectional morphology. This is something that crucially differentiates *ni* from the other copular forms discussed in this work, and is one of the central arguments for the analysis in Gatchalian (2023), where Kirundi *ni* is assumed to appear in copular clauses where  $v_{BE}$  is not projected. With that being said, there are a few reasons to think that *ni* may

be best analyzed as a  $v_{BE}$  allomorph.

Though invariant *ni* does not appear in any of the other languages investigated in this work, its distribution is somewhat similar to some of the copular forms in these languages. One of the key distributional facts about *ni* is that it only appears in the present tense. We have already seen that there are other copular forms that are similarly present-specific in this way, e.g., the non-limited copula *-ba* and the zero copula in Kihavu and Mashi. Like these forms, *ni* exists in contrastive distribution with another copular form in the present tense, specifically *-ri*. This follows a general pattern seen in the languages discussed where copular variation is most substantial in the present tense, as languages often exhibit multiple contrastive forms in this environment. Previous accounts of the *ni/ri* contrast in Kinyarwanda and Kirundi have generally characterized the alternation in rigid terms, positing that the two forms occur with distinct predicates; in copular clauses featuring a third person subject, *-ri* appears with locative predicates, and *ni* appears with nominal and adjectival predicates (Jerro 2015, Lafkioui et al. 2016, Gatchalian 2023). I have already shown however that this is not the case in Kinyamulenge, as both *ni* and *-ri* (and even *-ba* in some cases; see chapter 4) may be used with a locative predicate, with each yielding a distinct interpretation of the predication relation (175)-(176). Given the existence of present-tense specific interpretive contrasts among verbal copular forms in Kihavu, Mashi, and Kifuliiru, these facts provide some support for analyzing *ni* as a  $v_{BE}$  allomorph.

A more compelling argument in favor of this analysis stems from the diachronic origin of *ni* as a focus marker (McWhorter 1994, Schwarz 2003, Güldemann 2003). McWhorter (1994) attributes the development of the invariant copula *ni* in Swahili (G42) to its previous use as a predicate focus marker in cleft constructions. In later work, Güldemann (2003) provides a means for accounting for the uniquely present tense distribution of *ni* by positing a connection between predicate focus and the present (progressive) that facilitates the grammaticalization of a predicate focus marker as a marker of present progressive aspect.<sup>18</sup> One possibility then is that *ni* in JD60 languages has been reanalyzed as a copula that includes explicit information about tense and aspect; contra Gatchalian

---

<sup>18</sup>Specifically, Güldemann (2003) argues that the present progressive inherently involves focus in the sense that it emphasizes the ongoing nature of an event as the most salient information in the discourse.



(2023), *ni*-clauses are tensed clauses. If this is the case, it follows that  $v_{BE}$  must project in *ni*-clauses, as this is the only way to integrate the predicational small clause PredP with the functional heads T/Asp.

Assuming then that *ni* is a realization of  $v_{BE}$  that is unique to the present tense and can occur with predicates of any category (as in Kinyamulenge), there remain only two things to account for, namely that *ni* can only occur with a third person subject, and *-ri* can only occur with a third person subject if a locative predicate is also present. Addressing the first of these issues is relatively straightforward. As in Gatchalian (2023), the correlation between *ni* and the presence of a third person subject can be attributed to the PLC (255); since first and second person subjects must enter into an Agree relation to satisfy the PLC, they can only occur with a copular form that can inflect for agreement, e.g., *-ri*. The *-ri* copula that occurs with first and second person subjects therefore serves a purely inflectional purpose as a host for agreement morphology. Importantly, the applicability of the PLC here finds support from similar person-based effects in other Bantu languages where an inflectional verbal copula is specifically used for first and second person. This can be seen for example in Lamba (M54), where *-li* (cognate to Kinaymulenge/Kirundi *-ri*) is used with first and second person subjects, while a tonal marking strategy is used in lieu of an overt copula with third person subjects (41).

As for the correlation between *-ri* and locatives, an analysis is not so simple. Looking across Bantu languages, it is common to find reflexes of proto-Bantu *\*-de*, e.g., *-li/-ri*, used to express locative predication, with some even serving as a dedicated locative copula (McWhorter 1994, Gibson et al. 2019). However, the reason for this is not immediately clear. One possibility is that this stems from the semantics of the copula. In chapter 4, I argue that the general semantic function of *-li/-ri* in Eastern Bantu languages is to restrict the truth of a property ascription to a limited circumstance of evaluation defined by some relevant contextual parameter, e.g., time, location, world, etc.. Given the limited interpretive contrasts between copular forms in JD60 languages, it is possible that the meaning of *-li/-ri* has gradually been lost. However, this meaning may have been preserved with locative predicates due to its location-specific uses; the use of *-li/-ri* to express that

a property ascription is limited to a particular location resulted in the strong association between *-li/-ri* and locative predicates, which generally express information about locations. In this scenario, this resulted in a synchronic situation where Kinyamulenge/Kirundi *-ri* can only have an interpretive effect with a locative predicate. The reason this is limited to predication relations involving third person subjects is because this is the only context where it is in competition with *ni*. Such an account could feasibly also explain the fact that *-ba* may also appear in certain locative predications, e.g., those involving a class 1 human subject (184)-(185).

With this in mind, an account of  $v_{BE}$  allomorphy including *ni* can largely be derived from the existence of the PLC and the present-orientation of *ni*. In the present tense, the copula is uniformly realized as *-ri* with first and second person subjects in order to satisfy the requirements of the PLC. With a third person subject *ni* is generally used, however *-ri* (and *-ba* in some cases) can be used to make interpretive contrasts specifically with locative predicates. In other tense-mood environments, the pattern is the same as before: *-ri* is used in the past tense, and *-ba* in all irrealis environments.

This analysis is partially illustrated by the VI rules in (258). In order to capture the contrast between *ni* and *-ri* and still account for the effect of tense-mood, there is a need to distinguish between the *-ri* form that appears specifically with locatives (258a) and the homophonous form *-ri* that appears more broadly in the indicative (258c) (e.g., with first/second person subjects in the present tense and in all cases in the past tense).

(258) *Kinyamulenge/Kirundi VI rules - all copular clause types*

- a. [  $v_{BE}$  ]  $\rightarrow$  *-ri* / Mood<sub>[+indic]</sub> + T<sub>[+prs]</sub> \_\_\_\_\_ Pred<sub>STAGE</sub> Loc
- b. [  $v_{BE}$  ]  $\rightarrow$  *ni* / Mood<sub>[+indic]</sub> + T<sub>[+prs]</sub> \_\_\_\_\_
- c. [  $v_{BE}$  ]  $\rightarrow$  *-ri* / Mood<sub>[+indic]</sub> \_\_\_\_\_
- d. [  $v_{BE}$  ]  $\rightarrow$  *ba*

The rule in (258a) applies only in present tense locative constructions involving a Pred<sub>STAGE</sub> head, capturing the interpretation of *-ri* in this environment. The following rule (258b) applies in (most) other present tense environments so long as it is not independently ruled out by the PLC.<sup>19</sup>

<sup>19</sup>This excludes cases where *-ba* can occur in the present tense, which are attested in Kinaymulenge and Kirundi.

The remaining rules work as before, and serve to capture the observed tense-mood effects.

Although the proposed analysis for Kinyamulenge/Kifuliiru differs somewhat substantially from the analysis proposed for Kihavu and Mashi, the overall patterns remain quite similar. In both cases, there is a strong effect of tense/mood such that contrast between copular forms is limited to a single environment, e.g., the present tense. In all other tense/mood environments, only a single copular form is attested, e.g., *-li/-ba*. The reason why the Kinyamulenge/Kirundi copular system includes two present tense specific allomorphs is because it includes a copular form that is not present in Kihavu/Mashi: invariant *ni*. I suggest that, over time, the addition of *ni* as a copula (or rather, the reanalysis of *ni* from predicate focus marker to copula) resulted in the loss of interpretive contrasts in most environments. Though originally a copula used to yield limited interpretations, this resulted in *-ri* being primarily used as inflectional support; *-ri* appears in cases where agreement or tense morphology is obligatory, e.g., with first/second person subjects or in the past tense. The use of *-ri* with locatives is taken to be a remnant of its former, limited meaning; the interpretive effect of *-ri* only appears synchronically in clauses featuring locative predicates and third person subjects. As I discuss in section 5.4, this proposal falls in line with the general observation that the form of the copula in Great Lakes Bantu languages is influenced by both the syntax and the semantics; in some cases the form of the copula is tied to the interpretation of the predication relation, and in other cases it is tied to its morphosyntactic context.

## **5.4 Discussion and conclusion**

In this chapter, I showed that the expression of non-verbal predication in certain Great Lakes Bantu languages is sensitive to a number of different environmental factors, including interpretive context, copular clause type, and tense-aspect-mood environment, among others. In general, the effects of these different factors were found to be similar across languages, with consistent patterns in the expression of non-verbal predication observed in various environments. A particular emphasis in this chapter was placed on patterns of copular allomorphy, as the form of the copula was shown to be affected by various factors in its environment. That said, the effect of these environmental factors

was not limited to the form of the copula, and could in fact be seen in the expression of copular clauses more broadly.

To begin our discussion, let us focus on the latter of these observations. In addition to playing a role in determining the form of the copula, certain conditioning factors had a broader effect on the structure of copular clauses. Specifically, it was shown that different types of copular clauses involve distinct structures. One of the hallmarks of the four-way copular clause taxonomy proposed by Higgins (1979) is that certain clause types involve a reversal of arguments; in languages like English, specificational, identificational, and equational clauses (which all relate two nominal elements in a predication relation) differ from pure predicational clauses in that they situate the predicate in a pre-copular position and the logical subject in a post-copular position (56). Unsurprisingly, the same reversal of arguments is attested in the Bantu languages surveyed in this work; specificational, identificational, and equational copular clauses generally involve a reversal of the logical subject and predicate around the copula. Unlike in English however, the difference between pure predicational clauses and specificational, identificational, and equational clauses (which I will again jointly refer to as SCCs) in these languages is not limited to the relative ordering of the subject and predicate, as SCCs include an additional piece of structure that is absent from predicational clauses: the focus marker *-o*.

In all of the languages surveyed, SCCs differ from predicational clauses in that they employ a focus marker alongside the copula to express a non-verbal predication relation. This is similar to what Schneider-Zioga and Mutaka (2015a) and Schneider-Zioga (2018, 2021) report of SCCs in related Kinande (JD42). The focus marker is attested (albeit not always obligatory) across all types of SCCs (e.g., specificational clauses, equational clauses, and indentificational clauses) in all of the languages for which we have sufficient data. In most cases, the focus marker appears together with an overt form of the copula, however the focus marker exceptionally appears with a zero copula in present tense SCCs in two languages, namely Kihavu and Mashi. The form of the copula in all other cases largely follows the consistent patterns of copular allomorphy observed in pure predicational clauses (discussed below). From a typological perspective, it is relatively

unsurprising that specificational, identificational, and equational clauses pattern similarly, as this is regularly observed across languages (see subsection 2.3.2). However, if we assume that there are four distinct kinds of copular clauses as proposed by Higgins (1979), the fact that these three clause types pattern uniformly is somewhat unexpected, particularly given the substantial variation we see in the copular clauses overall in these languages. If there are in fact four distinct types of copular clauses, one would expect to find languages where these clause types are morphologically distinguished. However, just like many other languages that feature complex patterns of copular variation, there is only one broad division made between clause types in the languages discussed here, namely the division between pure predicational clauses on the one hand and specificational, identificational, and equational clauses on the other. In this sense, the patterns observed in this work broadly support a collapse of the four-way classification in Higgins (1979) and suggest that there may be as few as two types of copular clauses: predicational and specificational (see Arche et al. 2019 for further discussion on this topic).

Related to the question of copular clause type distinctions is a question about the semantic function of the copula: is the meaning of the copula the same across different copular clauses? For many, the existence of distinct copular clause types suggests the existence of semantically distinct copulas; the interpretive differences between predicational, specificational, equational, and identificational arise from semantically distinct copulas (see subsection 3.2.2). For others however, these differences are not attributed to the copula itself, but rather structural differences across clause types. For accounts of this type, there is a single meaningless copula in all copular clause types that serves to link the predicational small clause with higher functional projections (Myler 2016, 2018).

In the Great Lakes Bantu languages investigated in this project, copular clause type plays little role in determining the form of the copula. As mentioned previously, there is only one context in which the form of the copula in certain clause types does not adhere to more general patterns of allomorphy. Specifically, present tense SCCs in Kihavu and Mashi obligatorily involve a zero copula, while present tense predicational clauses exhibit multiple copular allomorphs. Outside of this one context, the form of the copula in SCCs patterns just like the copula in pure predicational

clauses (see discussion below). It is for this reason that I elect to treat the copular forms in all clause types as allomorphs of the same light verbal head  $v_{BE}$ . In this sense, I argue that the current data supports a single copula analysis in which interpretive contrasts between copular clause types are not tied to the copula itself. Instead, I assume that the interpretive differences between clause types — more specifically pure predicational clauses and SCCs — arise from the existence of two different predicational heads, e.g., Pred and F, that can be used to relate two DPs in a small clause (Heycock 2012, Schneider-Zioga 2021). As such, while there is a semantic difference between pure predication and specification/equation/identification, this does not correlate to two distinct copulas, but rather two distinct predicational heads. In all cases there is a single copula  $v_{BE}$ . The difference between clause types reduces to a difference between a predicational small clause PredP (e.g., in pure predicational clauses) and a specificational small clause FP (e.g., in SCCs).

Shifting our focus to the broader effect of environmental factors on the form of the copula, generalizable patterns were found to hold across languages. Specifically, it was shown that related copular forms in different languages tend to exhibit similar distributions, resulting in regular patterns of allomorphy. Across the languages investigated in this work, two key patterns of allomorphy were observed. The first of these patterns relates to the semantics; the form of the copula differs according to the interpretation of the predication relation. As discussed in chapter 4, different forms of the copula in the present tense encode distinct interpretations; reflexes of proto-Bantu *\*-de* (e.g., *-li/-ri*) correlate with limited interpretations of property ascriptions, while reflexes of proto-Bantu *\*-bà* (e.g., *-ba*) correlate with more general interpretations. In contrast, the second pattern relates more directly to the syntax; the form of the copula is sometimes determined by its morphosyntactic environment. The most salient example of this is the restrictive effect of tense and mood. Though each language surveyed exhibits multiple (interpretively distinct) forms of the copula in the present tense, this variation is fully neutralized in other tense/mood environments; only *-li/-ri* (i.e., reflex of proto-Bantu *\*-de*) is available in the past tense, while only *-ba* (i.e., reflex of proto-Bantu *\*-bà*) is available in irrealis environments. A similar sensitivity to morphosyntactic context is attested in a subset of the JD60 languages discussed in this work, e.g., Kinyamulenge and Kirundi. In

these languages, the form of the copula is restricted according to the person features of the subject; invariant *ni* can only occur with third person subjects.

The core similarity among the copular systems in these languages is their simultaneous sensitivity to these semantic and syntactic pressures. Although individual languages may involve different combinations of conditioning factors, the form of the copula in all languages is affected by both interpretive and morphosyntactic factors. The existence of interpretive contrasts among copular forms in the present tense suggests that there is a pressure to map the form of the copula to meaning; the form of the copula should reflect its interpretive function. At the same time, the neutralizing effect of tense/mood and even subject person suggests that this form to function mapping is sometimes overruled; the form of the copula is sometimes determined by its morphosyntactic environment, regardless of interpretative factors. It is the interaction of these pressures that yields the generalizable patterns observed across languages. In contexts where the syntax does not fundamentally restrict the form of the copula, e.g., the present tense, the form of the copula is free to map directly to the semantics. It is precisely in these environments where copular variation is most extensive; unless otherwise constrained, the pressure to map form to function results in the availability of multiple distinct copular forms in a single morphosyntactic environment.

The idea that the form of the copula is subject to these competing pressures is directly reflected in the analysis I propose in section 5.3. In brief, I propose a contextual allomorphy analysis that treats the various copular forms observed as allomorphs of the light verbal head  $v_{BE}$  that appear in particular contexts. To account for the sensitivity to both interpretation and tense/mood environment, I specifically analyze this as a case of span-conditioned allomorphy (Merchant 2015). Assuming that interpretive differences among copular forms is attributed to the existence of different Pred heads (Adger 2003, Markman 2008a, Balusu 2014), I take the copula to jointly realize the heads [ $v_{BE}$  + Pred]. In general, the form of the copula is determined by which Pred head is merged; *-li/-ri* appears when  $Pred_{STAGE}$  is merged, while *-ba* appears when  $Pred_{INDIV}$  is merged. In order to capture to the effect of tense and mood, an additional span [Mood + T] is proposed. This span serves to condition the realization of the lower span [ $v_{BE}$  + Pred], and is associated with its own

discrete exponent.<sup>20</sup> The particular heads and/or features included in this higher span determine whether or not the form of the copula reflects the Pred head that is merged. In the present tense, we see interpretive contrasts among copular forms. As such, if this span consists of the heads [Mood<sub>[+indic]</sub> + T<sub>[+prs]</sub>], the form of the copula is determined by Pred; there is a direct mapping from form to function. Note that in this case, there is no overt exponent of the conditioning span [Mood<sub>[+indic]</sub> + T<sub>[+prs]</sub>], as the present tense is not morphologically marked in any of the languages surveyed. In contrast, other tense/mood environments do not exhibit similar contrasts between copular forms. If the conditioning span includes either T<sub>[+pst]</sub> or Mood<sub>[+indic]</sub>, the form of the copula is not sensitive to which Pred head is merged. In these cases, the form of the copula is determined by the tense/mood environment itself; the copula is realized as *-li/-ri* when T<sub>[+pst]</sub> is present, but as *-ba* when Mood<sub>[+indic]</sub> is present. Importantly, the neutralizing effects of tense/mood in these contexts correlate with overt exponents of the span [Mood + T]. In this sense then, the form of the copula corresponds to its meaning unless [Mood + T] is overtly expounded, in which case the form of the copula is determined by the particular combination of heads and/or features that make up that span.<sup>21</sup>

Though the effect of subject person in Kinyamulenge and Kirundi is not overtly included in VI rules I propose in (258), a similar neutralization effect is observed; in copular clauses involving a first or second person subject, only *-ri* is attested. The other primary copular form in these languages, e.g., invariant *ni*, is restricted to clauses involving a third person subject. In my analysis, I follow Gatchalian (2023) and attribute this effect to the Person-Licensing Condition (Béjar and Rezac 2003) as described in (255); invariant *ni* is incompatible with first and second person subjects because they must enter into an agree relation with a functional head (e.g., T/Agr) to satisfy the PLC, and *ni* cannot bear inflectional morphology. In order to satisfy the PLC in these cases, the copula is instead realized as the verbal copula *-ri*, which crucially can host inflectional morphology.

<sup>20</sup>Discrete exponents of Mood and T do not co-occur in the data, suggesting that Mood and T may indeed serve as a single insertion domain.

<sup>21</sup>Importantly, I do not consider subject agreement morphology as part of the span [Mood + T] for the purposes of this argument. As I argue in the case of invariant *ni* in Kinyamulenge/Kirundi however, it is possible that overt subject agreement — whatever head it realizes in the syntax — similarly determines the form of the copula in the sense that it necessitates a verbal copula, e.g., *-ri/-ba*.



It is for this reason that Gatchalian (2023) describes *-ri* in these contexts as inflectional support; its sole function is to host inflectional morphology so that the PLC can be satisfied. Note that once again, this neutralizing effect correlates with overt inflectional morphology associated with higher functional heads, e.g., T/Agr. Though overt subject agreement does not restrict variation among verbal copulas in other languages like Kihavu and Mashi, it does restrict the form of the copula in Kinyamulenge and Kirundi to some degree; when subject agreement is necessitated by the PLC (e.g., with a first/second person subject), the copula can only be realized as *-ri*.<sup>22</sup> Taken together with the effect of tense/mood, this suggests that the neutralization of copular contrasts is closely tied to the presence of overt inflectional morphology. Whether this generalization is correct — and if so, why these two things correlate with one another — is not immediately clear. With that being said, I will leave any further discussion of this relationship to future work. For the time being, it suffices to say that different morphosyntactic factors impose similar restrictions on the form of the copula, and that these restrictions are often associated with overt inflectional morphology.

With respect to the applicability of the analysis proposed in section 5.3 to the current data, one positive is its ability to capture the extensive similarities observed across languages. The nearly identical effects of tense, mood, and interpretive context in the languages in question are reduced to a single set of generalizable VI rules. Differences across languages are assumed to be derived from subtle changes to these base rules or through the addition of supplementary language-specific VI rules. This allows the analysis to capture broad patterns of allomorphy that are consistent across languages while also accounting for variability in the set of copular forms and relevant conditioning factors observed in each language.

That being said, one of the weaknesses of the proposed analysis is that it doesn't always provide a satisfactory explanation of said language-specific differences. Take for example the three-way interpretive contrast in Kifuliiru. Assuming that interpretive contrasts are derived from the existence

---

<sup>22</sup>Note that the other verbal copula in Kinyamulenge and Kirundi, e.g., *-ba*, is only attested with third person subjects where the PLC is not relevant. This provides further support for the existence of two homophonous *-ri* allomorphs as proposed in (258). The more specific *-ri* allomorph applies uniquely with third person subjects and locative predicates, in which case it exhibits a limited interpretation of the predication relation (i.e., it alternates with *ni/-ba*). The more general *-ri* allomorph serves as inflectional support; it surfaces in cases where there is overt agreement morphology (e.g., with first/second person subjects) or overt tense morphology (e.g., in the past tense).

of distinct Pred heads, I speculate that there are two logical possibilities for why Kifuliiru exhibits three copular variants in the present tense. Either it is the case that there is a third Pred variant in addition to Pred<sub>STAGE</sub> and Pred<sub>INDIV</sub> that, when pronounced together with  $v_{BE}$ , yields *-ba* (i.e., non-limited and non-maximal), or  $v_{BE}$  is realized as *-ba* when  $v_{BE}$  and Pred are not jointly realized, for whatever reason. Both of these options face their own issues. While the first option ensures uniformity across the three forms, it is at odds with previous approaches that assume the existence of only two Pred heads (Adger 2003, Markman 2008a, Balusu 2014). On the other hand, though the second approach does not stray from the literature in this regard, it necessitates that the insertion span [ $v_{BE}$  + Pred] be arbitrarily broken up in cases where *-ba* surfaces. Though I note that this could explain instances where *-ba* is used regardless of which Pred head is merged (e.g., in irrealis environments), there is no clear explanation for why this span should sometimes be broken up in this way.

In a similar vein, this analysis provides a relatively limited account of certain morphosyntactic effects in Kinyamulenge and Kirundi. One of the primary issues in adapting the proposed analysis to the Kinyamulenge/Kirundi data relates to the appearance of *-ri* in present tense copular clauses featuring a third person subject and locative predicate. Though I suggest that the strong correlation between *-ri* and locative predicates in these languages may have something to do with its meaning as a limited copula (see chapter 4), the exact diachronic process that resulted in this correlation in Kinyamulenge and Kirundi — in addition to many other Bantu languages (Schneider-Zioga 2018, Gibson et al. 2019) — is unclear. Even the proposed account of subject person effects — which I attribute to the Person-Licensing Condition (PLC) as in Gatchalian (2023) — faces significant questions, as there is no explanation provided for why the PLC is relevant in Kinyamulenge/Kirundi but not in any of the other languages surveyed. Though this likely has to do with the absence of an invariant copula in those languages, the point remains that there are open questions regarding the analysis' ability to fully account for the Kinyamulenge/Kirundi copular system(s).

Ultimately further work is needed to fully address these issues, among others. Ideally a more comprehensive analysis of copular allomorphy in Great Lakes Bantu language would involve full

copular paradigms from a wider variety of languages than those included here. By extending the scope of this project, it is possible that some of the issues facing the current analysis would find simple solutions. Given the limitations of this project however, I will leave these issues unresolved. In any case, the analysis proposed in this chapter represents just one possible way of accounting for the clear and complex patterns of copular allomorphy attested in the languages surveyed in this work. Although alternate analyses of the data may be possible, these approaches would still need to address these patterns, meaning that they would likely face similar issues to those discussed here. Analysis aside, the major takeaway from this chapter is that the form of the copula in this group of Great Lakes Bantu languages is sensitive to two competing pressures, one semantic, and one syntactic. In some cases, the form of the copula is dictated by the semantics; unless otherwise restricted, the form of the copula reflects the interpretation of the predication relation it helps realize. In other cases however, the form of the copula is dictated by the syntax; certain morphosyntactic environments impose restrictions on the form of the copula regardless of interpretation.

## Chapter 6

### Conclusion

In this work, I presented a novel description of copular variation in five Great Lakes Bantu languages, namely Kihavu (JD52), Mashi (JD53), Kinyamulenge (JD61a), Kirundi (JD62), and Kifulliru (JD63). I demonstrated that copular variation in this group of languages is modulated by two factors: interpretive context and morphosyntactic context. Despite subtle differences across languages, I showed that there are generalizable patterns of variation that reflect these two sensitivities: cognate forms in different languages exhibit similar interpretive profiles and morphosyntactic sensitivities.

In chapter 4, I provided an overview and analysis of interpretive contrasts among copula forms in present tense predicational clauses. Across the five languages surveyed, three types of interpretive contrasts were discussed, e.g., the binary contrast between *-li* and *-ba*, the binary contrast between *-ri* and *ni*, and the ternary contrast between *-li*, *-ba*, and *-tula*. The same core interpretive contrast was observed in all three types: different copular forms distinguish property ascriptions that hold of limited contexts from property ascriptions that hold of more general contexts. Importantly, this general contrast was reflected in the similar interpretive profiles of cognate forms in different languages. In general, reflexes of the proto-Bantu copula *\*-de* (e.g., *-li/-ri*) were shown to express limited property readings, while reflexes of the proto-Bantu copula *\*-bà* (e.g., *-ba*) were shown to express more general readings. To account for these contrasts, I presented an adapted version of the semantic analysis proposed in Deo et al. (2017) that treats interpretively distinct copular forms as presuppositional variants.

In chapter 5, I turned my focus to the effect of morphosyntactic context on the form of the copula. I showed that the form of the copula in the surveyed languages is sensitive to different combinations of morphosyntactic factors, including copular clause type (or rather, the presence of a focus marker),

tense-aspect-mood environment, predicate category, and subject person. Despite these differences, similar patterns of variation were found across languages. Once again, cognate forms were shown to exhibit similar distributional sensitivities, particularly to tense/mood environment. Though all copular forms were attested in present indicative (predicational) clauses, other tense/mood environments were associated with particular forms: only reflexes of proto-Bantu *\*-de* (e.g., *-li/-ri*) appeared in the past tense, and only reflexes of proto-Bantu *\*-bà* (e.g., *-ba*) appeared in irrealis environments. To account for this data, I presented a generalizable contextual allomorphy analysis that captures the effect of tense/mood while also accounting for interpretive contrasts in the present tense. Ultimately I argued that the patterns of copular variation observed in these languages reflects a simultaneous sensitivity to interpretive context and morphosyntactic context; though the form of the copula generally maps to its semantic function, this mapping is lost in cases where the morphosyntactic context determines the form of the copula.

Alongside its broader typological contributions, this work also comments on some major theoretical questions regarding copulas and non-verbal predication. With respect to interpretation, this work provides novel support against a binary division between stage/individual-level predication in the sense of Kratzer (1995). Like Roy (2013), I suggest that copular contrasts involve more fine-grained distinctions, including at least three interpretively distinct kinds of predication relations. At the same time, this work also provides evidence against analyses that attribute interpretive contrasts to semantic or syntactic differences among predicates (Kratzer 1995, Roy 2013). Instead, I argue that the lack of distributional differences among copulas in the languages surveyed provides support for accounts that attribute interpretive contrasts to the semantics of the copula, specifically those that treat different copular forms as presuppositional variants (e.g., Clements 1988, 2006, Maienborn 2005, Deo et al. 2017).

In addition, this project also addresses pertinent questions related to the number of (semantically distinct) copulas and copular clause types. In general, the effect of copular clause type on the form of the copula was found to be minimal in the Bantu languages surveyed. As discussed in chapter 5, the only major effect of copular clause type is that it fully neutralizes interpretive

contrasts in the present tense; interpretive contrasts were only attested in present tense predicational clauses. Aside from this, identical patterns of morphosyntactically-conditioned variation were observed across all clause types. In light of this overlap and previous work that takes the copula to realize a light verbal head  $v_{BE}$  (Adger and Ramchand 2003, Baker 2003, Balusu 2014, Myler 2016, 2018, a.o.), I treat the copula identically across all copular clause types; in all cases, the copula realizes the light verbal head  $v_{BE}$ . In this sense, there are no overt differences in copular form across clause types to support any divisions between the four clause types (e.g., predication, specification, equation, and identification) in Higgins (1979). Though this does not directly resolve questions related to the number of (semantically distinct) copulas and copular clause types, it minimally shows that clause type differences are not reflected in the form of the copula in some Great Lakes Bantu languages. That being said, there is syntactic evidence that these languages broadly distinguish pure predication from specification, equation, and identification. As reported in other Bantu languages like Kinande (Schneider-Zioga 2018, 2021), specificational, equational, and identificational clauses in the languages surveyed involve the addition of a focus marker *-o* alongside the copula — something that does not occur in pure predicational clauses. This coincides with the commonly discussed difference in argument ordering between predicational clauses and other types of copular clauses; specificational, equational, and identificational clauses typically involve a reversal of arguments around the copula, e.g., a PREDICATE-SUBJECT orientation.

Moving forward, there are a number of questions that arise from the observations and analyses presented in this work. With respect to the typology, one relevant question is, How generalizable are these patterns of copular variation? Though we find that copular variation in this small group of Great Lakes Bantu languages is sensitive to certain morphosyntactic and interpretive factors, how generalizable are these sensitivities and the patterns that result from them? From a Bantu perspective, there is some evidence that these sensitivities extend beyond the languages discussed here. For example, tense- and (subject) person-based effects have been reported in many Bantu languages, including some that are not so closely-related to the languages discussed (see section 2.3). This suggests that some of the morphosyntactic sensitivities discussed in this work may be broadly

applicable to Bantu languages in general. Likewise, there is also evidence that similar interpretive effects to those discussed extend broadly across Bantu languages, at least to some extent. While overt interpretive contrasts between copular forms are rarely reported in Bantu languages, interpretive effects associated with copular forms can still be found. For example, reflexes of proto-Bantu *\*-de* (e.g., *-li/-ri*) are often reported to serve as dedicated locative copulas (Schneider-Zioga 2018, Gibson et al. 2019). As I argue in chapter 4, this may relate to the more general meaning of *-li/-ri* as a “limited” copula. A much larger typological survey would be necessary to determine the extent to which the observed morphosyntactic and interpretive sensitivities hold across Bantu languages (and beyond). We may find that many of the factors discussed in this work play a role in determining copular form across languages, however it is possible that these factors do not always have the same effect. Moreover, it is possible that there are additional factors relevant to copular form that were not considered in this project. In fact, two other factors are already known to play a role in determining copular form in Bantu languages, namely the polarity of clause (Gibson et al. 2019), and complementizer choice in embedded copular clauses (Gluckman 2024). Further work is needed to determine the full range of factors that are relevant to the form of the copula in Bantu languages — including the Great Lakes Bantu languages discussed in this project.

Another question that arises from the variation data is, How did these patterns develop in Bantu languages? In chapter 5, I speculated that the complex distribution of the copular form *-ri* in the JD60 languages Kinayamulenge and Kirundi (and potentially Kinyarwanda) relates to the general meaning of *-li/-ri* forms (i.e., reflexes of proto-Bantu *\*-de*) and the development of the invariant copula *ni*. Aside from this, I largely sidestep any discussion of the diachronic processes that resulted in the observed patterns of copular variation. That said, it is possible that some of these patterns relate to the interpretive profiles of the copular forms as I propose for JD60 *-ri*. For example, it is feasible that the correlation between tense/mood-environments and specific copular forms stems from the interpretive profile of the relevant copula form. One could imagine that the correlation between *-ba* and irrealis mood relates to its interpretive profile; if *-ba* is used to describe property ascriptions that hold of more general contexts that are not rigidly defined, it is not surprising to

find *-ba* used in hypothetical or future-oriented environments. Likewise, the correlation between *-li/-ri* and past tense could relate to its meaning as a limited copula; if *-li/-ri* is used to describe property descriptions that hold of rigidly defined contexts, it is not surprising to find it in past tense environments, which are temporally-bounded. Once again, further work is needed to determine if such patterns can be derived from the meanings of the different copular forms observed. Ideally a future investigation of tense effects in particular would also provide insight into the effect of aspect, which was largely ignored in this project.

In the bigger picture, the observations and analyses in this work contribute to our understanding of copulas and non-verbal predication, and set the stage for future typological investigations of copular variation in Bantu languages and beyond.



## Bibliography

- Abe, J. (2001). A o b da to omou to a o b to omou [“a o b da to omou” and “a o b to omou”]. *Nihongo to nihon bungaku*, 33, 14–24.
- Abe, J. (2015). Inyoukunai niokeru kopyura no hishutsugen nitsuite: Da to omou to to omou [on the non-occurrence of a copula in a quotative phrase: “da to omou” and “to omou”]. *Bunpou Danwakenkyu to Nihongo Kyouiku no Setten*, (pp. 57–78).
- Ackrill, J. (1963). *Aristotle’s categories and De Intrepretatione*. Oxford University Press.
- Adger, D. (2003). *Core Syntax: A minimalist approach*. Oxford University Press.
- Adger, D. & Ramchand, G. (2003). Predication and equation. *Linguistic Inquiry*, 34(3), 325–359, <https://doi.org/10.1162/002438903322247515>.
- Arche, M. J. (2006). *Individuals in time: Tense, aspect and the individual/stage distinction*. Amsterdam & Philadelphia: John Benjamins Publishing.
- Arche, M. J. (2012). On the aspectuality of the individual-level/stage-level dichotomy. *Borealis: An International Journal of Hispanic Linguistics*, 1(2), 109–132.
- Arche, M. J., Fábregas, A., & Gálvez, R. M. (2019). Main questions in the study of copulas: Categories, structures, and operations. In *The Grammar of Copulas Across Languages* (pp. 1–30).
- Arregi, K. (2000). Tense in basque. Ms. Massachusetts Institute of Technology. <https://home.uchicago.edu/karlos/Arregi-2001-focus.pdf>.
- Arregi, K., Francez, I., & Martinović, M. (2021). Three arguments for an individual concept analysis of specificational sentences. *Natural Language & Linguistic Theory*, 39, 687–708.

- Austin, J. L. (1950). Truth. In *Philosophical Papers* (3rd ed. 1979). Oxford University Press.
- Bach, E. (1967). Have and be in English syntax. *Language*, 43(2), 462–485, <https://doi.org/10.2307/411547>.
- Baker, M. (2003). On the loci of agreement: Inversion constructions in Mapudungun. In *NELS 33: Proceedings of the Thirty-Third Annual Meeting of the North East Linguistic Society* (pp. 25–49).
- Balusu, R. (2014). The overt predicator -gaa in Telugu. Talk at the NYU Syntax Brown Bag.
- Baranova, V. V. & Mishchenko, D. F. (2021). Non-verbal negation markers and the Negative Existential Cycle in Bashkir and Kalmyk with some typological parallels. *The Negative Existential Cycle*, (pp. 403–439).
- Beavers, J., Ponvert, E., & Wechsler, S. (2009). Possession of a controlled substantive: Light ‘have’ and other verbs of possession. In T. Friedman & S. Ito (Eds.), *Proceedings of Semantics and Linguistic Theory 18* (pp. 108–125).
- Béjar, S. & Kahnemuyipour, A. (2017). Non-canonical agreement in copular clauses1. *Journal of linguistics*, 53(3), 463–499.
- Béjar, S. & Rezac, M. (2003). Person Licensing and the Derivation of PCC Effects. *Amsterdam studies in the theory and history of linguistic science series 4*, (pp. 49–62).
- Benmamoun, E. (2000). *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford University Press.
- Birusha, A. (1985). *Description de la langue havu (bantout J52: grammaire et lexique*. PhD thesis, Université libre de Bruxelles.
- Bjorkman, B. (2011). *BE-ing Default: The Morphosyntax of Auxiliaries*. PhD thesis, Massachusetts Institute of Technology.
- Blom, A. & Daalder, S. (1977). *Syntaktische theorie en taalbeschrijving*. Coutinho.

- Bochnak, R., Grinsell, T., & Yu, A. (2011). Copula agreement and the stage-level/individual-level distinction in Washo. In M. J. Stein (Ed.), *Workshop on Structure and Constituency in Languages of the Americas (WSCLA)*, volume 16 (pp. 1–10).
- Borer, H. (2005). *The Normal Course of Events*, volume (Volume II). Oxford University Press.
- Botne, R. (1986). The temporal role of eastern Bantu -ba and-li. *Studies in African Linguistics*, 17(3), 303–318.
- Bowers, J. (1993). The syntax of predication. *Linguistic Inquiry*, 24(4), 591–656.
- Bowers, J. (2001). Predication. *The handbook of contemporary syntactic theory*, (pp. 299–333).
- Bruening, B. (2018). Depictive secondary predicates and small clause approaches to argument structure. *Linguistic Inquiry*, 49(3), 537–559, [https://doi.org/10.1162/ling\\_a\\_00281](https://doi.org/10.1162/ling_a_00281).
- Buell, L. C. & De Dreu, M. (2013). Subject raising in zulu and the nature of predp. *The Linguistic Review*, 30(3), 423–466.
- Camilleri, M. & Sadler, L. (2019). The grammaticalisation of a copula in vernacular Arabic. *Glossa*, 4(1), 1–33.
- Campbell, L. (1985). *The Pipil language of El Salvador*, volume 1. Berlin: Mouton.
- Carlson, G. (1977). *Reference to Kinds*. PhD thesis, University of Massachusetts, Amherst.
- Carnie, A. H. (1995). *Non-verbal predication and head-movement*. PhD thesis, Massachusetts Institute of Technology.
- Carstens, V. (2005). Agree and EPP in Bantu. *Natural Language & Linguistic Theory*, 23, 219–279.
- Chomsky, N. (1986). *Barriers*. Linguistic Inquiry Monograph. Cambridge, MA: MIT Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

- Choueiri, L. (2016). The pronominal copula in arabic. *Brill's Journal of Afroasiatic Languages and Linguistics*, 8(1), 101–135.
- Clarke, E. (1978). Locationals: Existential, locative and possessive constructions. In J. H. Greenberg (Ed.), *Universals of human language* (pp. 85–126). Stanford University Press.
- Clements, J. C. (1988). The semantics and pragmatics of the Spanish <copula + adjective > construction. *Linguistics*, 26(5), 779–822, <https://doi.org/doi:10.1515/ling.1988.26.5.779>.
- Clements, J. C. (2006). Ser-estar in the predicate adjective construction. *Functional approaches to Spanish syntax: Lexical semantics, discourse and transitivity*, (pp. 161–202).
- Comorovski, I. (2007). Constituent questions and the copula of specification. In *Existence: Semantics and syntax* (pp. 49–77). Dordrecht: Springer.
- Curnow, T. J. (2000). Towards a cross-linguistic typology of copula constructions. In *Proceedings of the 1999 Conference of the Australian Linguistic Society*, volume 11 (pp. 203–300). <https://www.als.asn.au/proceedings/als1999/curnow.pdf>.
- de Dreu, M. (2008). The internal structure of the Zulu DP. Master's thesis, Universiteit Leiden.
- den Dikken, M. (1998). Predicate inversion in DP. In A. Alexiadou & C. Wilder (Eds.), *Possessors, Predicates and Movement in the Determiner Phrase* (pp. 177–214). Amsterdam: John Benjamins.
- den Dikken, M. (2006). *Relators and linkers*. Cambridge, MA: MIT Press.
- Deo, A., Sanchez-Alonso, S., & Pinango, M. (2017). Alternative circumstances of evaluation and the *ser/estar* distinction in Spanish. Ms. Yale University. URL: <http://ling.auf.net/lingbuzz/003543>.
- Devitt, D. C. (1994). *Copula construction in cross-linguistic and diachronic perspective*. PhD thesis, State University of New York at Buffalo.

- Devlin, K. (2006). Situation theory and situation semantics. In *Handbook of the History of Logic*, volume 7 (pp. 601–664). Amsterdam: Elsevier.
- Diesing, M. (1990). *The syntactic roots of semantic partition*. PhD thesis, University of Massachusetts Amherst.
- Diesing, M. (1992a). Bare plural subjects and the derivation of logical representations. *Linguistic Inquiry*, 23(3), 353–380.
- Diesing, M. (1992b). *Indefinites*. Cambridge, MA: MIT Press.
- Dik, S. C. (1987). Copula auxiliarization: how and why. *Historical development of auxiliaries*, (35), 53–84.
- Doherty, C. (1996). Clausal structure and the Modern Irish copula. *Natural Language & Linguistic Theory*, 14(1), 1–46.
- Doke, C. M. (1922). *The grammar of the Lamba language*. London: Kegan Paul, Trench, Trübner.
- Eberhard, David M., G. F. S. & Fennig, C. D. (2024). Ethnologue: Languages of the World, nineteenth edition. Online version: <http://www.ethnologue.com>.
- Eid, M. (1983). The copula function of pronouns. *Lingua*, 59(2-3), 197–207.
- Eid, M. (1991). Verbless sentences in Arabic and Hebrew. *Perspectives on Arabic linguistics*, 3, 31–61.
- Finholt, A. (2022). Decomposing HAVE: Evidence from Possessive Predication in Mashi. In *NELS 52: Proceedings of the Fifty-Second Annual Meeting of the North East Linguistic Society*, volume 1.
- Finholt, A. (2023). Interpretive contrasts in the Kifuliiru copular system. *Proceedings of the 41st West Coast Conference on Formal Linguistics*, forthcoming.

- Finholt, A. (2024). Copulas, possession, and the temporary-permanent distinction in Mashi: Evidence for decompositional HAVE. *Glossa: a journal of general linguistics*, 9(1), <https://doi.org/10.16995/glossa.9977>.
- Fischer, G. (2003). The problem is/are your parents: Resolving number conflicts in equative sentences in Dutch and German. Undergraduate Honours dissertation, University of Edinburgh.
- Fortescue, M. D. (1984). *West Greenlandic*. Croom Helm London.
- Frajzyngier, Z. (1986). From preposition to copula. In *Annual Meeting of the Berkeley Linguistics Society*, volume 12 (pp. 371–386).
- Francez, I. & Koontz-Garboden, A. (2017). *Semantics and morphosyntactic variation: Qualities and the grammar of property concepts*. New York: Oxford University Press.
- Freeze, R. (1992). Existentials and other locatives. *Language*, 68, 553–396, <https://doi.org/10.2307/415794>.
- Frege, G. (1897). Logik. In K. F. Hermes, H. & F. Kaulbach (Eds.), *Nachgelassene Schriften (1969)* (pp. 137–163). Hamburg: Felix Meiner.
- Fukui, N. (1986). *A theory of category projection and its applications*. PhD thesis, Massachusetts Institute of Technology.
- Gambarage, J. J. (2019). *Belief-of-Existence Determiners: Evidence from the Syntax and Semantics of Nata Augments*. PhD thesis, University of British Columbia.
- Gatchalian, T. (2023). What’s in a copula? The syntax of Kirundi’s multiple copulas. *Toronto Working Papers in Linguistics*, 46.
- Geist, L. (2007). Predication and equation in copular sentences: Russian vs. English. In *Existence: Semantics and syntax* (pp. 79–105). Dordrecht: Springer.

- Gibson, H. (2012). *Auxiliary placement in Rangi: a dynamic syntax perspective*. PhD thesis, SOAS, University of London.
- Gibson, H., Guérois, R., & Marten, L. (2019). Variation in Bantu copula constructions. In A. Fábregas & R. Marín (Eds.), *The grammar of copulas across languages* (pp. 213–242). Oxford University Press.
- Gluckman, J. (2024). Clausal anchoring in Bantu languages. Ms. University of Kansas.
- Green, L. (2000). Aspectual be-type constructions and coercion in African American English. *Natural Language Semantics*, 8(1), 1–25, <https://doi.org/10.1023/A:1008392112529>.
- Greenberg, Y. (1998). An overt syntactic marker for genericity in Hebrew. In S. Rothstein (Ed.), *Events and Grammar* (pp. 125–143). Dordrecht: Kluwer Academic Publishers.
- Güldemann, T. (2003). Present progressive vis-à-vis predication focus in Bantu: A verbal category between semantics and pragmatics. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, 27(2), 323–360.
- Guthrie, M. (1948). *The Classification of the Bantu Languages*. London/New York: Routledge.
- Guthrie, M. (1967-1971). *Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages, 4 vols.* Gregg International.
- Hale, K. & Keyser, S. J. (2002). *Prolegomen to a Theory of Argument Structure*. Linguistic Inquiry Monograph. Cambridge, MA: MIT Press.
- Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In S. J. K. Kenneth Hale (Ed.), *The View from Building 20*, volume 20 (pp. 111–176). Cambridge, MA: MIT Press.
- Hallman, P. (2022). Head and dependent marking in clausal possession. *Linguistic Inquiry*, 53(3), 551–570, [https://doi.org/10.1162/ling\\_a\\_00416](https://doi.org/10.1162/ling_a_00416).

- Halpert, C. & Diercks, M. (2014). More on have and need. In J. Kandybowicz, T. Major, H. Torrence, & P. T. Duncan (Eds.), *African linguistics on the prairie: selected papers from the 45th Annual Conference on African Linguistics* (pp. 289–305).: Berlin: Language Science Press.
- Harley, H. (1995). *Subject, Events and Licensing*. PhD thesis, Massachusetts Institute of Technology.
- Harley, H. (1998). You're having me on: Aspects of have. In J. Guéron & A. Zribi-Hertz (Eds.), *La grammaire de la possession* (pp. 195–226). Paris: Université Paris X – Nanterre.
- Harley, H. (2005). How do verbs get their names? denominal verbs, manner incorporation and the ontology of verb roots in english. In *The syntax of aspect: Deriving thematic and aspectual interpretation*. Oxford Studies in Theoretical Linguistics.
- Hartmann, J. & Heycock, C. (2016). Evading agreement: A new perspective on low nominative agreement in icelandic. In *NELS 46: Proceedings of the Forty-Sixth Annual Meeting of the North East Linguistic Society* (pp. 67–80).: GLSA Publications.
- Hartmann, J. M. & Heycock, C. (2018). A remark on béjar & kahnemuyipour 2017: Specificational subjects do have phi-features. *Journal of Linguistics*, 54(3), 611–627.
- Haude, K. (2018). Nonverbal predication in movima. In R. V. Simon E. Overall & S. Gildea (Eds.), *Nonverbal predication in Amazonian languages* (pp. 217–244). Amsterdam: John Benjamins.
- Heggie, L. (1988). *The Syntax of Copular Structures*. PhD thesis, University of Southern California.
- Heim, I. & Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell.
- Heller, D. (2005). *Identity and information: Semantic and pragmatic aspects of specificational sentences*. PhD thesis, Rutgers University.
- Hengeveld, K. (1990). A functional analysis of copula constructions in Mandarin Chinese. *Studies in Language*, 14(2), 291–323.



- Hengeveld, K. (1992). *Non-verbal predication: Theory, typology, diachrony*, volume 15. Berlin: De Gruyter.
- Heycock, C. (1994). *Layers of predication: The non-lexical syntax of clauses*. New York: Garland.
- Heycock, C. (2012). Specification, equation, and agreement in copular sentences. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 57(2), 209–240.
- Heycock, C. & Kroch, A. (1998). Inversion and equation in copular sentences. *ZAS Papers in linguistics*, 10, 71–87.
- Heycock, C. & Kroch, A. (1999). Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry*, 30, 365–397.
- Higgins, F. R. (1979). *The Pseudo-Cleft Construction in English*. Routledge.
- Hope, E. (1974). *The deep syntax of Lisu sentences*. Canberra: Australian National University.
- Hulstaert, G. (1965). *Grammaire du Lomongo. Deuxième partie: Morphologie*. Tervuren, BEL: Musée royal de l’Afrique centrale.
- Jerro, K. (2015). Copulas and the semantics of location. In *Proceedings of 15th Meeting of the Texas Linguistic Society* (pp. 91–105).
- Junger, J. (1981). Copula constructions in modern hebrew. *Perspectives on Functional Grammar*, (pp. 117–134).
- Kahn, C. H. (1973). *The verb ‘Be’ in Ancient Greek*. Dordrecht: Reidel.
- Kayne, R. (1984). Principles of particle constructions. In *Grammatical Representations*. Dordrecht: Foris.
- Kayne, R. (1985). L’accord du participe passé en français et en italien. *Modèles linguistiques*, 7, 73–90.

- Kayne, R. S. (1993). Toward a modular theory of auxiliary selection. *Studia Linguistica*, 47(1), 3–31.
- Kiso, A. (2012). *Tense and aspect in Chichewa, Citumbuka and Cisená: A description and comparison of the tense-aspect systems in three southeastern Bantu languages*. PhD thesis, Stockholm University.
- Koopman, H. & Sportiche, D. (1991). The Position of Subjects. *Lingua*, 85, 211–258.
- Kratzer, A. (1995). Individual-level predicates. In G. Carlson & F. J. Pelletier (Eds.), *The generic book* (pp. 125–175). Stanford, CA: CSLI Publications.
- Kratzer, A. (1996). Severing the External Argument from its Verb. In J. Rooryck & L. Zaring (Eds.), *Phrase Structure and the Lexicon*, Studies in Natural Language and Linguistic Theory (pp. 109–138). Dordrecht: Kluwer Academic Publishers.
- Kratzer, A. (1998). Scope or pseudo-scope? Are there wide-scope indefinites? In S. Rothstein (Ed.), *Events in Grammar*. Dordrecht: Kluwer Academic Publishers.
- Kratzer, A. (2005). Building resultatives. In C. Maienborn & A. Wöllstein (Eds.), *Event arguments: Foundations and applications* (pp. 177–212). Max Niemeyer Verlag.
- Kratzer, A. (2023). Situations in Natural Language Semantics. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2023 edition.
- Kuroda, S.-Y. (1988). Whether you agree or not: rough ideas about the comparative grammar of english and japanese. *Lingvisticae Investigatione*, 25, 1–47.
- Lafkioui, M. B., Nshemezimana, E., & Bostoen, K. (2016). Cleft constructions and focus in Kirundi. *Africana Linguistica*, 22, 71–106.
- Laka, I. (1993). The structure of inflection: a case study in x0. *Generative studies in Basque linguistics*, 105, 21.

- Langacker, R. W. (1975). Functional stratigraphy. In *Papers from parasession on functionalism*, volume 351 (pp. 14–307). Chicago: Chicago Linguistic Society.
- Lanham, L. (1953). The copulative construction in bantu with special reference to Zulu. *African Studies*, 12(4), 141–162.
- Lasnik, H. (1981). Restricting the theory of transformations: A case study. In *The Logical Problem of Language Acquisition*. London: Longman.
- Lasnik, H. (1995). Verbal morphology: Syntactic structures meets the minimalist program. *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, (pp. 251–275).
- Levinson, L. (2011). Possessive WITH in Germanic: HAVE and the role of P. *Syntax*, 14(4), 355–393, <https://doi.org/10.1111/j.1467-9612.2011.00159.x>.
- Lohndal, T. (2006). The phrase structure of the copula. *Working papers in Scandinavian syntax*, 78, 37–75.
- Lyons, J. (1968). *Introduction to Theoretical Linguistics*. Cambridge University Press.
- MacFarlane, J. (2014). *Assessment sensitivity: Relative truth and its applications*. OUP Oxford.
- Mahajan, A. (1994). The ergativity parameter: have-be alternation, word order and split ergativity. In *NELS 24*, volume 1 (pp. 22).
- Maho, J. (2009). *NUGL Online: the web version of the new updated Guthrie List, a referential classification of the Bantu languages*. Gothenburg: Department of Oriental and African Languages, University of Gothenburg.
- Maienborn, C. (2005). A discourse-based account of Spanish ser/estar. *Linguistics*, 43(1), 155–180, <https://doi.org/10.1515/ling.2005.43.1.155>.

- Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st annual Penn linguistics colloquium*, volume 4 (pp. 201–225).
- Marantz, A. (2009a). Resultatives and re-resultatives: Direct objects may construct events by themselves. Paper presented at Penn Linguistics Colloquium.
- Marantz, A. (2009b). Roots, re-and affected agents: Can roots pull the agent below little v? Paper presented at Roots workshop, Universität Stuttgart.
- Markman, V. G. (2008a). The case of predicates (revisited): Predicate instrumental in russian and its restrictions. *Journal of Slavic Linguistics*, 16(2), 187–246, <https://doi.org/10.1353/jsl.0.0001>.
- Markman, V. G. (2008b). Pronominal copula constructions are what? reduced specificational pseudo-clefts. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, volume 26 (pp. 366–374).
- Marlo, M. R. (2013). Verb tone in Bantu languages: micro-typological patterns and research methods. *Africana Linguistica*, 19, 137–234.
- Marten, L. (2013). Structure and interpretation in swahili existential constructions. *Italian Journal of Linguistics*, 25, 45–73.
- Matushansky, O. & de Dreu, M. (2009). Some notes on non-verbal predication in Bantu. Paper presented at MUST, Utrecht University.
- McWhorter, J. (1994). From focus marker to copula in Swahili. In *Annual Meeting of the Berkeley Linguistics Society*, volume 20 (pp. 57–66).
- Meeussen, A. E. (1967). Bantu grammatical reconstructions. *Africana linguistica*, 3(1), 79–121.
- Meillet, A. (1912). L'évolution des formes grammaticales. *Scientia*, 6(12).

- Meinhof, C. (1906). *Grundzüge einer vergleichenden Grammatik der Bantusprachen*. D. Reimer (E. Vohsen).
- Merchant, J. (2015). How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry*, 46(2), 273–303, [https://doi.org/10.1162/LING\\_a\\_00182](https://doi.org/10.1162/LING_a_00182).
- Mikkelsen, L. (2005). *Copular Clauses: Specification, Predication, and Equation*. Amsterdam: John Benjamins.
- Mikkelsen, L. (2011). Copular clauses. In *Semantics: An International Handbook of Natural Language Meaning* (pp. 1805–1829). Berlin: De Gruyter.
- Milsark, G. (1974). *Existential Sentences in English*. PhD thesis, Massachusetts Institute of Technology.
- Mithun, M. (2001). *The languages of native North America*. Cambridge University Press.
- Moro, A. (1988). Per una teoria unificata delle frasi copulari. *Rivista di grammatica generativa*, 13(8), 1–110.
- Moro, A. (1990). There-raising: Principles across levels.
- Moro, A. (1995). Small clauses with predicative nominals. In *Small clauses* (pp. 109–132). San Diego: Academic Press.
- Moro, A. (1997). *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge University Press.
- Moro, A. (2000). *Dynamic antisymmetry*. Number 38. Cambridge, MA: MIT Press.
- Myler, N. (2016). *Building and Interpreting Possession Sentences*. Cambridge, MA: MIT Press.
- Myler, N. (2018). Complex copula systems as suppletive allomorphy. *Glossa: a journal of general linguistics*, 3(1), 1–43.

- Nakayama, M. (1988). Empty copulas. In Y. Otsu (Ed.), *Mita working papers in psycholinguistics*, volume 1 (pp. 1–128).: Mita Psycholinguistics Circle, Keio University.
- Nambu, S. (2023). A quantitative study on zero copula in Japanese. *Language Sciences*, 96, 101534.
- Nicolle, S. (2013). *A Grammar of Digo: a Bantu language of Kenya and Tanzania*. Dallas, TX: SIL International.
- Nicolle, S. (2015). Digo narrative discourse. *Language and Culture Documentation and Description*, 26.
- Nishiyama, K. (1999). Adjectives and the copulas in Japanese. *Journal of East Asian Linguistics*, 8(3), 183–222.
- Nurse, D. & Philippson, G. (2006). Common tense-aspect markers in Bantu. *Journal of African Languages and Linguistics*, 27(2), 155–196, <https://doi.org/doi:10.1515/JALL.2006.009> <https://doi.org/10.1515/JALL.2006.009> [cit. 2024-02-13].
- Otterloo, K. V. (2011a). *The Kifuliiru Language, Volume 1: Phonology, Tone, and Morphological Derivation*.
- Otterloo, R. V. (2011b). *The Kifuliiru Language, Volume 2: A Descriptive Grammar*.
- O’Neill, T. (2019). The support copula in the left periphery. In *The Grammar of Copulas Across Languages* (pp. 48–68). Oxford University Press.
- Partee, B. H. (1986). Ambiguous pseudoclefts with unambiguous be. In *NELS 16*, volume 1 (pp. 24).
- Pereltsvaig, A. (2001). *On the nature of intra-clausal relations: A study of copular sentences in Russian and Italian*. PhD thesis, McGill University.
- Persohn, B. (2017). *The verb in Nyakyusa: A focus on tense, aspect, and modality*. Language Science Press.
- Pesetsky, D. (1995). *Zero Syntax*. Cambridge, MA: MIT Press.

- Polack-Bynon, L. (1975). *A Shi grammar*. Tervuren: Musée Royal de l'Afrique Centrale, Annales – Sciences Humaines, No. 86.
- Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of ip. *Linguistic Inquiry*, 20(3), 365–424.
- Posthumus, L. (2000). When does 'become' become 'be'? *South African Journal of African Languages*, 20(4), 342–352.
- Poulos, G. & Msimang, C. (1998). A linguistic study of Zulu. *Via Afrika Limited*.
- Pustet, R. (2003). *Copulas: Universals in the Categorization of the Lexicon*. Oxford University Press.
- Ramchand, G. C. (1997). *Aspect and predication: The semantics of argument structure*. Oxford University Press.
- Ramsey, M. M. (1894). *A Textbook of Modern Spanish*. New York: Spaulding.
- Ritter, E. & Rosen, S. T. (1997). The Function of *have*. *Lingua*, 101, 295–321, [https://doi.org/10.1016/S0024-3841\(96\)00024-1](https://doi.org/10.1016/S0024-3841(96)00024-1).
- Rizzi, L. (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry*, (pp. 501–557).
- Romero, M. (2005). Concealed questions and specificational subjects. *Linguistics and Philosophy*, 28, 687–737.
- Roy, I. (2013). *Nonverbal predication: Copular sentences at the syntax-semantics interface*. Oxford Studies in Theoretical Linguistics. Oxford University Press.
- Russell, B. (1919). *Introduction to Mathematical Philosophy*. London: Allen and Unwin.
- Sangare, S. & Roy, I. (2022). Bambara inverse copular sentences. In *Workshop on copular sentences: Predication, Specification, Equation*.

- Schadeberg, T. C. (1986). Tone cases in Umbundu. *Africana linguistica*, 10(1), 423–447.
- Schlenker, P. (2003). Clausal equations (a note on the connectivity problem). *Natural Language & Linguistic Theory*, 21(1), 157–214.
- Schneider-Zioga, P. (2018). Non-verbal predication in Bantu. *Oxford Guide to the Bantu Languages*, forthcoming.
- Schneider-Zioga, P. (2021). A'-agreement is not only morphological: the case of focus agreement. Paper presented at the Move and Agree Forum, University of British Columbia and McGill University.
- Schneider-Zioga, P. & Mutaka, P. N. (2015a). Copular clauses in Kinande. *The Journal of West African Languages*, 42(2), 57–78.
- Schneider-Zioga, P. & Mutaka, P. N. (2015b). Small clauses and secondary predication in Kinande. *Journal of West African Languages*, 42(2), 101–124.
- Schwarz, F. (2003). Focus marking in Kikuyu. *ZAS papers in linguistics*, 30, 41–118.
- Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. *The handbook of phonological theory*, 1, 550–569.
- Seres, D. & Espinal, M. T. (2019). Russian definitional generic sentences. *Glossa: a journal of general linguistics*, 4(1).
- Shioda, T. (2010). Danuki kotoba?['da'deletion?]. *Online article at NHK Broadcasting Culture Research Institute*.
- Stassen, L. (1997). *Intransitive predication*. Oxford University Press.
- Stechow, A. v. (1996). The different readings of wieder 'again': A structural account. *Journal of semantics*, 13(2), 87–138.
- Stenson, N. J. (1981). *Studies in Irish syntax*. Tübingen: Gunter Narr Verlag.



- Stowell, T. (1981). *Origins of phrase structure*. PhD thesis, Massachusetts Institute of Technology.
- Stowell, T. (1982). The tense of infinitives. *Linguistic Inquiry*, 13(3), 561–570 <http://www.jstor.org/stable/4178293>.
- Stowell, T. (1995). Remarks on clause structure. In *Small clauses* (pp. 109–132). Brill.
- Tanomura, T. (2008). Daikibo-na densisiryoo-ni miru gendainihongo-no dootai. *A Corpus-Based Analysis of Some Time-Related Aspects of Contemporary Japanese* Machikaneyamaronkoo. *Bunkadootaironhen*, 42, 55–77.
- Urushibara, S. (1993). *Syntactic Categories and Extended Projections in Japanese*. PhD thesis, Brandeis University.
- Van der Wal, J. (2006). Predicative tone lowering in makhuwa. *Linguistics in the Netherlands*, 23(1), 224–236.
- Vydrin, V. (2020). Non-verbal predication and copulas in three Mande languages.. *Journal of West African Languages*, 47(1).
- Wald, B. (1973). Syntactic change in the Lake languages of Northeast Bantu. *Studies in African Linguistics; Vol 4, No 3 (1973)*.
- Warotamasikkhadit, U. (1972). *Thai syntax: An outline*, volume 68. Berlin: De Gruyter.
- Welmers, W. E. (1973). *African Language Structures*. University of California Press.
- Williams, E. (1983). Syntactic vs. semantics categories. *Linguistics and Philosophy*, 6(3), 423–446.
- Williams, E. S. (1975). Small clauses in English. In *Syntax and Semantics volume 4* (pp. 249–273). Brill.
- Willms, A. (1972). *Grammatik der südlichen beraberdialekte (SüdMarokko)*. Glückstadt: J.J. Augustin.

- Wongwattana, U. S. (2015). Complexities of Thai copular constructions. *Journal of the Southeast Asian Linguistics Society*, 8, 97–120.
- Wood, J. (2012). *Icelandic morphosyntax and argument structure*. PhD thesis, New York University.
- Wood, J. (2015). *Icelandic Morphosyntax and Argument Structure*, volume 90 of *Studies in Natural Language and Linguistic Theory*. Dordrecht: Springer.
- Zorc, R. D. & Nibagwire, L. (2007). *Kinyarwanda and Kirundi comparative grammar*. Dunwoody Press.

# Appendix

## List of Abbreviations

ACC	accusative case
AG	agentive nominal
AGR	agreement
AUG	augment
CL	classifier
COM	comitative
COMP	complementizer
COMPL	completive aspect
COP	copula
DEM	demonstrative
DEM_DIST	distal demonstrative
DEM_PROX	proximal demonstrative
DEM_REF	referential demonstrative
DPST	distant past tense
FEM	feminine
FOC	focus
FV	final vowel
GEN	genitive case
INDIC	indicative mood
INF	infinitival mood
LNK	linker

LOC	locative
MASC	masculine
MOD	modal
NMLZ	nominalizer
NEG	negation
NOM	nominative case
PERM	permanent
PFV	perfective aspect
PL	plural
POSS	possessive pronoun
PRD	predicative
PRO	pronoun
PRS	present tense
PST	past tense
REF	pronominal copula
REFL	reflexive
RP	relative pronoun
RPST	recent past tense
SBJV	subjunctive mood
SF	short form
SG	singular
SM	subject marker
TEMP	temporary
TOP	topic